The state of evidence for social and emotional learning: A contemporary meta-analysis of universal school-based SEL interventions

14 authors, including:

Christina Cipriano
Yale University
22 PUBLICATIONS   **169** CITATIONS

SEE PROFILE

Almut Zieher
Yale University
8 PUBLICATIONS   **82** CITATIONS

SEE PROFILE

Abigail Eveleigh
New York University
4 PUBLICATIONS   **14** CITATIONS

SEE PROFILE

Melissa Funaro
Yale University
2 PUBLICATIONS   **7** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   Teacher Awareness Survey View project

Project   Child Observational Mindfulness Measure refinement and adaptation View project

CHILD DEVELOPMENT

The state of evidence for social and emotional learning: A contemporary meta-analysis of universal school-based SEL interventions

Christina Cipriano[1] | **Michael J. Strambler[2]** | **Lauren H. Naples[1]** | **Cheyeon Ha[1]** | **Megan Kirk[1]** | **Miranda Wood[1]** | **Kaveri Sehgal[1]** | **Almut K. Zieher[1]** | **Abigail Eveleigh[1]** | **Michael McCarthy[3]** | **Melissa Funaro[4]** | **Annett Ponnock[1]** | **Jason C. Chow[5]** | **Joseph Durlak[6]**

[1]The Child Study Center, Yale School of Medicine, New Haven, Connecticut, USA

[2]The Consultation Center, Yale School of Medicine, New Haven, Connecticut, USA

[3]School of Education and Human Development, University of Rochester

[4]Harvey Cushing/John Hay Whitney Medical Library, Yale University, New Haven, Connecticut, USA

[5]The College of Education, University of Maryland at College Park, College Park, Maryland, USA

[6]Loyola University Chicago, Chicago, Illinois, USA

Correspondence
Christina Cipriano, Yale Child Study Center, Yale School of Medicine, 350 George Street New Haven, CT 06520, USA.
Email: christina.cipriano@yale.edu

Abstract

This article provides a systematic review and meta-analysis of the current evidence for universal school-based (USB) social and emotional learning (SEL) interventions for students in kindergarten through 12th grade available from 2008 through 2020. The sample includes 424 studies from 53 countries, reflecting 252 discrete USB SEL interventions, involving 575,361 students. Results endorsed that, compared to control conditions, students who participate in USB SEL interventions experienced significantly improved skills, attitudes, behaviors, school climate and safety, peer relationships, school functioning, and academic achievement. Significant heterogeneity in USB SEL content, intervention features, context, and implementation quality moderated student experiences and outcomes. Strengths and limitations of this evidence and implications for future USB SEL research, policy, and practice are discussed.

A little more than a decade ago, *Child Development* published the pivotal meta-analysis by Durlak et al. (2011) examining the state of the evidence for universal school-based (USB) social and emotional learning (SEL) interventions that was available through December 31, 2007. USB SEL interventions support the development of intra- and inter- personal skills to promote physical and psychological health for all students in a given school or grade. SEL includes fostering emotional intelligence, behavior regulation, identity formation, and the skills necessary for establishing and maintaining supportive relationships and making empathic and equitable decisions in the best interest of the entire school community (Cipriano et al., In Progress; CASEL, 2020; Greenberg et al., 2017). Durlak et al.'s meta-analysis of 213 USB SEL programs (involving 270,034 students in grades K-12) reported statistically significant improvements in students' social and emotional (SE) skills (standardized mean effect size [ES], Hedges' *g*=0.57), attitudes toward self and others (*g*=0.23), social behaviors (*g*=0.24) and academic achievement (*g*=0.27). In addition, analyses revealed significant reductions in students' conduct problems, such

---

Abbreviations: SE, social and emotional; SEL, social and emotional learning; USB, universal school-based.

as disruptive behaviors ($g=0.22$) and students' emotional distress ($g=0.24$).

Durlak et al.'s positive findings have been followed by substantial changes in educational practice and policy (Dusenbury et al., 2019; Hamilton et al., 2019; Jones et al., 2017; National Commission, 2018; Organization for Economic Cooperation and Development, 2017). At the time this manuscript was prepared, thousands of schools within and outside the United States had implemented USB SEL programs, with more than 70 countries participating in the 2022 International SEL Day (see selday.org), and more than half of US states had at least preschool and early elementary SEL state standards (58%; Dusenbury et al., 2019), an increase of 50% since 2011. Furthermore, federal, state, and local funding are increasing (Congress, 2020; Krachman & LaRocca, 2017; Weissberg et al., 2015). Ten years later, Durlak et al.'s (2011) paper remains the most frequently cited justification for SEL (cited by 10,124 to date; Google Scholar), although it only synthesizes studies through 2007. This paper provides an update of the evidence available for USB SEL programs from January 1, 2008 through December 31, 2020. We begin by describing significant developments in *what* aspects of SEL content and implementation make a difference in *which* outcomes and *who* benefits (or not) from USB SEL interventions.

Recent and relevant reviews of USB SEL

The first generation of reviews and meta-analyses since the 2011 paper concentrated on study outcomes in one or more of the six outcome areas assessed by Durlak and colleagues: skills, attitudes, conduct problems, prosocial behavior, emotional distress, and academic achievement. However, varying inclusion criteria and treatment effects across reviews have contributed to differential interpretations of what constitutes an SEL program and what effects can be expected (Table S1). In Sklad et al., 2012, Sklad and colleagues reviewed 75 studies published since Goleman, 1995 (based on the year that Goleman's *Emotional Intelligence* was published), including only studies that targeted at least one SE skill as defined by the World Health Organization in, 2003, to analyze the main effect of SEL interventions. Four years later, Wiglesworth and colleagues conducted a review of 89 studies based on an adaptation of Denham's (2006) framework of SE competency (see Denham, 2006; Wigelsworth et al., 2016). The following year, Domitrovich et al. (2017) reviewed five meta-analyses of USB SEL programs (inclusive of 300 studies, involving more than 300,000 students) and showed modest promise for promoting positive skills and reducing behavioral risk using a two-domain framing of intrapersonal and interpersonal SE competence, similar to Denham's (2006) framing.

Furthermore, recent studies have noted that students may benefit in related areas, such as developing civic attitudes, forming identity, and improving moral or ethical reasoning. For example, Grant et al.'s (2017) systematic review of 60 evidence-based SEL programs that met requirements under ESSA (2015) between January 1, 2002 and September 22, 2016, posited SEL competencies identified under four domains: intrapersonal competencies, interpersonal competencies, civic attitudes and behaviors, and school climate and safety. That same year, Taylor et al. (2017) reviewed the follow-up effects in all studies included in Durlak et al.'s, 2011 paper, identifying additional indicators of well-being including reduced rates of disciplinary actions, or pregnancies, and increased likelihood of graduation, employment, and overall health.

Concurrent to the recognition that supportive learning environments influence students' academic, personal, and social adjustment, emergent data suggest that the school or classroom climate changes positively following the administration of USB SEL programs (Berg, Osher, Moroney, & Yoder, 2017; National Commission, 2018; Wang et al., 2020). Considerable evidence regarding the dynamic interaction between positive school or classroom climate and student learning, and their personal and social development, has accumulated (Brown et al., 2010; Hamre & Pianta, 2007; Wang et al., 2020). Findings suggest that students in classrooms and schools that participate in SEL programs are more engaged, have better quality relationships with their teachers, and demonstrate increased prosociality (Brown et al., 2010; Hagelskamp et al., 2013). This has resulted in more USB SEL interventions aiming to actively improve climate to support whole student development (Berg, Osher, Moroney, & Yoder, 2017; Lawson et al., 2019; Wang et al., 2020).

The "What" of USB SEL: SEL content for skill development

Despite the volume of reviews, lack of consensus remains regarding the parameters of SEL (Berg, Osher, Same, et al., 2017; Cantor et al., 2019; Elias & Yuan, 2020; Immordino-Yang et al., 2019; Jones et al., 2017; National Commission, 2018; Osher et al., 2016; Schonert-Reichl, 2019). Upwards of 136 SEL frameworks (Grant et al., 2017) comprise more than 700 SEL-related competencies (Jones et al., 2017) that could represent critical content of USB SEL interventions (Dymnicki et al., 2020; Jones et al., 2019). For example, CASEL's five competencies (self-awareness, self-management, social awareness, relationship skills, responsible decision making; CASEL, 2020) are widely regarded as the standard. Lawson et al.'s (2019) review of 14 CASEL-designated programs differentiated these competencies into 12 components and 59 indicators.

Similarly, Jones et al. (2017) organized social and emotional skills and behaviors into three main categories: cognitive regulation, emotional processes, and social/interpersonal skills (Jones et al., 2017; Jones & Bouffard, 2012). Much like Lawson et al., Jones et al. (2017) also analyzed the content of 25 SEL programs to identify the SEL content they covered. The authors added two more categories, character and mindset skills, because related skills did not fit into the three primary categories. Further, Jones and colleagues categorized skills from 40 SEL-related frameworks to develop a taxonomy and nomological network to allow for clearer categorization of content that often uses similar terms to mean different things and different terms to mean similar things (EASEL, n.d.; Jones, n.d.; Jones et al., 2016). The three-category system was insufficient for some SEL-related skills, which resulted in the inclusion of three additional categories: intellectual, ethical, and civic values related to aspects of character education; perspectives sometimes associated with positive psychology and mindfulness (e.g., optimism, gratitude, openness) that are closely related to SEL; as well as aspects of identity formation related to self-knowledge, self-efficacy, and mindset (EASEL, n.d.; Elias & Yuan, 2020; Frank, 2020; Osher, Cantor, et al., 2020; Shankland & Rosset, 2017). The content delivered by SEL programs reflects one type of core component that could drive the effectiveness of programs among many other core components (Dymnicki et al., 2020).

From "What" to "How": advances in USB SEL implementation science

Advances in USB SEL implementation science, including evolved conceptions of SEL competencies and how program implementation and fidelity are associated with student outcomes, further require an update (Durlak, 2015, 2016; Jones et al., 2017; National Commission of Social, Emotional, and Academic Development, 2018). Growing evidence indicates implementation (i.e., how a program or set of practices is delivered), influences the effect of SEL approaches on student outcomes (Abry et al., 2017; Durlak, 2015, 2016; Rimm-Kaufman & Hulleman, 2015). Importantly, features of implementation go beyond content components to include program structure, such as the sequencing of content, as well as aspects of quality, such as fidelity and dosage (Abry et al., 2015; Jones et al., 2017; Low et al., 2016).

Program curriculum structure

Durlak et al. (2011) found that programs that "use a connected and coordinated set of activities…(Sequenced)," "use active forms of learning…(Active)," "have at least one component devoted to developing personal or social

skills…(Focused)," and "target specific SEL skills rather than targeting skills or positive development in general… (Explicit)"—agreed upon components reflected by the acronym SAFE—were more effective than those that did not include these components (p. 410). Similar to the SAFE components examined by Durlak et al. (2011), CASEL noted that evidence-based SEL programs tend to include "free-standing lessons that provide explicit, step-by step instructions…across the five core competency clusters," and focus on aspects of instruction that pertain to the classroom, including integrating skills into academic content and support for development of school policies and structures broadly (CASEL, 2017, p. 2). Jones et al.'s (2017) review of 25 SEL programs identified 10 common program components that echo those of CASEL, including integration of SEL into academics as well as others focusing on the SEL of adults and adaptability of SEL programs. Although investigation of these program features are now more common (Jones et al., 2017; Wanless & Domitrovich, 2015), questions remain regarding whether the number and combination of SEL content areas taught or the sequence of skills (e.g., emotion skills preceding social skills) relate to the effectiveness of USB SEL.

Program implementation

Durlak et al. (2011) found that programs that monitored implementation (i.e., fidelity) were more effective than those that did not. Most measures of implementation consider fidelity, or the extent to which the program was implemented as intended (e.g., lessons were delivered, core principles were implemented) and there is growing evidence supporting the important roles of implementation fidelity and implementation quality (Domitrovich et al., 2019; Granger et al., 2020; Low et al., 2016; Sutherland & McLeod, 2022). Given the substantial growth and focus on implementation, a contemporary meta-analysis is well suited to identify the role of monitoring other aspects of quality implementation, beyond fidelity, that have been associated with improved outcomes, such as the dosage delivered, quality of program delivery, or participant engagement (Low et al., 2016).

From "How" to "for Whom": USB SEL as a lever for educational equity

Concurrent with increasing USB SEL implementation, is the recognition of SEL's role in promoting or detracting from educational equity (Cipriano et al., 2023; Farrington, 2020; Jagers et al., 2019). Students' experience of education and their treatment in society is one of intersectionality (Crenshaw, 1991)—race, class, gender, ethnicity, sexuality, and disability interact to create overlapping and interdependent systems of

advantage and/or disadvantage for students. Although USB SEL has the potential to provide all students with the skills they need to thrive, the discourse of emotions in schools are situated within the politics and power of the education system, wherein significant disparities in SE outcomes between students persist (Cipriano et al., 2023; Hoffman, 2009). Students with disabilities, those growing up in economically disadvantaged communities and those with experiences of trauma have more difficulties in social relationships, recognizing and managing emotions, and awareness of strengths and needs (Cipriano et al., 2023; Daley & McCarthy, 2021)—all skills explicitly targeted by most USB SEL interventions (Grant et al., 2017; Lawson et al., 2019). And, common SEL program features, including teaching perspective taking, responsible decision making, identity development and affirmation, fostering a safe and inclusive school culture, and celebrating differences (Jones et al., 2017; Lawson et al., 2019) are critical for social justice (Farrington, 2020; Wanless & Barnes, 2020).

Yet, significant disparities are apparent in the little evidence we have available for students who have been historically marginalized in education (Cipriano et al., In Progress, 2023; Daley & McCarthy, 2021; Rowe & Trickett, 2018; Taylor et al., 2017). Academics in the field have scrutinized how programs suggest to and educators teach emotion and behavior regulation and how they set expectations of emotion expression, behavior, and compliance to identify how they may be contributing to culturally insensitive and oppressive practices, including the disproportionate use of punitive or exclusionary discipline with students of color and others historically marginalized (Cipriano et al., 2023; Duchesneau, 2020; Elias et al., 2019). Learners vary in how they perceive and navigate learning experiences and can encounter significant barriers structured by educators who have incongruous expectations or social norms about how learning experiences and environments "should" be navigated and how learners "should" behave within them. Identification of learner variability is critical to making the necessary cultural adaptations to ensure that USB SEL is supportive of all learners (Castro-Olivo, 2014; Jagers et al., 2019).

The present study

Our primary aim was to assess if USB SEL interventions for students in grades K-12 enhance student SE competencies and related outcomes. We hypothesized our review would yield significant mean effects across all indicators (Hypothesis 1; Durlak et al., 2011; Rimm-Kaufman & Hulleman, 2015; Taylor et al., 2017). To the extent data have become available since the Durlak et al. (2011), we explored the follow-up effects originally assessed at posttest and all additional long-term indicators of well-being

(Taylor et al., 2017). We hypothesized significant mean effects at follow-up across all outcomes and any additional indicators of wellbeing (Hypothesis 2; Durlak et al., 2011; Rimm-Kaufman & Hulleman, 2015; Taylor et al., 2017).

Further, although the promotion of social and emotional skills is the primary feature of SEL programs to date, no meta-analysis has sought to determine which specific content or content combinations yield the best outcomes. This meta-analysis is the first to address this question and explore if intervention effects are associated with specific skills, attitudes, and/or beliefs targeted by SEL programs, in terms of the number of different skill domains that are targeted, and/or the number of discrete skills that are targeted within any one domain. We expected programs with instruction in more than one skill domain to yield significantly better effects than programs with instruction in only one skill domain (Hypothesis 3a; Abry et al., 2017; Durlak, 2015), and programs that included instruction with greater depth of coverage (higher number of discrete skills within domains) to yield significantly better effects at than programs that instruct fewer targeted skills within domains (Hypothesis 3b; Lawson et al., 2019).

Following Durlak et al.'s (2011) meta-analysis, we examined if intervention effects are moderated by delivery format, inclusion of effective program features (i.e., SAFE practices), the integration of the intervention into academic instruction, or intervention dosage and sequencing. We hypothesized that: program delivery by teachers and multicomponent approaches will yield significantly higher effects than those delivered by outside school personnel or solely classroom-based approaches (Hypothesis 4a; Durlak et al., 2011; Taylor et al., 2017); interventions containing all four SAFE features will yield significantly higher effects than interventions that did not (Hypothesis 4b; Durlak et al.); programs that are sequenced to teach intrapersonal skills before interpersonal skills will yield significantly higher effects than programs not containing this sequence (Hypothesis 4c; Denham et al., 2003; Izard et al., 2001); programs that are integrated into academic instruction will yield significantly higher effects than those that are not (Hypothesis 4d; Abry et al., 2015; CASEL, 2020; Newman & Dusenbury, 2015); and programs that are longer in duration will yield significantly higher effects than those that are shorter (Hypothesis 4e). Lastly, given the increasing recognition that study design features can influence study outcomes, we explored if intervention effects are associated with indicators of study design, or the quality of implementation. We hypothesized that studies with higher quality study designs (evidenced by an index of study quality) will yield significantly higher effects at posttest than those with lower quality designs (Hypothesis 5; Joanna Briggs Institute, 2020). We further hypothesized that studies with higher

quality implementation will yield significantly higher mean effects than lower quality studies (Hypothesis 5b; Durlak, 2015).

Exploratory inquiries

When adequately powered, we investigated if intervention effects differed when studies are implemented in the country where the intervention originated, for programs that include other program features, and for different subgroups of students.

METHOD

This systematic review and meta-analysis followed the contemporary PRISMA guidelines (Page et al., 2021) and adhered to a predetermined, peer-reviewed protocol to strengthen transparency through pre-registration (Nosek et al., 2012). The analytic files reflected in this paper are available on Open Science Framework (OSF) alongside two supplemental papers, data, instructions, and code for analysis to support reproducibility. Table S3 provides a glossary of paper terminology. Here we outline the data collection methods used in the following subsections: inclusion criteria, search strategy, data screening and extraction, and risk for bias.

Inclusion and exclusion criteria

Inclusion criteria

We included published or unpublished reports of universal interventions that target one or more intrapersonal and interpersonal skills (e.g., self-management and communication skills). Furthermore, each study was (a) a randomized or quasi-experimental design, (b) contained a control group (wait list, or attention placebo), (c) involved at least six sessions if it was classroom-based or lasted for at least 4 months if a whole school approach; (d) appeared in English in any country between January 1, 2008 and December 31, 2020 (including online first publications), and for the meta-analysis (e) allowed for the calculation of effect sizes by data in print or provided by authors.

Exclusion criteria

Studies comparing two interventions without a control, the SEL intervention was not the primary component, those delivered only to specific students within the classrooms or delivered to students removed from class, were excluded. We also excluded studies whose primary purpose was to promote academic achievement through a specific instructional strategy, as well as sex education and interventions to prevent drug use, AIDS, or obesity.

Search strategy

We conducted a systematic literature search with terms developed through an analysis of the controlled vocabulary terms of known key articles and through scoping searches in each database (Table S4). Our approach used an iterative process of translating and refining the searches. To maximize sensitivity, our formal search used controlled vocabulary terms and synonymous free-text words to capture the concepts of "SEL programs" and "school." The search strategy was peer reviewed using the Peer Review of Electronic Search Strategies (PRESS) Checklist (McGowan et al., 2016).

A comprehensive search of multiple databases was then performed by an experienced university librarian of APA Psycinfo (Ovid), MEDLINE (Ovid), Education Resources Information Center (ERIC), Dissertations & Theses Global (ProQuest) and Web of Science databases. Our initial search returned 41,002 studies, and after removing duplicates (9745), advanced 31,257 studies forward to title and abstract screening (Figure 1).

To mitigate risk for publication bias, we followed Higgins et al. (2011) and manually searched for unpublished studies. We (a) searched the American Educational Research Association's (AERA) database, (b) put out a call for unpublished studies on the AERA Social Emotional Learning Special Interest Group and Cognitive Development Society list-servs, and (c) searched three prominent repositories of unpublished and published papers of SEL interventions: CASEL's (2013) guide *Effective Social and Emotional Learning Programs: Preschool and Elementary School Edition*, CASEL's (2015) updated list of interventions to the 2013 report, and the RAND report *Social and Emotional Learning Interventions Under the Every Student Succeeds Act* (Grant et al., 2017). These steps did not yield any additional articles beyond the initial electronic search (*n* = 41,002).

Data screening

Search results (*n* = 41,002) were pooled in EndNote [www.endnote.com] to remove duplicates (*n* = 9745) and then the final set of articles were uploaded (*n* = 31,257) to Covidence [www.covidence.org] for screening. Articles were double screened: first title and abstracts, then full-text. Conflicts were resolved through discussion and consensus. To ensure reliability on study exclusion, two authors independently reviewed all excluded studies and resolved any conflicts by discussion. If there was not sufficient detail in the title and abstract to determine inclusion or exclusion, the article was assessed against the inclusion criteria
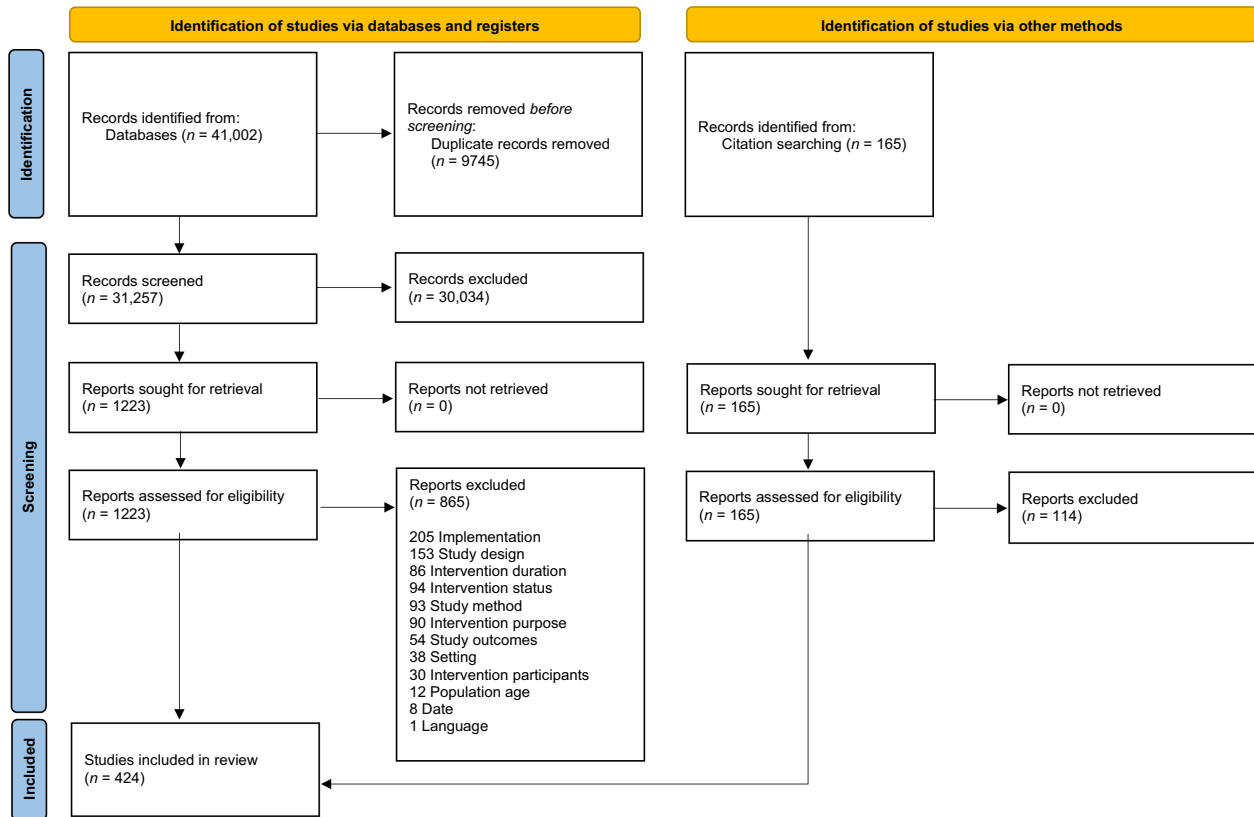
FIGURE 1   PRISMA diagram of included studies.

in the full-text stage of screening. Once abstract and title screening was complete, the review team transitioned to full-text review of eligible articles ($n = 1223$). To ensure reliability on study exclusion at this phase, all articles were independently reviewed by two of the authors, and conflicts were resolved through discussion. Authors reached "almost perfect" interrater reliability for both stages of screening (0.93%–1% for title and abstract, and 0.81%–0.98% agreement for full text; McHugh, 2012), resulting in 382 studies screened for inclusion.

Finally, we contacted all first authors of the 382 studies identified through the search to inquire for additional studies while four researchers conducted ancestral searches of the references lists of the 382 studies that met eligibility for inclusion (searches for articles that had cited these studies that could meet criteria for inclusion in the present review) using Web of Science and Google Scholar tools. Through backward and forward searching, a potential 165 studies were identified for review. After removing duplicates and double screening, 42 studies met the criteria for inclusion for extraction, bringing our total to 424 studies (see Figure 1).

Coding procedures

A team of twelve coders (including two undergraduates, four graduate students, and six of the authors) reviewed articles using the codebook (Supporting Information File A) to code studies in formatted Qualtrics forms. We extracted relevant information from the articles to form our data set. Of note, codebook indicators had an 'other' open text option whenever possible, and extractors manually cataloged the study description of an indicator. These open text responses were aggregated into categories by the authors in the final data set. Coded dataset is available [insert here] and full raw data are available from the first author upon request.

SEL intervention content

Discrete SEL skills were coded individually and clustered into four contemporary SEL content frameworks. Discrete skill indicators were derived from adaptations to the EASEL Taxonomy project (EASEL, n.d.) and integration of indicators reflecting the CASEL Competency Areas identified by Lawson et al. (2019). The Two-Domain Framework classifies discrete skills into *intrapersonal* processes, such as emotion regulation, and *interpersonal* processes such as conflict resolution. For the CASEL Competency Areas, discrete skills are classified into five content areas: *self-awareness*, which includes skills related to identifying one's own feelings, understanding how emotions relate to thoughts and behaviors, as well as aspects of self-knowledge and

acceptance; *self-management*, which includes skills related to coping, setting and reaching goals, and focusing one's attention; *social awareness*, which includes skills related to identifying others' emotions, perspective taking and empathy, and embracing diversity; *relationship skills*, which encompass social skills broadly, including assertiveness; and *responsible decision-making*, which primarily reflects problem-solving skills (Lawson et al., 2019). The Framework for Social and Emotional Learning (Jones & Bouffard, 2012) reflects: *cognitive regulation*, including skills related to cognition and planning, such as attention and critical thinking skills; *emotional processes*, including activities that focus on developing knowledge and skills related to defining emotions and their expression and regulation, as well as empathy and perspective taking skills; and *social/interpersonal skills* that primary focus on developing students' prosocial and cooperative behavior, including understanding social cues and resolving conflicts (Jones et al., 2017; Jones & Bouffard, 2012). The EASEL Taxonomy includes these plus three supplemental categories (EASEL, n.d.; Jones et al., 2017): *values* instruction related to ethical, performance, civic and intellectual values; activities that focus on *perspectives*, such as optimism, gratitude, and openness; and *identity* instruction to develop self-knowledge, a sense of purpose, self-efficacy, and self-esteem (EASEL, n.d.). From the discrete codes, we compiled the *skill frequency* and *skill domains*.

Intervention characteristics

We coded for *primary* and *secondary change agents* (who primarily delivered the intervention). We captured dosage with the *intervention duration* (in sessions), *average session duration* (in minutes), *number of discrete sessions*, and if the intervention lasted for *more than one school year*. We coded for *features of effective programs* (i.e., SAFE: sequenced, active, focused, and explicit; Durlak et al., 2011; Taylor et al., 2017), including the *order of intrapersonal and interpersonal skills* (no order, intrapersonal skills precede, interpersonal skills precede). We coded for many other program features, including *classroom-based, free-standing lessons, integration of SEL into academic instruction, classroom activities beyond core lessons, multi-component, efforts to improve classroom or school climate, family engagement* and *types of family engagement, community engagement* and *types of community engagement, focus on adult social and emotional competence, multi-phased, tiered*, and *adaptability to context*.

Quality

We assessed three categories of quality: study design, study quality, and study implementation quality. Study design variables included *report format*, and an examination of risk of bias using the adapted Joanna Briggs Institute Critical Appraisal tool (2020) for quasi-experimental and randomized controlled trials. Study quality included the *control/comparison condition*, and if there *was more than one control comparison group*, as well as *missing data handling*, and if the *assumption of independence* was met (Carbonneau et al., 2013). Implementation quality included *steps to increase effective program implementation* (e.g., pre-program training, ongoing support or training, supervisory or support meetings, detailed program manuals), *data on the level of implementation* and *level of implementation achieved* (i.e., not mentioned, satisfactory, low, moderate, high, variable), whether *aspects of implementation* were measured (e.g., fidelity, dosage, quality of delivery, student engagement) and what methods were used to assess program implementation. Each of these domains also included an indicator for a combination or other unlisted outcomes. We double coded each indicator and merged codes into a composite index score for analyses (Kremer et al., 2015; Lester et al., 2020).

Sample ecological variables

Sample ecological variables referred to setting and participant characteristics. Setting characteristics included the *country* where the study was conducted, the *type of community* the intervention was conducted in (urban, suburban, rural, mixed), and the *type of school* (public, private [secular, religious], charter, magnet, or other) and *average school size* (large, midsized, small, or numeric) of the schools the intervention was conducted in. For participant characteristics, we coded for participants' mean *age, age range, developmental level (childhood, early adolescence, adolescence), SES reporting* and overall *SES level, percentage female*, reporting of *LGBTQIA+ identification* (i.e., sexuality-related minority status or related metric), *race/ethnicity, language status*, and constructs (or any risk designation) related to *disability*.

Dependent variables: Intervention effects

The dependent variable categories included: (a) *SEL skills*, (b) *attitudes/beliefs*, (c) *prosocial behaviors*, (d) *externalizing behaviors*, (e) *civic attitudes and behaviors*, (f) *peer relationships*, (g) *emotional distress*, (h) *school functioning*, (i) *disciplinary outcomes*, (j) *school climate and safety*, (k) *family relationships*, and (l) *physical health outcomes*. Each domain includes an indicator for other outcomes that were identified as each presented (File B).

Given the volume of data, three teams concurrently coded for general, intervention, and effects

information, respectively. All studies were double coded and due to wide disparities in SEL science and reporting, more than half of the studies were triple coded to estimate coding reliability. Further, 116 corresponding authors were contacted to provide additional data required for the meta-analysis. Of these, 90 (77.6%) authors did not respond to our request, and 20 (17.2%) email addresses were undeliverable. A total of six (5.2%) authors responded and three (2.6%) provided sufficient data to be included in the meta-analysis. To ensure coder consistency, we conducted calibration exercises before review and weekly throughout full-text coding, which lasted 9 months. Discrepancies were resolved through discussion between pairs of coders, and a third coder arbitrated when needed. Pairs independently coded a randomly selected 25% sample of the studies to estimate reliability of coding (Durlak et al., 2011). Kappa coefficients corrected for chance agreement were acceptable across all codes reported in this review (mean kappa was 0.89). Raters' agreements on continuous variables were strong, all above 0.90, and agreement on categorical variables ranged from .78–.92, respectively.

Analytic strategy

Power analysis

An a priori power analysis with a calculator (Quintana & Tiebel, 2019) was conducted based on equations from Valentine et al. (2010) using conservative values from Durlak et al. (2011) of (a) anticipated summary effect size, (b) average group sample size, and (c) number of effect sizes. Specifically, we set a summary effect size of 0.20, which was the lowest found by Durlak et al. (2011). We chose a value of 150 for the average group sample size which is a reasonable estimate from Durlak et al. (2011). Regarding number of effect sizes, the lowest number of studies across the outcomes examined by Durlak and colleagues was $N=3$ for academic outcomes among SEL programs delivered by non-school personnel; $N=7$ for the SEL skills outcome among multicomponent SEL programs, and $N=10$ for academic outcomes among SEL programs delivered by teachers. However, the remaining 21 mean effects had substantially higher sample sizes, ranging from $N=11$ to $N=112$. Thus, we examined power for a range of $N=7$, $N=10$, and $N=12$. All three power analyses assume high heterogeneity of $I^2=0.75$. With these conservative assumptions, power was 0.63 for $N=7$, 0.78 for $N=10$, and 0.91 for $N=15$. These estimates provided guidance regarding analyses involving smaller cells, for which we adjusted in our discussion relative to the final number of observed studies synthesized.

Effect size calculation

While our index of effect was Hedges' g we first calculated Cohen's d using available information from each report. When means and standard deviations were not presented, effect sizes were calculated using other information (e.g., F or t test values) following Borenstein et al. (2009). If an effect size could not be calculated from data in the report, we contacted authors twice for the necessary information. Whenever possible, we calculated adjusted post and follow-up effect sizes using any pre-intervention differences that occurred between intervention and control groups. This procedure improves the precision of intervention effect estimates over time and all results reported herein are based on this measure. Following our calculation of all Cohen's d and adjusted effects sizes, we used the equation presented in Lakens (2013; equation #4) to covert Cohen's d values to Hedges' g effect sizes. We also calculated the variances and standard errors of effect sizes to use in subsequent analyses. We identified effect size outliers as effect sizes that were more than two standard deviations (SD) on either side of the grand mean effect size. There were 76 effect sizes that were more than 2 SDs on the right tail of the distribution and 1 effect size that was outside of 2 SDs on the left side of the distribution. We reset these outliers to be exactly at the 2 SD points to guard against them unduly affecting results. Hedges' g variances and SE were then recalculated for these cases.

Post and follow-up meta-analyses

To estimate mean effects, we conducted Correlated and Hierarchical Effects (CHE) models (Pustejovsky & Tipton, 2022) with Robust Variance Estimation (RVE). CHE models allow for combining effect sizes from the same study into clusters (the "hierarchical" aspect of the approach) and accounts for the correlated nature of the effect size sampling error for effect sizes drawn from the same study (Harrer et al., 2021). This part of the analysis requires us to define ρ, or the assumed mean correlation within and across studies. To be conservative, we assigned a high value of ρ = .60 for post and follow-up meta-analyses an assumed correlation within and between studies. We used the "rma.mv" command in the "metafor" package for R to conduct the CHE models for each outcome domain. We then applied a sandwich estimator using the "clubSandwich" package (Pustejovsky, 2022) to employ RVE for obtaining robust confidence intervals and p-values (Hedges et al., 2010; Tipton, 2015; Tipton & Pustejovsky, 2015). Lastly, we examined if overrepresentation of certain interventions influenced effect sizes (see Table S11); the most frequently represented were *Incredible Years* (10) and *PATHS* (8) and 22 interventions appear more than three times.

Moderator analyses

We assessed whether program effects varied by intervention features by conducting moderation analyses for each of the ten moderators indicated in Table S18. We took a similar approach to the mean effects approach of conducting CHE models with RVE and in this case, we entered dummy-coded moderators as predictors. We used the same assumption of $\rho = .60$ for all analyses, and to assess the robustness of our analyses we conducted sensitivity analyses for SEL skills (the outcome with the largest sample size) where we also examined outcomes at $\rho = .40$ and $\rho = .80$ (Tanner-Smith & Tipton, 2014).

Publication bias

Lastly, we assessed the extent publication bias may have impacted findings through reviewing funnel plots, contour-enhanced funnel plots, and Egger's regression tests where we accounted for the nested structure of the data (Egger et al., 1997). Further, given evidence that average effect sizes differ based on the nature of a publication (Chow & Ekholm, 2018; Polanin et al., 2016), we compared overall effect sizes for published to unpublished studies.

In sum, we followed the PRISMA framework and methodological guidance for research in education and psychology to report findings and reach conclusions (Page et al., 2021; Pigott & Polanin, 2020; Polanin et al., 2020). These practices of data treatment allowed us to build upon a comprehensive catalog of data for review, and reduced researcher bias while maintaining methodological transparency (Pigott & Polanin, 2020; Rodgers & Pustejovsky, 2021).

RESULTS

Descriptive characteristics of reviewed studies

We systematically reviewed 424 studies that involved 575,361 students. Table 1 summarizes investigation features. The sample years ranged from 2008 through 2020, and 47.1% of the sample reflect the last 5 years, with 2020 most frequently represented at 62 studies (14.6%). Studies included 355 peer-reviewed published articles (83.7%), 51 unpublished theses (12.0%), 12 technical reports/program evaluations (2.8%), 3 conference papers (0.7%), and 3 books or chapters (0.7%). Studies represent 196 discrete outlets (Table S6), 178 studies did not report funding and 36.9% reported Federal and other sources (Table 1).

Studies reflect 53 identified countries, and 233 studies (55%) were conducted in the same country where the SEL intervention was developed (see Table S7). Students ranged from 5 to 17 years of age (Table S8).

TABLE 1 Descriptive characteristics of 424 studies of universal school-based SEL interventions.

General study features	N	%
Date of report		
2008–2011	89	21.0
2012–2015	135	31.8
2016–2019	138	32.5
2020	62	14.6
Source of report		
Published article	355	83.7
Dissertation/master's thesis	51	12.1
Technical report, program evaluation	12	2.8
Conference paper	3	0.7
Book or book chapter	3	0.7
Dissemination outlets	196	
Peer reviewed		
Yes	356	84.0
No	68	16.0
Funding source		
No funding indicated	178	42.0
Federal	154	36.9
Private foundation	53	12.5
State	16	3.7
University grant	23	5.5
Study design		
Randomized control trial	237	55.9
Quasi-experimental design	187	44.1
Reported implementation data		
Yes	189	44.6
Level of implementation		
Satisfactory	19	4.4
Low	2	0.5
Moderate	13	3.1
High	76	17.9
Variable across sites	33	7.8
Sample ecology		
Country	53	
Intervention studied in the same country where intervention was developed	233	55.0
Age of students at pre		
Elementary School (Kindergarten/ 5 yrs old—5th grade/10 years old)	241	56.8
Middle School (6th grade / 11 yrs old-8th grade / 13 years old)	104	24.5
High School (9th grade / 14 years old 12th grade / 17 years old)	79	18.6

(Continues)

TABLE 1   (Continued)

General study features	N	%
Types of schools		
Unknown	217	51.2
Public	164	38.7
Private	17	4.0
Charter, Alternative, Technical, Community, Government	14	3.2
School setting		
Unknown	217	51.2
Urban, not otherwise clarified	76	17.9
Urban, inner-city, low-income area	21	5.0
Suburban	30	7.1
Rural	19	4.5
Combination of settings	61	14.3
Average school size		
Unknown	350	82.5

More than half the studies did not report the type of school (e.g., public, 51.2%) nor setting in which the intervention was conducted (51.2%). Among those that did, most were in public schools (164; 38.7%), urban (97; 22.9%) or a combination of urban, suburban, or rural environments (59; 13.9%), and most (350; 82.5%) did not report the average school size of the intervention sample (Table 2).

There was limited and variable representation in how studies reported student demographic characteristics (Table S9). Two hundred and sixteen studies (50.9%) reported the race/ethnicity of students, 347 studies reported student gender (81.9%), with only two making mention of student gender non-binarily (Mogro-Wilson & Tredinnick, 2020; Silverstone et al., 2017), and no studies reported on student sexual identity. Fifty-five studies (13.0%) reported on student's native language status, with 8 (1.9%) excluding participation in the study based on student or guardian's language status. Further, 3 studies reported actively excluding students based on a specific special education status (0.7%), and 70 studies (16.5%) mentioned students with disabilities in their study, with mostly general identifiers such as Special Education Needs, Individual Education Plan, and students with disabilities (15%) and limited mention of disability classifications including Learning, Emotional or Behavioral, Physical, or Developmental Disability, and even fewer specific diagnoses (<2%). One hundred seventy-three studies reported student socio economic status (SES, 40.8%). SES was reported using categorical (low, high-income, good or bad finances) and

TABLE 2   Descriptive characteristics of the 252 universal school-based SEL interventions.

	N	%
Discrete SEL interventions		
Named	210	90.1
Unnamed	42	9.9
Reflected more than once in study	57	
Country intervention was developed in	35	
Based on theory	237	55.9
Primary change agent		
Teachers/aides	290	68.4
School staff other than teachers/aides (school mental health professionals)	30	7.1
Researcher and/or their staff	75	17.7
Entire school staff (this was a whole school program)	5	1.2
Community members or parents	3	0.7
Other/Unknown	21	5.0
Intervention duration		
Up to half a school year (4 months)	190	44.8
Up to 1 whole school year	161	38.0
More than 1 whole school year	63	14.9
Features of effective programs		
SAFE (Sequenced, active, focused, explicit)	267	62.7
Free-standing	353	83.2
Classroom-based	407	96.0
Beyond the lesson	220	51.9
Supports school climate	285	67.2
Supports SEL integration in academics	151	35.6
Tiered intervention	145	34.2
Multi-phased implementation	271	63.9
Multi-component intervention	160	37.7
Instructs skills in an order	365	86.1
Supports adult SEL	99	23.3
Family engagement		
No	173	40.8
Letters/emails sent home	175	41.3
HW	21	5.0
Explicit training+	29	6.8
Community engagement		
No	362	85.4
Service-Learning projects	39	9.2
Community volunteers+	21	4.9
Other	2	0.5

TABLE 2 (Continued)

	N	%
Steps to increase program implementation		
Pre-program training was provided	41	9.7
Ongoing, support or training was provided once the program began	2	0.5
Detailed program manuals or protocols provided to guide providers	8	1.9
Combination of above (list)	309	72.7
Other	14	3.3
None of these steps or strategies	45	10.6
SEL skills and strategies reflected in interventions		
Two domain framework	424	100.0
Interpersonal skills	350	82.5
Intrapersonal skills	341	80.4
CASEL 5 competency framework	411	96.9
Self-awareness	312	73.6
Self-management	349	82.3
Social awareness	273	64.4
Relationship skills	318	75.0
Responsible decision making	335	79.0
Framework for social and emotional learning	401	94.6
EASEL taxonomy	412	97.2
Cognitive regulation	340	80.2
Emotional processes	348	82.1
Social/interpersonal skills	341	80.4
Values	210	49.5
Perspectives	128	30.2
Identity	145	34.1

numerical (free or reduced-price lunch (FRPL), aid to families, household income, disadvantaged status, poverty line) constructs.

Studies were roughly split between randomized control trials (237; 55.9%) and quasi-experimental designs (187; 44.1%), and most studies used an education as usual or waitlist control (381; 72.9%). Only 44.6% reported any implementation data, with 21% of those reporting the level of implementation as Moderate or High. The studies analyzed reflect 2049 SEL outcomes. Just over half of the outcomes reported across studies were reported by children (53.5%, 1185), followed in frequency by teachers (305, 13%) and caregivers (120, 5.4%), among others. Outcome measures spanned the 12 domains of interest in frequency. Most outcomes used rating scales, checklists, or questionnaires (1302, 58.7%). See Table S10.

Descriptive characteristics of reviewed SEL interventions

We identified 252 different SEL interventions, 210 named and 42 unnamed, in our review. Fifty-seven interventions are in the sample more than once, and 14 interventions were in the sample 5 of more times: *Second Step* (16), *Positive Action* (15), *Promoting Alternative Thinking Strategies* (PATHS; 12), the *Good Behavior Game* (12), *Incredible Years* (10), *INSIGHTS* (9), *Strong Kids* (7), *RULER* (6), *Positive Adolescent Training through Holistic Social Programmes* (P.A.T.H.S., 6), *FRIENDS* (6), *Aussie Optimism* (6) *Zippy's Friends* (5), the *4Rs Program* (5), and *Olweus Bullying Prevention Program* (5). See Table S11. Interventions were developed in 35 countries. The most frequently reported countries of development were the United States (46.9%), Australia (6.4%), Spain (2.6%), England (1.9%), Turkey (1.7%), Canada (1.7%), Hong Kong (1.7%), and Norway (1.7%). See Table S7. Across the entire sample, 160 studies (37.7%) explicitly identified the intervention with the term SEL In the study, with more than half of these studies in the last 5 years (57.7%). The most frequent terms other than SEL used to described interventions present in this sample included *Prevention Programs* (64; 15.1%), *Mindfulness* (18; 4.2%), *Classroom Management Programs* and *Positive Youth Development Programs* (11; 2.6%, respectively). Slightly more than half of studies reported being based on an explicit theory or set of theories, with many using Social Cognitive Theory, mindfulness frameworks (general mindfulness, Mindfulness Based Cognitive Therapy [MBCT], and/or Mindfulness-Based Stress Reduction [MBSR]), Cognitive Behavior Therapy (CBT), Emotional Intelligence, Theory of Triadic Influence, and Temperament, among others. Interventions were primarily delivered by teachers/aides (290; 68.4%), researchers (75; 17.7%), and school staff other than teachers/aides (e.g., mental health professionals; 30; 7.1%). Interventions lasted from up to half a school year (190; 44.8%), up to one whole school year (161; 38.0%), or more than 1 year (63; 14.9%). Lessons within interventions ranged in duration from 5 to 120 minutes per session: childhood (grades K-5; $M=43.88$ minutes, SD=24.63), early adolescence (grades 6–8; $M=51.06$ minutes, SD=16.08) and adolescence (grades 9–12; $M=55.83$ minutes; SD=22.87). Intervention length was from 6 to 180 sessions: childhood ($M=30.72$, SD=35.08), early adolescence ($M=19.04$, SD=17.18) and adolescence ($M=19.58$; SD=26.27). See Table S12.

Regarding the features of effective programs (CASEL, 2015; Durlak et al., 2011), 267 studies met the criteria as SAFE (62.7%). Most interventions were classroom-based (407, 96%), included free-standing lessons (353, 83.2%), and more than two thirds focused on improving school climate (285, 67.2%). Two hundred twenty interventions provided content beyond the core lessons (51.9%), 160

TABLE 3   Meta-analytic results of mean effect sizes (Hedges' g) and 95% confidence intervals at post-intervention between intervention and control conditions (Robust random-effect model by study).

		Overall	1. SEL skills	2. Attitudes/ beliefs	3. Prosocial behaviors	4. Externalizing behaviors
Total sample with clustering by paper	ES	0.194	0.219	0.209	0.178	0.162
	95% CI	[0.166, 0.221]	[0.171, 0.267]	[0.160, 0.258]	[0.126, 0.229]	[0.121, 0.204]
	p	<.001***	<.001***	<.001***	<.001***	<.001***
	k	1862	322	220	199	367
	N	258	114	87	89	124
	Q (df)	27101.93 (1861)***	4405.353 (321)***	2163.912 (219)***	1780.063 (198)***	4885.813 (366)***
	Egger (df)	3.59 (1860)	4.10 (320)	1.82 (218)	3.30 (197)	0.63 (365)
	p	<.001***	<.001***	.070	.001**	.531

Note: To estimate the overall effect size across 12 outcome domains, in this table, domains 4, 7, and 9 (Externalizing Behaviors, Emotional Distress, and Disciplinary Outcomes, respectively) were reported with converted positive scores of ES (*g*). The outcome 'Academic Achievement' was also evaluated as a subset of Domain 8 (School Functioning) Hedges' $g=0.112$; 95% CI=[0.038, 0.185]; $p=.004$.

*p<.05; **p<.01; ***p<.001, and marginal ᵀp<.10.

(37.7%) were multi-component programs, and 145 (34.2%) offered tiered programming. Ninety-nine (23.3%) focused on Adult SEL and 151 (35.8%) explicitly integrated SEL into academic instruction. Regarding family engagement, 173 interventions reported no family component (40.8%). Among those that did, the most frequent mode was explicit training to families (29, 6.9%) or engaging families with materials at home (21, 5.0%), Regarding community involvement, the majority of studies did not have a community engagement component (362, 85.4%), and those that did reported service-learning projects (39, 9.2%) or community volunteering (21, 4.9%). Most interventions (309, 72.7%) included tactics to increase program implementation including pre-program training, ongoing training, and detailed manuals, with 45 studies (10.6%) endorsing none of these tactics.

Most interventions were sequenced (311, 73.4%), and instructed intrapersonal skills before interpersonal skills (86.3%), and the content of the interventions ranged widely. Three-hundred thirty-eight (79%) studies endorsed both dimensions of the *Two-Domain Framework of SEL* (Denham, 2006), with 341 studies endorsing *intrapersonal skills* (80.4%) and 351 studies endorsing *interpersonal skills* (82.5%). Four hundred and one studies endorsed at least one of the dimensions of *the Framework for Social and Emotional Learning* (94.6%; Jones & Bouffard, 2012), with 340 studies (80.2%) endorsing *cognitive regulation*, 348 (82.1%) endorsing *emotional processes*, and 341 (80.4%) endorsing *social/interpersonal skills*. The EASEL Taxonomy includes three supplemental categories based on theoretical SEL frameworks and content (Jones et al., 2017; EASEL, n.d.), and activities that focus on *values* were endorsed in 210 studies (49.5%), activities that focus on *perspectives* were endorsed in 125 studies (30.2%), and *identity* instruction was endorsed in 145 studies (34.1%). For the *CASEL Competency Areas* (CASEL, n.d.; Lawson et al., 2019), 411 studies endorsed

at least one CASEL Competency (96.9%). Three hundred twelve studies (73.6%) endorsed *self-awareness*, 349 (82.3%) endorsed *self-management*, 273 (64.4%) endorsed *social awareness*, 318 (75.0%) endorsed *relationship skills*, and 335 (79.0%) endorsed *responsible decision-making*. See Table S13.

Meta-analytic sample

Two hundred fifty-eight (60.8%) studies provided sufficient statistical data to be included in the meta-analytic review. Studies were excluded if (a) there was insufficient data (e.g., qualitative data only, missing sample sizes, presented inappropriate statistics [e.g., median/range, structural equation modeling]) required for meta-analysis (136, 32.1%), (b) reported either follow-up or baseline statistics only with no post-intervention data (14, 3.3%), (c) had questionable or unclear reporting of statistics (e.g., aggregated intervention and control group data, typos/errors in labelling) that made interpretation difficult (8, 1.9%), (d) did not report child-level outcomes (7, 1.7%), or was a protocol paper of an RCT with no statistical results (1, 0.2%). The meta-analytic sample did not differ from the full sample in general study features, sample ecology, or USB SEL intervention features (See Table S5 for comparison).

Do SEL programs significantly improve students' skills, attitudes, and behaviors?

To begin the analysis of mean effects we assessed the degree of variance at each of the level of the three-level model. The estimated variance components were $\tau^2_{Level\ 3}=.029$ and $\tau^2_{Level\ 2}=.046$. This translates to $I^2_{Level\ 3}=57.58\%$ of the total variation being attributed

5. Civic attitudes /behaviors	6. Peer relationship	7. Emotional distress	8. School functioning	9. Disciplinary outcomes	10. School climate/ safety	11. Family relationships	12. Physical health
0.255	0.222	0.140	0.122	0.183	0.293	0.061	0.160
[0.043, 0.466]	[0.143, 0.301]	[0.103, 0.177]	[0.065, 0.178]	[−0.237, 0.602]	[0.198, 0.388]	[−0.034, 0.155]	[−0.014, 0.333]
.022*	<.001***	<.001***	<.001***	.328	<.001***	.177	.068^T
15	84	305	151	11	116	21	51
13	55	121	62	7	43	13	16
76.374 (14)***	632.935 (83)***	1289.242 (304)***	838.517 (150)***	1825.934 (10)***	2635.581 (115)***	75.010 (20)***	1647.184 (50)***
0.27 (13)	0.96 (82)	1.38 (303)	1.83 (149)	−0.33 (9)	4.13 (114)	1.58 (19)	1.76 (49)
.794	.341	.169	.070	.753	<.001***	.130	085

to between-cluster (study), and $I^2_{\text{Level 2}} = 36.11\%$ being attributed to within-cluster (effect size) heterogeneity. Results are summarized in Table 3. Overall, the three-level meta-analytic results, reflecting 1862 outcomes across 12 domains, revealed a statistically significant intervention effect (Hedges' $g = 0.194$; 95% CI = [0.166, 0.221]; $p < .001$) in favor of the SEL intervention condition compared to the control group at post-intervention. Assessment of statistical heterogeneity was significant (Q (1861) = 27,101.93, $p < .001$, $I^2 = 93.7\%$) indicating substantial heterogeneity across studies and suggesting the presence of one or more moderators.

We hypothesized our review would yield significant effect size improvements across all twelve outcomes domains in favor of USB SEL interventions. Table 3 presents Hedges' g effect sizes and 95% confidence intervals at post-intervention across all studies in each outcome category. Nine of the twelve domain effect sizes (g range = 0.122–0.293) were statistically significant and partially confirm our first hypothesis. Results are based on 7–124 SEL interventions depending on the domain outcome category, and represent 11–367 outcome measures, respectively (See Table 3). Students who participated in USB SEL interventions demonstrated significant improvement (in order of greatest magnitude) for school climate/safety ($g = 0.293$, $k = 116$), civic attitudes/behaviors ($g = 0.255$, $k = 15$), SEL skills ($g = 0.219$, $k = 322$), peer relationships ($g = 0.222$, $k = 84$) attitudes/beliefs ($g = 0.209$, $k = 220$), prosocial behaviors ($g = 0.178$, $k = 199$), externalizing behaviors ($g = 0.162$, $k = 367$), emotional distress ($g = 0.140$, $k = 305$), and school functioning ($g = 0.122$, $k = 151$), compared to controls. We further report that students who participated in USB SEL interventions demonstrated significant mean effect size improvement in academic achievement ($g = 112$, $k = 76$). The effect size improvements equate to a small to moderate effect (Baird & Pane, 2019; Cohen, 1988; Kraft, 2020).

We found no evidence of significant mean differences on discipline, physical health, nor family relationships between the intervention and control groups at post-intervention ($p > .05$).

Follow-up meta-analytic effects

Forty-seven studies (11%) reported follow-up data at least 6 months after the intervention ended. The average follow-up was 52.8 weeks (range = 24–156 weeks). Meta-analytic results of follow-up between-group effects were conducted on 327 outcomes and are presented in Table S14. An additional 9 studies (2.1%) reported *only* follow-up effects. These studies are noted in Table S15. Overall, follow-up effect sizes were in favor of SEL intervention condition (Hedges' $g = 0.167$; 95% CI = [0.103, 0.230]; $p < .001$). For six outcome domains (50%) follow-up effect sizes remained significant and yielded a small to small-moderate effect in favor of the SEL intervention condition. Students in SEL intervention programs demonstrated improved SEL skills ($g = 0.178$, $k = 67$), attitudes/beliefs ($g = 0.200$, $k = 28$), and peer relationships ($g = 0.267$, $k = 13$), and reduced emotional distress ($g = 0.122$, $k = 76$) and externalizing behaviors ($g = 0.218$, $k = 76$). Improved prosocial behaviors ($g = 0.141$, $k = 20$) and school functioning ($g = 0.116$, $k = 21$) were marginally significant ($p < .10$). Follow-up effects for civic attitudes/behaviors, disciplinary outcomes, family relationships and physical health were not applicable for meta-regression analysis due to limited clustering data. No other SEL domains had significant follow-up effects. Analysis revealed significantly high cross-study heterogeneity (Q (326) = 3293.00, $p < .001$, $I^2 = 92.5\%$). Given the limited number of studies with follow-up effects, all subsequent analyses were conducted at post only (Durlak et al., 2011).

Do SEL skills targeted in interventions significantly improve student outcomes?

Table S18 presents the moderation analysis of effect sizes obtained for three contemporary SEL content frameworks: The Framework for SEL (three domains), the CASEL Framework (5 domains), and the EASEL Taxonomy (six domains). We expected programs with instruction that target all the skill domains of each framework to yield better effects than programs with instruction in only one skill domain within each framework. Programs with instruction in all three of the skill domains of the Framework for SEL produced significantly better peer relationships ($g=0.249$) and improvements in emotional distress ($g=0.080$) than programs that instructed one skill domain of the Framework for SEL ($p<.05$). Programs with instruction in all five CASEL domains demonstrated significantly better outcomes in peer relationships ($g=0.225$) than programs that instructed skills in only one domain ($p<.001$). We did not find evidence that programs that instructed skills in all six of the EASEL domains produced significantly better outcomes than those instructing skills in any one domain ($p>.05$).

We further hypothesized that programs that included instruction with greater depth of coverage (higher number of discrete skills within domains) would yield larger improvement effects than programs that offer instruction in fewer targeted skills within domains. Due to limitations in data reported about intervention content, we were unable to reliably collect these data for analyses as proposed. We summarize best available data on depth of skills covered within domains in Table S13 and report these results here descriptively only. Best available data includes skills endorsed in-text or available through one step further of due diligence. We reviewed intervention content if linked to a program website (i.e., www.therulerapporach.org; Rivers et al., 2013) or when studies referenced a previously published paper (i.e., Fossum et al., 2017 references Webster-Stratton & Herman, 2010). Programs endorsed 116 discrete skill categories at least once, and we cataloged 4626 discrete skills across all domains, respectively. The most frequently endorsed skills were behavior emotion regulation (225), cognitive emotion regulation (201), emotion awareness (169), emotion expression (147), problem solving (138), and conflict resolution (114; Table S13).

Are SEL effects moderated by delivery format, program features, dosage, and sequencing?

We hypothesized that program delivery by teachers and multicomponent approaches would yield significant effect size improvements compared to those delivered by outside school personnel or solely classroom-based approaches. Programs delivered by teachers were more effective in promoting school functioning ($g=0.112$) than programs delivered by non-school personnel ($p<.05$). We further hypothesized that interventions containing all four SAFE features would yield larger effects than those that did not. This hypothesis was supported: interventions that met all SAFE criteria were more effective at improving SEL skills ($g=0.118$) and externalizing behaviors ($g=0.098$) compared to programs that did not ($p<.05$). Meta-analytic effect sizes of all moderators are in Table S18.

We further hypothesized that sequenced programs that teach intrapersonal skills *before* interpersonal skills would result in significant effect size improvements compared to programs with another sequence. Programs that taught intrapersonal followed by interpersonal skills produced significantly better experiences of school climate ($g=0.814$) than programs that did not use this sequence ($p<.001$), and showed marginal effect size improvement in student SEL skills ($g=0.097$) attitudes and beliefs ($g=0.162$), and reductions in externalizing behaviors ($g=0.081$) compared to programs that did not use this sequence (, $p<.10$; $p=.06$, respectively).

We further hypothesized that programs that integrate SEL into academic instruction would yield larger intervention effects than those that did not. This hypothesis was not supported. Lastly, we hypothesized that programs longer in duration would produce significant effect size improvements compared to shorter programs. Findings supported the reverse relationship. Interventions lasting up to 16 weeks produced significantly better reductions in student externalizing behaviors ($g=-0.12$) compared to interventions lasting longer than 16 weeks ($p<.05$).

Are SEL effects associated with indicators of study design or the quality of implementation?

We hypothesized higher quality study designs (as evidenced by an index of study quality) would yield larger intervention effects at posttest than those with lower quality designs. Evidence was weak, but there was a trend toward significance showing that high quality RCTs produced greater improvements in SEL skills ($g=0.119$) and school functioning ($g=0.128$) compared to low quality RCTs ($p=0.09$, respectively). When examining the overall quality index, results revealed that high quality studies produced significantly better reductions in externalizing behaviors ($g=-0.078$) compared to low quality studies ($p<.05$). No significant findings for high quality quasi-experimental studies were found. Given the significant heterogeneity across studies, findings must be interpreted with caution.

Do effects significantly differ as a function of intervention and study country?

About half of the studies reviewed (55.6%) implemented programs in the same country the program was

developed. We hypothesized that interventions developed in the country of implementation would result in larger effect size improvements than those implemented in another country. Our hypothesis was supported for one outcome: interventions with programs developed and delivered in the same country showed a marginal effect size reduction in student externalizing behaviors $(g=0.082)$ compared to interventions implemented in a country different from where the program was developed $(p=0.06)$. No other outcomes emerged as significant.

Do effects significantly differ as a function of student age?

Lastly, we explored if intervention effects were moderated by student age. Results revealed no significant differences based on student age and among all relationships only early adolescents trended toward marginally improved civic attitudes and behaviors $(g=0.371)$ compared to adolescents $(p=0.07)$. Further research is needed to substantiate these findings.

Potential publication bias

A funnel plot of effects across outcome domains and plots within domains can be found in Table S16. The results from the Egger's regression test indicated a potential for publication bias across all domains $(t[1860]=3.59, p<.001)$, and in the outcome domains of SEL skills $(t[320]=4.10, p<.001)$, prosocial behavior $(t[197]=3.30, p=.001)$, and school climate $(t[114]=4.13, p<.001)$. This suggests that if effect size estimates are biased by the size of the study, it is mostly a result of bias in the three aforementioned domains. Results show less evidence of publication bias in the other outcome domains. Lastly, we compared overall effect sizes separately for published versus not published studies. On average, published studies had an average ES of $g=0.220$ (95% CI [0.191, 0.256]) whereas unpublished was $g=0.164$ (95% CI [0.096, 0.231]) indicating a meaningful difference between these two classes of studies. This finding is consistent with other meta-review research that found the same trend (Chow & Ekholm, 2018).

DISCUSSION

The current findings update the state of the evidence for USB SEL interventions for students in Kindergarten through Twelfth grade. We document a broad literature examining over half a million students with widely varying intervention content, features, and outcomes. Overall, students who participated in USB SEL interventions experienced improved academic achievement, school climate, school functioning, social emotional

skills, attitudes, and prosocial and civic behaviors, and reduced internalizing and externalizing problems. We further found that SEL intervention effects were moderated by program and implementation features. Programs that met SAFE criteria had high quality implementation, were delivered by classroom teachers, focused on school climate, used a multicomponent approach, taught intrapersonal skills first, and integrated SEL into academic content, and those studies with high quality designs, differentially improved student's skills, attitudes, beliefs, and academic outcomes. These findings come at a critical time in international discourse around SEL; they confirm and advance what we know and need to know to support student academic, social, and emotional thriving (Durlak et al., 2022). Our comprehensive search strategy and expansive inclusion criteria resulted in a highly heterogeneous sample in terms of SEL intervention content, implementation, and outcomes. We found a significant overall effect of USB SEL interventions on student outcomes despite this variability, suggesting the robustness of findings. Here we contextualize our findings, challenges, and opportunities for future directions for USB SEL interventions for students in grades K-12.

What we know now: Interpreting effect sizes in the context of the field

The reported effect sizes in this paper reflect the constellation of contemporary SEL interventions and outcomes since 2008. We found significant intervention effects in both *attitudes/beliefs* $(g=0.224)$, including self-efficacy, self-esteem, mindset, perseverance, and optimism, among others, and *civic attitudes and behaviors* $(g=0.255)$, including understanding civic processes and systems, social justice, understanding of current events, moral or ethical reasoning, among others, which is in line with previously reported increases in attitudes toward self and others $(g=0.23$; Durlak et al., 2011). We report statistically significant improvements in students' *SEL skills* $(g=0.228)$, *prosocial behaviors* $(g=0.193)$, and a reduction in *emotional distress* $(g=0.149)$, that are comparable, albeit smaller, than those reported previously (Durlak et al., 2011). We further found that participation in USB SEL programs results in a small but significant reduction in students *externalizing behaviors* $(g=0.171)$, including violence/aggression, noncompliance/behavior problems, and bullying, and that SEL programming improves *school functioning* $(g=0.131)$, which in our analyses includes academic achievement (in the form of tests, grades/GPA), study skills (including attending behavior), indices of behavioral school adjustment (attendance, tardiness), and on-task behaviors. Once more, SEL programs yielded a small effect size improvement on *academics* $(g=0.111)$ which was smaller than the effect size of 0.27 reported previously (Durlak et al., 2011).

Possible explanations for the different effect sizes found in this contemporary meta-analysis include the breadth of domains and outcomes that were investigated, the high statistical heterogeneity across studies included in review, and the varied designs and intervention delivery reported. Importantly, a significant intervention effect found in favor of USB SEL interventions that represent over 500,000 students across the globe despite this heterogeneity is meaningful and points toward the value of SEL. In addition to the broader inclusion criteria of our study, the type of achievement measure reflected in our review may have contributed to differences in effect size. For example, we expect grades to be more readily impacted by SEL given that grades include a teacher's evaluations of students' learning behaviors and are a more proximal indicator of student growth from an SEL intervention (Durlak, 2015; McKown, 2019). In our review, student grades only represented one third of the academic indicators, with standardized tests accounting for two thirds of the distribution of scores included in analysis (See Table S10). It is possible that our study had a larger ratio of test scores to grades than Durlak et al.' (2011) study, and this could partially explain our different effect sizes. Future studies might consider analyzing these types of outcomes separately to explore this hypothesis further. In addition, recent advances in interpreting the magnitude of effect sizes in the context of experimental educational intervention effects on student achievement suggests a modified benchmark is warranted; less than 0.05 is small, 0.05 to less than 0.20 is medium, and 0.20 or greater is large (see Kraft, 2020 for review). Within this contemporary heuristic, which accounts for intervention, outcome, and student heterogeneity, the effect of USB SEL on student achievement found in our review is to moderate. Once more, published studies had an average ES of $g=0.220$. whereas unpublished were $g=0.164$, which is a substantial difference. It is further possible this difference could be contributing to the different ES we report in this review, but in alignment with contemporary best-practice in meta-analytic review, this more conservative unbiased approach is warranted (Pigott & Polanin, 2020).

Lastly, the current report includes four outcome domains that have not been previously examined. We did not find evidence of effects for two domains, *family relationships* and *physical health* outcomes. However, we did find that SEL interventions had a significant effect on *peer relationships* ($g=0.231$), and the largest evidence of effects among all outcomes explored on *school climate and safety* ($g=0.301$). This suggests that finding places and spaces of safety has important implications on student flourishing, which we discuss in the next section.

attachment, liking of, or connectedness to school; general attitudes toward school and education; perceptions of classroom supportiveness and school safety, such as attitudes or feelings about teachers and school staff, such as students feeling that they care or are helpful, school policies or rules are fair, school structure is organized or they feel like they can take risks; and the quality of student relationships with teachers, among others (Table S10). Findings confirm the accumulated evidence: climate changes positively following the administration of USB SEL programs (Berg, Osher, Moroney, & Yoder, 2017; National Commission, 2018; Wang et al., 2020); and students who participate in SEL programs in their classrooms and schools are more engaged and have better quality relationships with their teachers (Brown et al., 2010; Hagelskamp et al., 2013). Importantly, our review's evidence for the impact of USB SEL interventions on students' experiences of inclusion and belonging is significant. Given the dynamic interaction and bidirectionality between healthy climates and student academic, personal, and social development, results encourage the implementation of USB SEL to support whole child development. It is well studied that threats to physical and psychological safety diminishes student attention, impairs working memory, and can result in academic disengagement, disidentification, and underachievement (Berg, Osher, Moroney, & Yoder, 2017; Elias et al., 2019; Shackman et al., 2006). Worse, these effects can extend across a student's developmental trajectory (Berg, Osher, Moroney, & Yoder, 2017; Osher, Pittman, et al., 2020). The possibility that USB SEL interventions can mitigate these trajectories is promising; USB SEL programming may be a way to address related inequities, especially given the lack of evidence in the effectiveness of Diversity, Equity, and Inclusion initiatives (DEI; Paluck et al., 2021; Forscher et al., 2019; Lai et al., 2016). Lastly, given our findings and the dynamic systems within which USB SEL interventions are operating, investigations of the bidirectional relationship between SEL and climate the mediating role of school climate on student achievement, is warranted (Durlak et al., 2022; Osher, Pittman, et al., 2020; Yang et al., 2020).

To support the practical implications of this dynamic impact, we transformed effect sizes into an improvement index, which reflects the expected change in percentile rank for an average comparison group student if the student had received the intervention (WWC, 2017). In the case of school climate and safety, our reported effect size of 0.30 translates to an improvement index of a 12.5 percentile increase in a control group student's experience of school climate had they received an SEL intervention.

USB SEL interventions improve student experiences of school climate and safety

In this review, school climate and safety included: students' feelings of inclusion or belonging, such as bonding,

The role of content of USB SEL intervention outcomes

We report on previously unexamined relations between USB SEL intervention content and student outcomes, and

further consider the role of SAFE features, which were associated with greater SEL skills, school climate and safety, and externalizing behaviors. We more closely examined the roles of the skill sequence (S), focused (F) personal or social skill development, and explicit (E) or specific skill instruction. We found that programs that taught intrapersonal followed by interpersonal skills resulted in better student SEL skills, school climate and safety, and externalizing behaviors (like those meeting SAFE criteria), and this sequence also resulted in better student SEL attitudes and beliefs. Further, we present modest and novel results that teaching skills across all domains of prominent SEL frameworks produced increased and differential effects. These findings point toward both unique and synergistic relations between the skills included in USB SEL interventions and the outcomes SEL is setting out to achieve. Although our analysis was unable to illuminate the role of specific domains of skills due to limited data, it does importantly advance the field by offering evidence of the benefit for learning intrapersonal skills first. Future research should examine whether this is because intrapersonal skills enhance the efficacy of the interpersonal skills that are subsequently learned, as well as explore the synergistic and unique role of particular SEL skill instruction on student outcomes. Greater clarity in this area will support the strategic structuring of USB SEL interventions to optimize outcomes important to students, schools, and communities.

The content delivered by an SEL program reflects just one type of core component that drives effectiveness (Dymnicki et al., 2020). Programs that fulfilled the CASEL framework demonstrated better peer relationships; programs that fulfilled the Framework for Teaching produced better peer relationships and reduced emotional distress among students. These findings are important to consider further the role of how what is taught in an SEL intervention produces differential effects on students' academic, social, and emotional development. It is possible that depth of skill coverage might produce similar benefits for students, but further examination is warranted. When reporting results, programs should be specific about the sequence and content delivered so that future analysis can disentangle the contributions of sequence and content. Our findings specifically suggest the value of teaching emotion skills before social skills, as this produced the strongest effects of SEL programming (Denham et al., 2003; Izard et al., 2001). These findings are important despite the limited, best available data for this report and point to the potential for the SEL field to benefit by increasing the specificity of differentiation in program features, including the content and sequence of instruction.

Student identity and USB SEL interventions

Significant work remains to understand if and how USB SEL intervention effects differed for subgroups of students based on gender, race/ethnicity, native language

learner status, socioeconomic status, LGBTQIA+, or disability status. To advance our knowledge in this area, we proposed to explore moderation analyses on these key identity variables to determine if, when, and for whom USB SEL works. However, the complexities of identity compounded in extraction of these data: the high missingness of identifiers, alongside significant heterogeneity in identity descriptors would result in biased estimates of effects if aggregated as proposed.

For example, all studies that reported participants' *gender* did so in binary terms, obscuring the identity of participants identifying as non-binary or transgender/transsexual. Two studies alluded to the presence of additional gender identities in text: Mogro-Wilson and Tredinnick (2020) noted that 7% of their sample indicated "other gender" in their evaluation of Connect with Kids in the United States, and Silverstone et al. (2017) reported that 34% declined to identify their gender in binary terms in their evaluation of EMPATHY in Canada. Further, only 16% of studies mentioned *students with disabilities*, and the ways in which they were represented varied, with most representations using non-specific indicators such as Special Education Needs, Students with Disabilities, or Individual Education Plan. And about half of the sample did not present any *racial or ethnic identifie*r, and among those that did, country and region-specific designations within studies were complicated by reported generalizations such as *minorities, diverse, other, unspecified, mixed*, and *representative* (see Table S9).

Also, the way race is often reported obscures meaningful in-group distinctions such as the "voluntary" vs "involuntary" behavioral and attitude differences documented by Ogbu and colleagues (Ogbu & Simons, 1998). This same issue holds for ethnicity, such as how the common Hispanic/Latino designation hides enormous linguist and cultural variability among those with different national origins. Simply knowing one's identity says little about how central that identity is to oneself versus other aspects of one's identity; the experience of identity is personal and varied, and intersects with how students interact with and develop skills from SEL programming (Cipriano et al., 2023). Our field, and meta-analyses such as this one, would greatly benefit from reporting identifiers in more multi-dimensional ways as well as applying or developing more sophisticated methodological approaches for capturing the multifaceted complexities of identity. As one example, person-centered analyses (Lanza & Rhoades, 2013) that allow for capturing *profiles* of various identities as latent constructs shows promise for addressing this nuance as profiles likely more accurately represent the nature of identities.

Challenges and opportunities for social and emotional learning

This report documents several challenges for the SEL field to address to support our understanding of the

effectiveness of USB SEL interventions. We discuss these challenges here as opportunities to advance the evaluation, implementation, and effectiveness of USB SEL.

SEL is not a homogeneous construct

Our report documents the vast variability in SEL interventions, comprising: program content including the skill domains and strategies focused on; features such as SAFE criteria, classroom-based, multi-component, focus on School Climate, Adult SEL, or tiered; elements of implementation such as the duration, dosage, and sequencing of the intervention; and who implements the interventions, why it was developed (theoretical foundation) and how implementation was evaluated (domains of outcomes and constructs and the measures embedded within; Durlak et al., 2011; Dymnicki et al., 2020; Lawson et al., 2019). Despite this heterogeneity, our report found significant evidence for the effectiveness of USB SEL interventions for students in schools worldwide. These findings suggest that we can speak with confidence about the robust effectiveness of USB SEL interventions and encourage the field to evolve a more specific discourse around the constellation of USB SEL interventions. This is especially the case given the growing attention being given to concepts such as mindsets and social justice, which differ in approach from more traditional forms of SEL that prioritize competencies (Strambler, 2023; Strambler et al., 2023).

To some extent, heterogeneity in program content is influenced both by the specific theories, conceptual models, and frameworks that researchers use to guide their studies as well as the theorical foundations that SEL programs are based upon. Even among SEL interventions targeting a similar skill and/or subject, there is variability. For instance, studies guided by a mindfulness framework were common in the literature, but these studies differed based on the type of mindfulness framework used, such as general mindfulness, Mindfulness Based Cognitive Therapy (MBCT), and Mindfulness-Based Stress Reduction (MBSR). Although there is a diversity of theories reflected in the SEL literature, notably absent are frameworks that specifically attend to the expected variability in student identities, with developmental, ecological, social, cognitive, behavioral, and emotional frameworks predominately shaping the current USB SEL research. This may explain in part the underrepresentation of student variability in contemporary reports of SEL effects found in the United States (Cipriano et al., 2023; Daley & McCarthy, 2021). An identified area for future exploration is how all students and the SEL field may benefit from pairing frequently used theories and conceptual models in the USB SEL intervention literature with frameworks that explicitly

focus on culturally based (i.e., culturally sustaining pedagogy), accessible and inclusive (i.e., universal design for learning), and sensitive (i.e., trauma informed pedagogy) teaching and learning practices (see CAST, 2018; Gay, 2018; TransformEd, 2020). Increased attention to the diversity of programmatic components would be a helpful and productive contribution to the international SEL discourse (Karanga, n.d.).

Evidence-based practice requires evidence

Our review is limited by the data available, and the evidence available was replete with missingness. Indicators of implementation influences the effect of SEL approaches on student outcomes (Abry et al., 2017; Durlak, 2015), yet more than half of the sample did not monitor implementation in any way (55.5%), and among those that did, more than half did not mention the level of implementation achieved (59.6%). We found evidence of moderation by study quality indices, which predicted significant mean differences in student externalizing behaviors. However, this finding warrants much deeper analyses as it remains unclear to what degree quality of implementation impacted the effectiveness USB SEL interventions due to the lack of data (Domitrovich et al., 2019; Granger et al., 2020; Sutherland & McLeod, 2022).

Relatedly, the significant effects of SEL interventions on student social, emotional, and academic outcomes reported in this review are promising but limited. For example, although we report a significant effect of USB SEL interventions on student academic achievement exclusively ($g=0.11$), only 42 studies (14.8%) included student achievement data for analyses. This is only slightly more than the 35 studies reported in 2011 (Durlak et al.). Further, SEL Skills (Domain 1) were assessed in less than half of the interventions (114; 44%), and *no studies in our review* reported outcomes in every examined domain. Despite the momentum and perceived progress USB SEL intervention design, implementation, and evaluation over the past decade, the reports of effectiveness of these interventions do not reflect such progress. The complete articulation of the benefits of USB SEL interventions requires increased collection and reporting of outcomes that are most meaningful in their practical implications, including contextual measures that are both macro structural and proximal to the interactions and experiences in SEL programs that contribute to what are statistically discerned as impacts so that more sophisticated meta-analyses can be conducted (Durlak et al., 2022; Greenberg, 2023; Osher, Pittman, et al., 2020).

USB SEL measurement matters

The variety of the measures and constructs presented in this review reflects a similar area of growth for the

field. We document a research literature that is simultaneously oversaturated and under-developed in its' use of measures of SEL effects. The same measures were operationalized as different constructs across and within the twelve outcome domains examined, and they were inconsistently presented and justified. This issue is trifold; researchers must clarify *what* construct an assessment measures, *how* it measures that construct, and *who* the measure benefits, to be most effective (McKown, 2019; Ng et al., 2022).

For example, we cataloged the use of over 1800 outcomes of student SEL (Table S10) and know little about the quality of these measures. Less than half of studies reported the reliability and validity of measures used in their analyses, even though only 56 measures were unnamed, unspecified, or study developed. Furthermore, within-domain heterogeneity in the operationalized constructs measured was complex. Domain 1 (SEL Skills) includes 119 different measures reflecting 20 constructs and 30 outcomes that combined 2 or more constructs. This heterogeneity (e.g., social emotional competence vs. emotional competence vs. social and emotional skill) is further complicated by the use of global indicators versus discrete measures. As a further example, the Devereux Student Strengths Assessment (DESSA) was used 5 times to measure SEL skills and twice for prosocial behaviors, whereas the Panorama SEL survey total was used to measure SEL skills, and the Grit and Growth Mindset subscales were within *attitudes/beliefs*, respectively. Once more, the use of macro indicators like these and the hundreds of others reflected in our review are distal to the real-time learning and interactions in classrooms that may statistically discern impacts of an SEL intervention more robustly (Ng et al., 2022; Osher, Pittman, et al., 2020).

The most frequently reported assessment in our review was the Strengths and Difficulties Questionnaire (SDQ; Goodman, 2001), which is a student self-report or adult-report measure of youth prosocial behavior and psychopathology organized into emotional, conduct, hyperactivity-inattention, peer, and prosocial subscales. The SDQ has been translated into more than 40 languages and was developed and normed on a study of just over 10,000 British youth ages 5–15 years old in 1999. The SDQ is reflected 154 times across 7 outcome domains in our study. A recent review of emotion regulation measure development and validation studies used in US school-based studies from past two decades documented similar challenges; reporting divergent conceptualizations of emotion regulation, trade-offs between methodological rigor and practicality, under-sampling of marginalized youth, and insufficient psychometric evidence across assessments (See Ng et al., 2022 for review). A manuscript providing methodological guidance for conducting and reporting an evaluation of an USB SEL intervention based on

the findings of this report is in preparation (Cipriano et al., 2023).

Attend to the intersection of business and efficacy of USB SEL

Lastly, USB SEL intervention science would benefit from attending to the intersection of industry and efficacy of USB SEL interventions. Our review found evidence of potential publication bias across all outcome domains, and that published studies had larger average effects than unpublished reports of USB SEL. Although these findings are consistent with other recent meta-analyses of school-based interventions (Chow & Ekholm, 2018), further examination is warranted to consider if published studies of USB SEL effects are more likely to be from commercialized USB SEL interventions. Relatedly, a question remains whether program evaluations of SEL interventions commissioned by SEL program developers produce larger effect sizes than evaluations conducted by independent third parties. Once more, measure development by USB SEL program providers, or the coupling of measures by different entities, to support district, state, and country evaluations of SEL interventions is increasingly more common (RAND, 2020). Wolf and Harbatkin (2023) used data from the What Works Clearinghouse to categorize outcome measures in educational interventions and documented how average effect sizes varied for outcome measure type by whether the group who developed the measure potentially had a stake in the intervention (non-independent) or not (independent). Research designs and the possibility of a conflict of interest further complicates an already complex constellation of potential outcomes in evaluations of USB SEL interventions (Cheung & Slavin, 2016; Greenberg, 2023; Wolf & Harbatkin, 2023). Subsequent review of best available data is underway and available upon request.

Fostering social–emotional health of children and adolescents with USB SEL

The social and emotional health of young people continues to be a global concern. At the time this manuscript was prepared, an estimated 68 million youth ages 10–19 were suffering from depression and anxiety (UNICEF, 2021), with rates of youth suicide being more prevalent worldwide than ever reported before (CDC, 2019; UNICEF, 2021). In the wake of a global tragedy, where an estimated 10.5 million children lost a parent or caregiver to the COVID-19 pandemic (Hillis et al., 2022), when students worldwide experienced unprecedented interruptions to education as usual, and rates of gun violence spiked, the world's attention turned

to social and emotional learning (Cipriano et al., In Progress).

The promise of USB SEL interventions to support the social, emotional, and academic trajectories of the next generations of global citizens is ripe. The amassed data of this report only reflects the beginning of the contemporary evidence suggesting SEL interventions support healthy and inclusive school climates, increased school achievement, and healthy attitudes and behaviors among students. Once more, an important takeaway from this review is that SEL interventions vary widely and are not one cohesive approach. Rather, our evidence suggests it is more useful to think of USB SEL interventions as a wide range of implementation approaches focusing on a wide range of outcomes in the service of student social, emotional, and academic development. Stakeholders worldwide will benefit from understanding the best fitting SEL approaches for their context and the supporting evidence for such approaches; it is critical for future research to rigorously study the effectiveness of SEL across context and features. Addressing these areas in SEL research and practice show promise for fostering the social–emotional health of children and adolescents worldwide.

CONFLICT OF INTEREST STATEMENT

We have no known conflicts of interest to disclose.

DATA AVAILABILITY STATEMENT

All data and code for analyses presented in this manuscript are available on Open Science Framework at https://osf.io/6pqx2/. Additional data available upon request.

ORCID

Christina Cipriano https://orcid.org/0000-0002-7414-1821
Michael J. Strambler https://orcid.org/0000-0002-5576-0408
Lauren H. Naples https://orcid.org/0000-0002-9467-5898
Cheyeon Ha https://orcid.org/0000-0002-4966-7055
Megan Kirk https://orcid.org/0000-0002-2069-2177
Miranda Wood https://orcid.org/0000-0002-3973-3549
Kaveri Sehgal https://orcid.org/0000-0001-8855-9818
Almut K. Zieher https://orcid.org/0000-0002-2909-5598
Abigail Eveleigh https://orcid.org/0000-0002-3526-8328
Michael McCarthy https://orcid.org/0000-0002-4007-2020
Melissa Funaro https://orcid.org/0000-0001-6846-4846
Annett Ponnock https://orcid.org/0000-0002-5983-7767
Jason C. Chow https://orcid.org/0000-0002-2878-7410

REFERENCES

Abry, T., Hullerman, C. S., & Rimm-Kaufman, S. E. (2015). Using indices of fidelity to intervention core components to identify program active ingredients. *American Journal of Evaluation*, *36*(3), 320–338. https://doi.org/10.1177/1098214014557009

Abry, T., Rimm-Kaufman, S. E., & Curby, T. W. (2017). Are all program elements created equal? Relations between specific social and emotional learning components and teacher–student classroom interaction quality. *Prevention Science*, *18*(2), 193–203.

Baird, M. D., & Pane, J. F. (2019). Translating standardized effects of education programs into more interpretable metrics. *Educational Researcher*, *48*(4), 217–228. https://doi.org/10.3102/0013189X19848729

Berg, J., Osher, D., Moroney, D., & Yoder, N. (2017). *The intersection of school climate and social and emotional development*. American Institutes for Research https://www.air.org/sites/default/files/downloads/report/Intersection-School-Climate-and-Social-and-Emotional-Development-February-2017.pdf

Berg, J., Osher, D., Same, M. R., Nolan, E., Benson, D., & Jacobs, N. (2017). *Identifying, defining, and measuring social and emotional competencies*. American Institutes for Research.

Borenstein, M., Cooper, H., Hedges, L., & Valentine, J. (2009). Effect sizes for continuous data. In *The handbook of research synthesis and meta-analysis* (Vol. *2*, pp. 221–235). Russell Sage Foundation.

Brown, J. L., Jones, S. M., LaRusso, M. D., & Aber, J. L. (2010). Improving classroom quality: Teacher influences and experimental impacts of the 4rs program. *Journal of Educational Psychology*, *102*(1), 153–167. https://doi.org/10.1037/a0018160

Cantor, P., Osher, D., Berg, J., Steyer, L., & Rose, T. (2019). Malleability, plasticity, and individuality: How children learn and develop in context1. *Applied Developmental Science*, *23*(4), 307–337. https://doi.org/10.1080/10888691.2017.1398649

Carbonneau, K. J., Marley, S. C., & Selig, J. P. (2013). A meta-analysis of the efficacy of teaching mathematics with concrete manipulatives. *Journal of Education and Psychology*, *105*(2), 380–400. https://doi.org/10.1037/a0031084

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

Center for Applied Special Technology [CAST]. (2018). *Universal design for learning guidelines version 2.2.* http://udlguidelines.cast.org

Centers for Disease Control and Prevention [CDC]. (2019). *Youth risk behavior survey: Data summary and trends report 2009–2019.* https://www.cdc.gov/healthyyouth/data/yrbs/pdf/YRBSDataSummaryTrendsReport2019508.pdf

Cheung, A. C., & Slavin, R. E. (2016). How methodological features affect effect sizes in education. *Educational Researcher*, *45*(5), 283–292. https://doi.org/10.3102/0013189X16656

Chow, J. C., & Ekholm, E. (2018). Do published studies yield larger effect sizes than unpublished studies in education and special education? A meta-review. *Educational Psychology Review*, *30*, 727–744. https://doi.org/10.1007/s10648-018-9437-7

Cipriano, C., Ha, C., Polanin, J., Chow, C., & Strambler, M. (In Progress). Methodological Guidance Paper: Conducting a High Quality Evaluation of an SEL Intervention.

Cipriano, C., Naples, L. H., Eveleigh, A., Funaro, M., Cook, A., Cassidy, C., McCarthy, M., & Schlichtmann, G. (2023). A systematic review of student disability and race representation in universal school-based SEL interventions for elementary school students. *Review of Educational Research*, *93*(1), 73–102. https://doi.org/10.3102/00346543221094079

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Earlbaum Associates.

Collaborative for Academic, Social, and Emotional Learning. (2013). *Effective social and emotional learning programs: Preschool and elementary school edition.* Author.

Collaborative for Academic, Social, and Emotional Learning. (2015). *Effective social and emotional learning programs: Middle and high school edition.* Author.

Collaborative for Academic, Social and Emotional Learning [CASEL]. (2017). *Sample teaching strategies to support core competencies of social and emotional learning.* https://casel.org/wp-content/uploads/2017/08/Sample-Teaching-Activities-to-Support-Core-Competencies-8-20-17.pdf

Collaborative for Academic, Social and Emotional Learning [CASEL]. (2020). *What is SEL?* https://casel.org/what-is-sel/

Congress, U. S. (2020). Coronavirus Aid, Relief, and Economic Security Act (CARES Act). P.L., 116-136.

Crenshaw, K. (1991). Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stanford Law Review*, *43*(6), 1241. https://doi.org/10.2307/1229039

Daley, S. G., & McCarthy, M. F. (2021). Students with disabilities in social and emotional learning interventions: A systematic review. *Remedial and Special Education*, *42*(6), 384–397. https://doi.org/10.1177/0741932520964917

Denham, S. A. (2006). Social-emotional competence as support for school readiness: What is it and how do we assess it? *Early Education and Development*, *17*(1), 57–89. https://doi.org/10.1207/s15566935eed1701_4

Denham, S. A., Blair, K. A., DeMulder, E., Levitas, J., Sawyer, K., Auerbach–Major, S., & Queenan, P. (2003). Preschool emotional competence: Pathway to social competence? *Child Development*, *74*(1), 238–256. https://doi.org/10.1111/1467-8624.00533

Domitrovich, C. E., Durlak, J. A., Staley, K. C., & Weissberg, R. P. (2017). Social-emotional competence: An essential factor for promoting positive adjustment and reducing risk in school children. *Child Development*, *88*(2), 408–416.

Domitrovich, C. E., Li, Y., Mathis, E. T., & Greenberg, M. T. (2019). Individual and organizational factors associated with teacher self-reported implementation of the PATHS curriculum. *Journal of School Psychology*, *76*, 168–185. https://doi.org/10.1016/j.jsp.2019.07.015

Duchesneau, N. (2020). *Social, emotional, and academic development through an equity lens.* Education Trust.

Durlak, J. A. (2015). Studying program implementation is not easy but it is essential. *Prevention Science*, *16*(8), 1123–1127. https://doi.org/10.1007/s11121-015-0606-3

Durlak, J. A. (2016). Program implementation in social and emotional learning: Basic issues and research findings. *Cambridge Journal of Education*, *46*(3), 333–345.

Durlak, J. A., Mahoney, J., & Boyle, A. (2022). *What we know, and what we need to find out about universal, school-based social and emotional learning programs for children and adolescents: A review of meta-analyses and directions for future research.* Psychological Research Online First.

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

Dusenbury, L., Yoder, N., Dermody, C., & Weissberg, R. P. (2019). *An examination of frameworks for social and emotional learning (SEL) reflected in State K-12 learning standards.* https://measuringsel.casel.org/wp-content/uploads/2019/02/FrameworkC.3.pdf

Dymnicki, A., Trivits, L., Hoffman, C., & Osher, D. (2020). *Advancing the use of core components of effective programs: Suggestions for researchers publishing evaluation results.* Office of Assistant Secretary for Planning and Evaluation, US Department of Health and Human Services.

Ecological Approaches to Social Emotional Learning [EASEL]. (n.d.). *Explore EASEL: Identify related skills across frameworks.* http://exploresel.gse.harvard.edu/compare-terms/

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634.

Elias, M. J., Brackett, M. A., Miller, R., Jones, S., Kahn, J., Mahoney, J. L., Weissberg, R. P., & Chung, S. Y. (2019). Developing social and emotional skills and attitudes and ecological assets. In D. Osher, M. J. Mayer, R. J. Jagers, K. Kendziora, & L. Wood (Eds.), *(Eds.), Keeping students safe and helping them thrive: A collaborative handbook on school safety, mental health, and wellness* (pp. 185–209). Praeger.

Elias, M. J., & Yuan, M. (2020). Creating coherent conceptualization and communication among character education kinfolk: Averting a "tower of babel". *Journal of Character Education*, *16*(2), 19–26.

Every Student Succeeds Act of 2015. (2015). Pub. L. No. 114-95 § 114 Stat. 1177.

Farrington, C. A. (2020). Equitable learning and development: Applying science to foster liberatory education. *Applied Developmental Science*, *24*(2), 159–169.

Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., Devine, P. G., & Nosek, B. A. (2019). A meta-analysis of procedures to change implicit measures. *Journal of Personality and Social Psychology*, *117*(3), 522–559. https://doi.org/10.1037/pspa0000160

Fossum, S., Handegård, B. H., & Britt Drugli, M. (2017). The incredible years teacher classroom management programme in kindergartens: Effects of a universal preventive effort. *Journal of Child and Family Studies*, *26*(8), 2215–2223. https://doi.org/10.1007/s10826-017-0727-3

Frank, J. L. (2020). School-based practices for the 21st century: Noncognitive factors in student learning and psychosocial outcomes. *Policy Insights from the Behavioral and Brain Sciences*, *7*(1), 44–51. https://doi.org/10.1177/2372732219898703

Gay, G. (2018). *Culturally responsive teaching: Theory, research, and practice.* Teachers college press.

Goleman, D. (1995). *Emotional intelligence.* Bantam Books, Inc.

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(11), 1337–1345.

Granger, K. L., Sutherland, K. S., Conroy, M. A., Hetrick, A. A., & Parnell, E. (2020). Barriers and facilitators to implementation of BEST in CLASS. *Exceptionality*, *28*(3), 209–221.

Grant, S., Hamilton, L. S., Wrabel, S. L., Gomez, C. J., Whitaker, A., Leschitz, J. T., Unlu, F., Chavez-Herrerias, E. R., Baker, G., Barrett, M., Harris, M., & Ramos, A. (2017). *Social and emotional learning interventions under the every student succeeds act: Evidence review. Research report. RR-2133-WF.* RAND Corporation.

Greenberg, M. (2023). *The evidence for social and emotional learning.* The Learning Policy Institute.

Greenberg, M., Domitrovich, C., Weissberg, R., & Durlak, J. (2017). Social and emotional learning as a public health approach to education. *The Future of Children*, *27*(1), 13–32.

Hagelskamp, C., Brackett, M. A., Rivers, S. E., & Salovey, P. (2013). Improving classroom quality with the ruler approach to social and emotional learning: Proximal and distal outcomes. *American Journal of Community Psychology*, *51*, 530–543.

Hamilton, L. S., Doss, C. J., & Steiner, E. D. (2019). *Teacher and principal perspectives on social and emotional learning in America's schools: Findings from the American educator panels.* RAND Corporation.

Hamre, B. K., & Pianta, R. C. (2007). Learning opportunities in preschool and early elementary classrooms. In R. C. Pianta, M. J. Cox, & K. L. Snow (Eds.), *School readiness and the transition to kindergarten in the era of accountability* (pp. 49–83). Paul H Brookes Publishing.

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2021). *Doing meta-analysis with R: A hands-on guide.* Chapman & Hall/CRC Press.

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta- regression with dependent effect size estimates. *Research Synthesis Methods*, *1*(1), 39–65.

Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savović, J., Schulz, K. F., Weeks, L., & Sterne, J. A. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*, *343*, d5928.

Hillis, S., N'konzi, J. P. N., Msemburi, W., Cluver, L., Villaveces, A., Flaxman, S., & Unwin, H. J. T. (2022). Orphanhood and caregiver loss among children based on new global excess COVID-19 death estimates. *JAMA Pediatrics*, *176*(11), 1145–1148.

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.

Immordino-Yang, M. H., Darling-Hammond, L., & Krone, C. R. (2019). Nurturing nature: How brain development is inherently social and emotional, and what this means for education. *Educational Psychologist*, *54*(3), 185–204.

Izard, C., Fine, S., Schultz, D., Mostow, A., Ackerman, B., & Youngstrom, E. (2001). Emotion knowledge as a predictor of social behavior and academic competence in children at risk. *Psychological Science*, *12*(1), 18–23. https://doi.org/10.1111/1467-9280.00304

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

Joanna Briggs Institute. (2020). *Checklist for systematic reviews and research syntheses.* https://jbi.global/critical-appraisal-tools

Jones, S. (n.d.). *The Taxonomy Project.* https://easel.gse.harvard.edu/files/gse-easel-lab/files/taxonomy_handout_0.pdf

Jones, S., Bailey, R., Brush, K., Nelson, B., & Barnes, S. (2016). What is the same and what is different? Harvard graduate school of education. https://easel.gse.harvard.edu/news/what-same-and-what-different

Jones, S., Brush, K., Bailey, R., Brion-Meisels, G., McIntyre, J., Kahn, J., Nelson, B., & Stickle, L. (2017). *Navigating SEL from the inside out: Looking inside & across 25 leading SEL programs—A practical resource for schools and OST providers.* Harvard Graduate School of Education.

Jones, S. M., & Bouffard, S. M. (2012). Social and emotional learning in schools: From programs to strategies and commentaries. *Social Policy Report*, *26*(4), 1–33.

Jones, S. M., McGarrah, M. W., & Kahn, J. (2019). Social and emotional learning: A principled science of human development in context. *Educational Psychologist*, *54*(3), 129–143.

Karanga. (n.d.). https://karanga.org/

Krachman, S. B., & LaRocca, B. (2017). *The scale of our investment in social and emotional learning.* Transforming Education https://www.transformingeducation.org/wp-content/uploads/2017/09/Inspire-PaperTransformingEd.pdf

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, *49*(4), 241–253.

Kremer, K. P., Maynard, B. R., Polanin, J. R., Vaughn, M. G., & Sarteschi, C. M. (2015). Effects of after-school programs with at-risk youth on attendance and externalizing behaviors: A systematic review and meta-analysis. *Journal of Youth and Adolescence*, *44*(3), 616–636.

Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., Calanchini, J., Xiao, Y. J., Pedram, C., Marshburn, C. K., Simon, S., Blanchar, J. C., Joy-Gaba, J. A., Conway, J., Redford, L., Klein, R. A., Roussos, G., Schellhaas, F. M. H., Burns, M., … Nosek, B. A. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General*, *145*(8), 1001–1016. https://doi.org/10.1037/xge0000179

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology*, *4*, 1–12.

Lanza, S. T., & Rhoades, B. L. (2013). Latent class analysis: An alternative perspective on subgroup analysis in prevention and treatment. *Prevention Science*, *14*(2), 157–168.

Lawson, G. M., McKenzie, M. E., Becker, K. D., Selby, L., & Hoover, S. A. (2019). The core components of evidence-based social emotional learning programs. *Prevention Science*, *20*(4), 457–467. https://doi.org/10.1007/s11121-018-0953-y

Lester, A. M., Chow, J. C., & Melton, T. N. (2020). Quality is critical for meaningful synthesis of afterschool program effects: A systematic review and meta-analysis. *Journal of Youth and Adolescence*, *49*(2), 369–382.

Low, S., Smolkowski, K., & Cook, C. (2016). What constitutes high-quality implementation of SEL programs? A latent class analysis of second step® implementation. *Prevention Science*, *17*(8), 981–991. https://doi.org/10.1007/s11121-016-0670-3

McGowan, J., Sampson, M., Salzwedel, D. M., Cogo, E., Foerster, V., & Lefebvre, C. (2016). PRESS peer review of electronic search strategies: 2015 guideline statement. *Journal of Clinical Epidemiology*, *75*, 40–46.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282.

McKown, C. (2019). Challenges and opportunities in the applied assessment of student social and emotional learning. *Educational Psychologist*, *54*(3), 205–221.

Mogro-Wilson, C., & Tredinnick, L. (2020). Influencing social and emotional awareness and empathy with a visual arts and music intervention for adolescents. *Children & Schools*, *42*(2), 111–119. https://doi.org/10.1093/cs/cdaa008

National Commission on Social, Emotional, and Academic Development. (2018). *From a nation at risk to a nation at hope: Recommendations from the National Commision.* The Aspen Institute http://nationathope.org/report-from-the-nation/

Newman, J., & Dusenbury, L. (2015). Social and emotional learning (SEL): A framework for academic, social, and emotional success. In *Prevention science in school settings* (pp. 287–306). Springer.

Ng, Z. J., Willner, C. J., Mannweiler, M. D., Hoffmann, J. D., Bailey, C. S., & Cipriano, C. (2022). A systematic review of emotion regulation assessments in US schools: Bridging the

gap between researchers and ducators. *Educational Psychology Review*, *34*, 2825–2865. https://doi.org/10.1007/s10648-022-09691-4

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science*, *7*(6), 615–631.

Ogbu, J. U., & Simons, H. D. (1998). Voluntary and involuntary minorities: A cultural-ecological theory of school performance with some implications for education. *Anthropology & Education Quarterly*, *29*(2), 155–188.

Organization for Economic Cooperation and Development. (2017). *Social and emotional skills well-being, connectedness and success.* https://www.oecd.org/education/school/UPDATED%20Social%20and%20Emotional%20Skills%20-%20Well-being,%20connectedness%20and%20success.pdf%20(website).pdf

Osher, D., Cantor, P., Berg, J., Steyer, L., & Rose, T. (2020). Drivers of human development: How relationships and context shape learning and development1. *Applied Developmental Science*, *24*(1), 6–36.

Osher, D., Kidron, Y., Brackett, M., Dymnicki, A., Jones, S., & Weissberg, R. P. (2016). Advancing the science and practice of social and emotional learning: Looking back and moving forward. *Review of Research in Education*, *40*(1), 644–681.

Osher, D., Pittman, K., Young, J., Smith, H., Moroney, D., & Irby, M. (2020). *Thriving, robust equity, and transformative learning & development: A more powerful conceptualization of the contributors to youth success.* American Institutes for Research and Forum for Youth Investment https://www.air.org/sites/default/files/Thriving-Robust-Equity,-and-Transformative-Learning-and-Development-July-2020.pdf

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., & McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, *372*.

Paluck, E. L., Porat, R., Clark, C. S., & Green, D. P. (2021). Prejudice reduction: Progress and challenges. *Annual Review of Psychology*, *72*, 533–560.

Pigott, T. D., & Polanin, J. R. (2020). Methodological guidance paper: High-quality meta- analysis in a systematic review. *Review of Educational Research*, *90*(1), 24–46.

Polanin, J. R., Hennessy, E. A., & Tsuji, S. (2020). Transparency and reproducibility of meta-analyses in psychology: A meta-review. *Perspectives on Psychological Science*, *15*(4), 1026–1041. https://doi.org/10.1177/1745691620906416

Polanin, J. R., Tanner-Smith, E. E., & Hennessy, E. A. (2016). Estimating the difference between published and unpublished effect sizes: A meta-review. *Review of Educational Research*, *86*(1), 207–236.

Pustejovsky, J. (2022). *clubSandwich: Cluster-Robust (sandwich) variance estimators with small-sample corrections.* https://CRAN.R-project.org/package=clubSandwich

Pustejovsky, J. E., & Tipton, E. (2022). Meta-analysis with robust variance estimation: Expanding the range of working models. *Prevention Science*, *23*(3), 425–438.

Quintana, D. S., & Tiebel, J. (2019). How to calculate statistical power for your meta-analysis. https://doi.org/10.17605/OSF.IO/5C7UZ

RAND. (2020). *RAND Education Assessment Finder.* https://www.rand.org/education-and-labor/projects/assessments/tool.html

Rimm-Kaufman, S. E., & Hulleman, C. S. (2015). Social and emotional learning in elementary school settings: Identifying mechanisms that matter. In J. A. Durlak, C. E. Domitrovich, R. P. Wissberg, & T. P. Gullotta (Eds.), *Handbook of social and emotional learning: Research and practice* (pp. 151–166). Guilford Press.

Rivers, S. E., Brackett, M. A., Reyes, M. R., Elbertson, N. A., & Salovey, P. (2013). Improving the social and emotional climate of classrooms: A clustered randomized controlled trial testing the RULER approach. *Prevention Science*, *14*(1), 77–87. https://doi.org/10.1007/s11121-012-0305-2

Rodgers, M. A., & Pustejovsky, J. E. (2021). Evaluating meta-analytic methods to detect selective reporting in the presence of dependent effect sizes. *Psychological Methods*, *26*(2), 141–160. https://doi.org/10.1037/met0000300

Rowe, H. L., & Trickett, E. J. (2018). Student diversity representation and reporting in universal school-based social and emotional learning programs: Implications for generalizability. *Educational Psychology Review*, *30*(2), 559–583. https://doi.org/10.1007/s10648-017-9425-3

Schonert-Reichl, K. A. (2019). Advancements in the landscape of social and emotional learning and emerging topics on the horizon. *Educational Psychologist*, *54*(3), 222–232.

Shackman, A. J., Sarinopoulos, I., Maxwell, J. S., Pizzagalli, D. A., Lavric, A., & Davidson, R. J. (2006). Anxiety selectively disrupts visuospatial working memory. *Emotion*, *6*(1), 40–61. https://doi.org/10.1037/1528-3542.6.1.40

Shankland, R., & Rosset, E. (2017). Review of brief school-based positive psychological interventions: A taster for teachers and educators. *Educational Psychology Review*, *29*(2), 363–392. https://doi.org/10.1007/s10648-016-9357-3

Silverstone, P. H., Bercov, M., Suen, V. Y. M., Allen, A., Cribben, I., Goodrick, J., Henry, S., Pryce, C., Langstraat, P., Rittenbach, K., Chakraborty, S., Engles, R. C., & McCabe, C. (2017). Long-term results from the empowering a multimodal pathway toward healthy youth program, a multimodal school-based approach, show marked reductions in suicidality, depression, and anxiety in 6,227 students in grades 6-12 (aged 11–18). *Frontiers in Psychiatry*, *8*, 81. https://doi.org/10.3389/fpsyt.2017.00081

Sklad, M., Diekstra, R., De Ritter, M., Ben, J., & Gravesteijn, C. (2012). Effectiveness of school-based universal social, emotional, and behavioral programs. Do they enhance students' development in the area of skill, behavior, and adjustment? *Psychology and Schools*, *49*, 892–909.

Strambler, M. J. (2023). *The intellectually honest case for social and emotional learning.* Discourse Magazine. https://www.discoursemagazine.com/culture-and-society/2023/05/12/open-inquiry-initiative-the-intellectually-honest-case-for-social-and-emotional-learning/

Strambler, M. J., Schonert-Reichl, K. A., & Rimm-Kaufman, S. E. (2023). Conclusion: Trends, challenges, and future directions in SEL. In S. E. Rimm-Kaufman, M. J. Strambler, & K. A. Schonert-Reichl (Eds.), *Social and emotional learning in action: Creating systemic change in schools* (pp. 281–297). The Guilford Press.

Sutherland, K. S., & McLeod, B. D. (2022). Advancing the science of integrity measurement in school mental health research. *School Mental Health*, *14*(1), 1–6.

Tanner-Smith, E. E., & Tipton, E. (2014). Robust variance estimation with dependent effect sizes: Practical considerations including a software tutorial in Stata and SPSS. *Research Synthesis Methods*, *5*(1), 13–30.

Taylor, R., 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*, 1156–1171.

Tipton, E. (2015). Small sample adjustments for robust variance estimation with meta- regression. *Psychological Methods*, *20*(3), 375–393.

Tipton, E., & Pustejovsky, J. E. (2015). Small-sample adjustments for tests of moderators and model fit using robust variance estimation in meta-regression. *Journal of Educational and Behavioral Statistics*, *40*(6), 604–634.

TransformEd. (2020). *Trauma-informed SEL toolkit.* https://www.transformingeducation.org/trauma-informed-sel-toolkit/

United Nations Children's Fund. (2021). *Seen, counted, included: Using data to shed light on the well-being of children with disabilities.* UNICEF https://data.unicef.org/resources/children-with-disabilities-report-2021/

Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics, 35*(2), 215–247.

Wang, M. T., Degol, J. L., Amemiya, J., Parr, A., & Guo, J. (2020). Classroom climate and children's academic and psychological wellbeing: A systematic review and meta-analysis. *Developmental Review, 57*, 100912.

Wanless, S. B., & Barnes, T. N. (2020). *Why social justice and equity is essential to SEL.* Office of Child Development, University of Pittsburgh.

Wanless, S. B., & Domitrovich, C. E. (2015). Readiness to implement school-based social-emotional learning interventions: Using research on factors related to implementation to maximize quality. *Prevention Science, 16*(8), 1037–1043. https://doi.org/10.1007/s11121-015-0612-5

Webster-Stratton, C., & Herman, K. C. (2010). Disseminating incredible years series early intervention programs: Integrating and sustaining services between school and home. *Psychology in the Schools, 47*, 36–54. https://doi.org/10.1002/pits.20450

Weissberg, R. P., Durlak, J. A., Domitrovich, C. E., & Gullotta, T. P. (2015). Social and emotional learning: Past, present, and future. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds.), *Handbook of social and emotional learning: Research and practice* (pp. 3–19). Guilford Press.

What Works Clearinghouse. (2017). *Procedures handbook (Version 4.0).* https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_procedures_handbook_v4_draft.pdf

Wigelsworth, M., Lendrum, A., Oldfield, J., Scott, A., Ten Bokkel, I., Tate, K., & Emery, C. (2016). The impact of trial stage, developer involvement and international transferability on universal social and emotional learning programme outcomes: A meta-analysis. *Cambridge Journal of Education, 46*(3), 347–376.

Wolf, B., & Harbatkin, E. (2023). Making sense of effect sizes: Systematic differences in intervention effect sizes by outcome measure type. *Journal of Research on Educational Effectiveness, 16*(1), 134–161.

World Health Organization. (2003). *Skills for health: Skills-based health education including life skills: An important component of a child-friendly/health-promoting school.* World Health Organization. https://apps.who.int/iris/handle/10665/42818

Yang, C., Chan, M. K., & Ma, T. L. (2020). School-wide social emotional learning (SEL) and bullying victimization: Moderating role of school climate in elementary, middle, and high schools. *Journal of School Psychology, 82*, 49–69.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Cipriano, C., Strambler, M. J., Naples, L. H., Ha, C., Kirk, M., Wood, M., Sehgal, K., Zieher, A. K., Eveleigh, A., McCarthy, M., Funaro, M., Ponnock, A., Chow, J. C., & Durlak, J. (2023). The state of evidence for social and emotional learning: A contemporary meta-analysis of universal school-based SEL interventions. *Child Development*, *00*, 1–24. https://doi.org/10.1111/cdev.13968