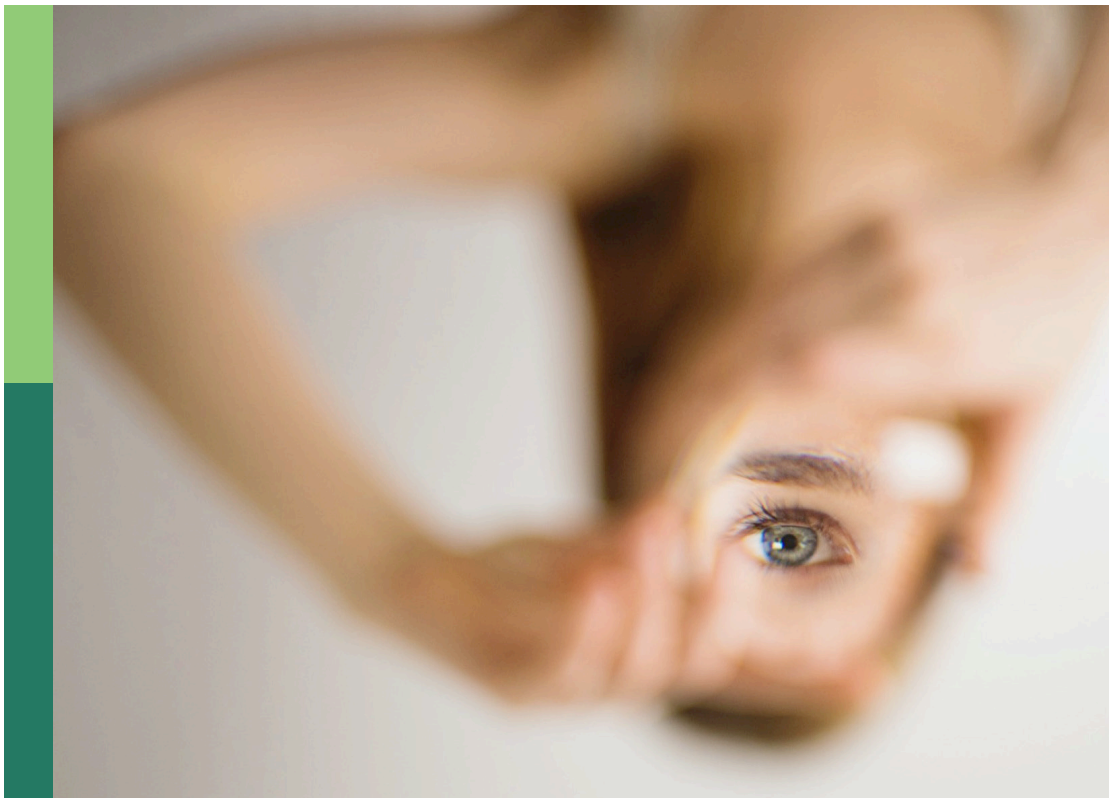
frontiers | Research Topics

Implementation of social and emotional learning interventions in applied settings: Approaches to definition, measurement, and analysis

Edited by
Stephanie M. Jones, Celene Domitrovich and Andres Molano

Coordinated by
Sophie Barnes

Published in
Frontiers in Psychology
Frontiers in Education



ISSN 1664-8714
ISBN 978-2-8325-3542-4
DOI 10.3389/978-2-8325-3542-4

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Implementation of social and emotional learning interventions in applied settings: Approaches to definition, measurement, and analysis

Topic editors

Stephanie M. Jones — Harvard University, United States
Celene Domitrovich — Georgetown University, United States
Andres Molano — University of Los Andes, Colombia

Topic coordinator

Sophie Barnes — Harvard University, United States

Citation

Jones, S. M., Domitrovich, C., Molano, A., Barnes, S., eds. (2023). *Implementation of social and emotional learning interventions in applied settings: Approaches to definition, measurement, and analysis*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-3542-4

Table of contents

frontiers | Frontiers in Psychology

TYPE Editorial
PUBLISHED 08 September 2023
DOI 10.3389/fpsyg.2023.1281083

OPEN ACCESS

EDITED AND REVIEWED BY
Pamela Bryden,
Wilfrid Laurier University, Canada

*CORRESPONDENCE
Stephanie M. Jones
✉ jonesst@gse.harvard.edu

RECEIVED 21 August 2023
ACCEPTED 29 August 2023
PUBLISHED 08 September 2023

CITATION
Barnes SP, Domitrovich CE and Jones SM (2023) Editorial: Implementation of social and emotional learning interventions in applied settings: approaches to definition, measurement, and analysis.
Front. Psychol. 14:1281083.
doi: 10.3389/fpsyg.2023.1281083

Editorial: Implementation of social and emotional learning interventions in applied settings: approaches to definition, measurement, and analysis

Sophie P. Barnes[1], Celene E. Domitrovich[2] and Stephanie M. Jones[1]*

[1]Harvard Graduate School of Education, Harvard University, Cambridge, MA, United States, [2]Early Childhood Innovation Network, Department of Psychiatry, Georgetown University Medical Center, Washington, DC, United States

KEYWORDS

implementation, social and emotional learning (SEL), intervention, social and emotional learning (SEL) skills, measurement, schools

Editorial on the Research Topic
Implementation of social and emotional learning interventions in applied settings: approaches to definition, measurement, and analysis

Implementation matters for SEL intervention effectiveness

More than two decades of meta-analytic research documents the effectiveness of social and emotional learning (SEL) interventions for improving social-emotional competencies and longer-term academic outcomes, behavioral functioning, and mental health (Durlak et al., 2011, 2022; Jones et al., 2021; Cipriano et al., 2023). Implementation research suggests that outcomes are more robust when interventions are implemented with adherence to their intended model (Durlak and DuPre, 2008). In a meta-analysis of 213 studies of SEL interventions, programs implemented with fidelity produced greater improvements in children's outcomes than studies that reported challenges with implementation (Durlak et al., 2011).

There are a number of ways to characterize and measure implementation. Dane and Schneider (1998) defined five dimensions (adherence, exposure, quality, participant responsiveness, and program differentiation) of program integrity that have remained a part of most current definitions of implementation outcomes (Proctor et al., 2011). The term fidelity has emerged over time as a broader term with adherence and dosage as indicators within that dimension (Century et al., 2010; Proctor et al., 2011). Most studies of SEL program implementation assess fidelity or dosage while fewer focus on quality or participant responsiveness (O'Donnell, 2008; Berkel et al., 2011).

Several conceptual frameworks have been developed to illustrate the multiple factors at various ecological levels that influence the implementation process (Wandersman et al., 2008; Damschroder et al., 2009; Meyers et al., 2012). Domitrovich et al. (2008) developed a three-level ecological framework for organizing factors that relate specifically to the implementation of school-based interventions: macro-level factors (e.g., policies

and financing, community capacity and empowerment), school-level factors (e.g., organizational functioning, school and classroom culture, and climate), and individual-level factors [e.g., psychological functioning (burnout and self-efficacy) and perceptions of and attitudes toward the intervention]. Several studies have empirically validated the importance of multi-level factors as predictors of SEL implementation (e.g., see Malloy et al., 2015; Domitrovich et al., 2019; Musci et al., 2019; Cramer et al., 2021).

Studies that include fidelity data often report variability at both the individual level between implementers and at the organizational level across the settings (e.g., schools) where programs are delivered, suggesting the need to more deeply understand associations between implementation and outcomes by examining potentially relevant individual, school, and community factors (Durlak and DuPre, 2008). While research underscoring the importance of implementation fidelity for SEL interventions is growing, it is rarely the primary focus of SEL research and is typically not measured or described in sufficient detail. However, we believe that some of what is relatively equivocal in research on SEL programs (e.g., Jones et al., 2019) can be addressed by documenting and understanding implementation with greater precision and depth. In this Research Topic, authors addressed a wide variety of topics including: (1) approaches to measuring SEL implementation and engagement, (2) applying innovative quantitative and qualitative methods, (3) investing in partnerships with practitioners, (4) multi-level factors predicting implementation, (5) equity in defining and measuring SEL and its implementation, and (6) understanding implementation in global contexts (see Table 1 for an overview of all the manuscripts included in this issue). In the sections below we describe key themes and ideas across the many papers in this Research Topic.

Measuring SEL implementation and engagement

Two papers described comprehensive implementation measurement strategies that monitored implementation of both the intervention and the support system. Choles et al. developed a two-level conceptual implementation framework and aligned measures to capture fidelity of implementation of a mindfulness-focused SEL program for pre-school children. The model is innovative because it focuses on program implementation supports for teachers (e.g., coaching) as well as teacher implementation of the curriculum in the classroom. Martinsone et al. describe their approach to creating a system for monitoring the implementation of a comprehensive school-based mental health program for elementary and middle school students that included universal curriculum lessons as well as home-based activities. The implementation supports included teacher training and the development of school teams whose members helped ensure program fidelity and quality. The program was delivered in six countries and the authors paid close attention to capturing cultural adaptations as part of fidelity monitoring.

A number of papers proposed new approaches to measuring SEL implementation. Wu et al. designed an approach to capture nuanced features of implementation of non-curricular, flexible approaches organized as brief activities across SEL domains (Mindfulness and Brain Games in this study) in humanitarian settings. The paper presents three dimensions of dosage: quantity (how much), duration (for how long), and temporal pattern (how often). This approach can capture (1) how often activities targeting the same SEL domain are repeated and (2) how many activities are implemented before at least one activity is attempted from each available SEL domain—providing insight into patterns of intervention implementation and exposure, in addition to quantity and duration. Devlin et al. noted that most implementation measures focus on the implementer even though children's engagement during implementation likely influences children's outcomes. The authors developed a four-step protocol designed to capture active child engagement by observing children's behavior. The protocol focused on identifying points of active child engagement, operationalizing and measuring those dimensions, and analyzing the data by linking child engagement to other meaningful variables. Bodrova et al. discuss the importance of play as a context for SEL development during early childhood and the challenges of monitoring the fidelity of this activity when it is a component of an SEL intervention. They argue that play components of SEL interventions need to be made "visible" and that nuanced measures of play need to be developed so research can isolate the important characteristics of play that predict positive social and behavioral development over time. All three papers underscore the need to consider implementation as a multi-faceted, dynamic process that requires attention to multiple aspects of implementation (e.g., implementer and recipient, multiple intervention components, implementation context, etc.).

Applying innovative quantitative and qualitative methods

Studies employed a variety of innovative qualitative and quantitative analytic approaches. Two studies used latent profile analysis to create descriptive profiles (or categories) of implementation and explore associations between teacher/classroom profiles and children's outcomes. Zhao et al., used measures of dosage, adherence, quality of delivery, and student engagement to identify three latent profiles of implementation (high, moderate, and low). Classrooms with moderate- and high-level implementation practices showed significantly higher gains in student outcomes than those with low-level implementation. Similarly, Gómez et al. identified two latent profiles: below average implementation and above average implementation using measures of teacher responsiveness (teacher evaluation of the training sessions) and amount of exposure to implementation supports (ratings of coaching and time spent with coach). Teachers in the below average profile were less responsive to training and received less support than teachers in the above average profile. Using propensity scores, the authors found that more experienced teachers and teachers reporting lower levels of burnout were more likely to implement the intervention as intended.

Integrating SEL and youth participatory action research (YPAR), a form of critical participatory action research (CPAR), represents another novel and promising methodological approach (Meland and Brion-Meisels). In YPAR, youth are full participants

TABLE 1 Overview of manuscripts included in the SEL implementation Research Topic.

	Method			Age Group				Location		Focus
	Theoret.	Qual.	Quant.	ECE	Elem.	M.S./H.S.	Adults	U.S.	Global	
Bodrova et al.										Theory and predictors
Braun et al.										Theory and predictors
Choles et al.										Measurement and methods
Devlin et al.										Measurement and methods
Dyson et al.										Educator perspectives and partnerships
Gómez et al.										Measurement and methods
Grant et al.										Educator perspectives and partnerships
Harker Roa et al.										Theory and predictors
Hunter et al.										Theory and predictors
Lin et al.										Educator perspectives and partnerships
Martinsone et al.										Measurement and methods
McCoy and Hanno										Theory and predictors
Meland and Brion-Meisels										Measurement and methods
Partee et al.										Educator perspectives and partnerships
Spacciapoli et al.										Educator perspectives and partnerships
Thierry et al.										Theory and predictors
Ulla and Poom-Valickis										Theory and predictors
White et al.										Theory and predictors
Wu et al.										Measurement and methods
Zhao et al.										Measurement and methods

10.3389/fpsyg.2023.1281083

in the research process and seen as the experts of their own lives and contexts. The authors propose a set of four core commitments as the mechanisms of YPAR that nurture SEL (e.g., democratic participation that centers youth expertise) and conceptualize SEL implementation integrity as adherence to a set of commitments rather than fidelity to a specific set of activities. This approach can provide more flexible ways to think about implementation that centers youth empowerment and voice.

We were also pleased to receive several papers using qualitative methods that focus on understanding individual perceptions and experiences. Using interviews, focus groups, and observations, Dyson et al. explore educators' views on SEL in a rural, high-needs elementary school setting. While educators "bought-in" to SEL, they reported lack of time, lack of preparation and development, home-school disconnection, and pushback from students as significant constraints. Another study used a mixed-methods approach to study the Early Childhood Mental Health Consultation pilot in Virginia. Partee et al. interviewed participants who chose not to participate in ECMHC or opted out after consultation began and conducted focus groups with participants who had sustained engagement. These qualitative papers provide perspectives from a wider set of voices often not included in traditional impact and implementation research including rural settings and those who opt *not* to participate in interventions.

Investing in partnerships with practitioners

Building on the themes of incorporating previously undervalued voices in new ways, a handful of studies centered partnerships with educators. Grant et al. shared findings from a multi-year research-practice partnership (RPP) designed to support SEL implementation in a district. The authors offer key lessons learned related to developing feasible and meaningful implementation measures, identifying structures that can support the collection and use of implementation data to improve practice, presenting data for various audiences, and creating systems for sustainable data use. Spacciapoli et al. conceptualize fidelity as part of an ongoing professional development feedback ecosystem. Teachers record short videos across the school year, review and reflect on their video, and receive targeted feedback from a coach. The method approaches fidelity of implementation as a developmental journey with the expectation that teachers will improve over time and develop a nuanced set of indicators across the school year. These examples demonstrate that mutually beneficial relationships between researchers, practitioners, and other stakeholders can create conditions for iterative cycles of design and testing and the development of sustainable systems of SEL implementation, data collection, and use.

Multi-level factors predicting implementation

A number of studies examined predictors of implementation. Ulla and Poom-Valickis conducted a systematic review and identified four categories of contextual factors that can influence implementation: program support, school, teacher, and student level factors. Their analysis focused on the relative importance of these factors and found that the most frequent statistically significant factors included modeling activities during coaching and teacher-coach working relationship. Teacher burnout was uniquely related to program dosage. In community-based childcare centers, Hunter et al. examined factors that predict the implementation of a comprehensive pre-school program that includes curricular components and teaching practices designed to promote the social-emotional development of young children. They found that baseline teaching practices and responsiveness to intervention (and not teacher education or experience) predicted quality of intervention activities and teaching strategy delivery. Workplace factors (e.g., classroom resources and job satisfaction) predicted multiple features of implementation. Braun et al. also examined workplace factors (i.e., occupational health) that predict the implementation of a universal social-emotional learning program implemented in elementary schools by early career teachers. The authors went a step beyond typical research examining the main effects of implementation predictors to test interactions among factors. They found perceptions of program feasibility moderated the relationship between job stress and implementation quality in unexpected ways. In a unique study of implementation predictors, Thierry et al. conducted secondary analyses of a national dataset that included information on school district and macro-level factors to explore factors associated with teacher and counselor-facilitated delivery of a universal social-emotional learning program conducted in elementary and middle schools. The studies included in this section highlight the multi-level factors that shape the conditions for SEL implementation in schools.

Equity in defining and measuring SEL and its implementation

Three papers included equity as a central component. Lin et al. examined how pre-service educators define SEL. Educators conceptualized SEL as broader than competency-based models. They instead considered SEL as an opportunity to respond to student and community needs, center humanity, and advance social justice. Participants advocated for a co-created, humanizing SEL approach that honors identity, promotes justice, and ultimately can dismantle inequitable systems. YPAR, the approach described by Meland and Brion-Meisels, elevates individual voices and empowers youth to engage in collective action that aims to disrupt systems of inequity and promote positive change in communities. Underlying the approach is a trusting, equitable, reciprocal relationship that remains a central part of the entire YPAR process. Finally, Spacciapoli et al. noted that measures of implementation often leave teachers "in the dark" about observation goals and items as well as implementation strengths and areas for improvement. The authors' transparent approach includes teachers as collaborators in implementation measurement and tracking by engaging them in observing and reflecting on their practice in order to create equitable partnerships.

Understanding implementation in global contexts

The focus of papers in this Research Topic describe SEL efforts across the globe including programming that was delivered in humanitarian settings in Sierra Leone (Wu et al.), community-based programs conducted in Rio de Janeiro, Brazil (McCoy and Hanno), and Colombia (Harker Roa et al.), and a comprehensive school-based mental health program implemented in several European countries (Italy, Latvia, Romania, Croatia, Greece, and Portugal; Martinsone et al.). The two papers describing SEL efforts in South America both discuss the challenges associated with delivering programs in extreme conditions including community violence, forced displacement, and extreme adversity and the importance of cultural adaptation and flexibility. Harker Roa et al. identified cultural adaptation as an implementation enabler that was successful when it was conducted intentionally following a structured protocol. They also found that with sufficient training and support, their parent-focused program could be delivered by paraprofessionals, an example of "task-shifting" that was necessary because of the shortage of mental health professionals in Colombia. McCoy and Hanno also identify culture as a key factor that influenced the delivery of their SEL programming in Brazil. Their perspectives on the importance of this and other macro-level factors including timing and government support came from their work delivering an SEL program in elementary schools and from their review of similar research conducted in low-income, conflict affected settings.

Considerations for the future

Put quite simply, the collection of articles in this Research Topic tell us unequivocally that the quality and quantity of implementation of social and emotional learning programs, strategies, and practices is the cornerstone to their efficacy. But that is not all these innovative and penetrating articles tell us. Indeed, they go beyond relatively "simple" questions of whether and how implementation factors make a difference and stretch our body of knowledge, pushing us to (1) consider new ways of measuring, operationalizing, and analyzing implementation data, (2) incorporate the perspectives of those often not represented in our implementation data, (3) expand our view well beyond our own borders and bring multiple contexts and settings into the broader body of work, and (4) sustain SEL by developing innovative support systems that address both individual and contextual factors that influence the process of implementation. In addition, these papers leave us with some directions to consider for future work. Among many possible directions for future research, we highlight three key questions that build from these papers and, in our view, are central to moving the field forward.

- What are the common and unique, multi-level, predictors of implementation quality and quantity that ultimately represent universal- and program-specific factors?
- Can we embed implementation data collection, reflection, and adapted practice into program design and delivery in ways that create meaningful improvements, rather than considering implementation as an added and separate effort?
- What do cross-cutting and persistent patterns of implementation (i.e., lower than expected dosage, fidelity, and overall engagement) suggest about potential changes required to program design, delivery expectations, and pre-implementation training?

Author contributions

SB: Writing—original draft, Writing—review and editing. CD: Writing—review and editing. SJ: Writing—review and editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Berkel, C., Mauricio, A. M., Schoenfelder, E., and Sandler, I. N. (2011). Putting the pieces together: an integrated model of program implementation. *Prevent. Sci.* 12, 23–33. doi: 10.1007/s11121-010-0186-1

Century, J., Rudnick, M., and Freeman, C. (2010). A framework for measuring fidelity of implementation: a foundation for shared language and accumulation of knowledge. *Am. J. Eval.* 31, 199–218. doi: 10.1177/1098214010366173

Cipriano, C., Strambler, M. J., Naples, L. H., Ha, C., Kirk, M., Wood, M., et al. (2023). The state of evidence for social and emotional learning: a contemporary meta-analysis of universal school-based SEL interventions. *Child Dev.* 1–24. doi: 10.1111/cdev.13968. [Epub ahead of print].

Cramer, T., Ganimian, A., Morris, P., and Cappella, E. (2021). The role of teachers' commitment to implement in delivering evidence-based social-emotional learning programs. *J. Sch. Psychol.* 88, 85–100. doi: 10.1016/j.jsp.2021.08.003

Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., and Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement. Sci.* 4, 50. doi: 10.1186/1748-5908-4-50

Dane, A. V., and Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clin. Psychol. Rev.* 18, 23–45. doi: 10.1016/S0272-7358(97)00043-3

Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., et al. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: a conceptual framework. *Adv. Sch. Mental Health Promot.* 1, 6–28. doi: 10.1080/1754730X.2008.9715730

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

Durlak, J. A., and DuPre, E. P. (2008). Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am. J. Commun. Psychol.* 41, 327. doi: 10.1007/s10464-008-9165-0

Durlak, J. A., Mahoney, J. L., and Boyle, A. E. (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. *Psychol. Bull.* 148, 765. doi: 10.1037/bul0000383

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

Jones, S. M., Brush, K. E., Ramirez, T., Mao, Z. X., Marenus, M., Wettje, S., et al. (2021). *Navigating SEL From the Inside Out: Looking Inside and Across 33 Leading SEL Programs: A Practical Resource for Schools and OST Providers (Revised and Expanded Second Ed. Elementary School Focus).* New York, NY: The Wallace Foundation. Available online at: https://www.wallacefoundation.org/knowledge-center/Documents/Navigating-Social-and-Emotional-Learning-from-the-Inside-Out.pdf (accessed June 15, 2023).

Jones, S. M., McGarrah, M. W., and Kahn, J. (2019). Social and emotional learning: a principled science of human development in context. *Educ. Psychol.* 54, 129–143. doi: 10.1080/00461520.2019.1625776

Malloy, M., Acock, A., DuBois, D. L., Vuchinich, S., Silverthorn, N., Ji, P., et al. (2015). Teachers' perceptions of school organizational climate as predictors of dosage and quality of implementation of a social-emotional and character development program. *Prevent. Sci.* 16, 1086–1095. doi: 10.1007/s11121-014-0534-7

Meyers, D. C., Durlak, J. A., and Wandersman, A. (2012). The quality implementation framework: a synthesis of critical steps in the implementation process. *Am. J. Commun. Psychol.* 50, 462–480. doi: 10.1007/s10464-012-9522-x

Musci, R. J., Pas, E. T., Bettencourt, A. F., Masyn, K. E., Ialongo, N. S., and Bradshaw, C. P. (2019). How do collective student behavior and other classroom contextual factors relate to teachers' implementation of an evidence-based intervention? A multilevel structural equation model. *Dev. Psychopathol.* 31, 1827–1835. doi: 10.1017/S095457941900097X

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K—12 curriculum intervention research. *Rev. Educ. Res.* 78, 33–84. doi: 10.3102/0034654307313793

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., et al. (2011). Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Admin. Policy Mental Health Mental Health Serv. Res.* 38, 65–76. doi: 10.1007/s10488-010-0319-7

Wandersman, A., Duffy, J., Flaspohler, P., Noonan, R., Lubell, K., Stillman, L., et al. (2008). Bridging the gap between prevention research and practice: the interactive systems framework for dissemination and implementation. *Am. J. Commun. Psychol.* 41, 171–181. doi: 10.1007/s10464-008-9174-z

frontiers | Frontiers in Psychology

TYPE Methods
PUBLISHED 26 October 2022
DOI 10.3389/fpsyg.2022.1043001

OPEN ACCESS

EDITED BY
Herman Hay-Ming Lo,
Hong Kong Polytechnic University,
Hong Kong SAR, China

REVIEWED BY
Margrét Sigmarsdóttir,
University of Iceland, Iceland
Khalida Akbar,
Durban University of Technology,
South Africa

*CORRESPONDENCE
Baiba Martinsone
baiba.martinsone@lu.lv

SPECIALTY SECTION
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

RECEIVED 14 September 2022
ACCEPTED 12 October 2022
PUBLISHED 26 October 2022

CITATION
Martinsone B, Stokenberga I and Grazzani I (2022) Monitoring system of implementation of the Promoting Mental Health at Schools (PROMEHS) program.
Front. Psychol. 13:1043001.
doi: 10.3389/fpsyg.2022.1043001

Monitoring system of implementation of the Promoting Mental Health at Schools (PROMEHS) program

Baiba Martinsone[1]*, Ieva Stokenberga[1] and Ilaria Grazzani[2]

[1]Department of Psychology, University of Latvia, Riga, Latvia, [2]Department of Human Sciences for Education "R. Massa", University of Milano-Bicocca, Milan, Italy

Effective school-based mental health programs are a research field with growing interest and great social value. At the stage of development and initial testing of the program, as well as during dissemination, and adaptation in other cultures, it is important that the implementation is carried out in the way that was originally intended. Fidelity or adherence is the most often used concept relating to the extent to which the implemented intervention corresponds to the originally intended program. Therefore, monitoring of the implementation is an essential element necessary to integrate into contemporary evidence-based program. The current paper describes the monitoring system developed for the Promoting Mental Health at Schools (PROMEHS) project. The monitoring was done on both the structural and procedural aspects of the program's implementation, involving the evaluation of five core aspects: fidelity, dosage, quality, responsiveness, and adaptation. This methods article aims to describe the development of the monitoring system and to analyze the strengths of the qualitative-quantitative multi-informant approach in the monitoring of the intervention's implementation. In the future, this would support further research on effectiveness of the PROMEHS program.

KEYWORDS

monitoring, social-emotional learning, mental health, fidelity, dosage, quality, responsiveness, adaptation

Introduction

Monitoring as a key aspect of qualitative/reliable program implementation

Evidence connecting school-based mental health program outcomes with implementation components are increased rapidly during last years, especially in the US (Rojas-Andrade and Bahamondes, 2019). It was supported by growing body of the scientific studies of the implementation field and following recommendation for testing,

implementing, and disseminating evidence-based programs (Domitrovich et al., 2008; Proctor et al., 2013). Several models guiding implementation and monitoring of the implementation fidelity are prevalent in the literature, demonstrating broad scope and variability of components (Fixsen et al., 2021). Focus on defining usable innovation (active components hypothesized to cause effect) and implementations drivers (actors) and stages (procedure) characterize majority of them.

Factors in macro-level, school-level, and individual level can affect successful program implementation in schools (Domitrovich et al., 2008). Several of them has been recognized in the literature and proved to be crucial for the school-based mental health interventions (e.g., teacher competence and support from the head of the school) (Lendrum et al., 2013). Among the factors depending on the implementation process, there are several that should be emphasized: The support system of program providers (i.e., training and assistance during the implementation), compatibility of the innovation, providers' attitudes and beliefs, community resources, and general and specific organizational factors (e.g., Stith et al., 2006; Durlak and DuPre, 2008; Wandersman et al., 2008). Teachers' positive attitude toward the program and understanding of the core components is crucial because it allows to make them necessary adaptation without negative cost for quality and predicts fidelity of the program implementation in the long term (Sørlie, 2021).

It is known that the implementation process is related to the outcomes of programs when their effectiveness is evaluated (e.g., DuBois et al., 2002; Wilson et al., 2003; Durlak et al., 2011). Moreover, faithful replication is even more important when programs are disseminated to use in the field, where development and testing of the program is not the focus. Even a well-developed program could become less effective or even ineffective over time without proper dissemination, introducing it to the potential implementers, support for the acceptance of the program, and investment in its sustainability. Recent study in Norway supports necessity to start implementation monitoring in the early stages of the intervention, because these data predict fidelity of the program in the long term (Sørlie, 2021). Thus, the validity of an intervention should be ensured by consistent monitoring of the implementation process.

There are several implementation components important for monitoring described in the literature (Durlak and DuPre, 2008; Durlak, 2015). The criteria relating to a program's implementation are fidelity (correspondence of the implemented program to the originally intended one), dosage (quantity of delivered content of the intervention), quality (how well the program has been conducted), and the responsiveness of participants. Some authors also note the differentiation between or the extent to which the content and methodology of a program are distinct from other programs as a considerable aspect (Dane and Schneider, 1998; Durlak, 2015). In the recent literature (e.g., Mohr et al., 2014) the necessity to monitor the control group, participation rate, and the representativeness of groups involved, as well as the extent of adaptations to or modifications of the program during the implementation process is also emphasized.

Several components should be included in the monitoring because we do not know which are the most important implementation factors. In previous studies, different components have been found to be the most significant implementation factors. It has been proved that interventions implemented in high fidelity show stronger effect on outcome (Durlak et al., 2011). Recent analysis found that students' exposure (number of classes) and receptiveness (student commitment) are among those with the strongest impact (Rojas-Andrade and Bahamondes, 2019).

In most cases, only few components have been assessed during the monitoring of the implementation of different preventive programs (for a review, see Durlak and DuPre, 2008; Rojas-Andrade and Bahamondes, 2019). Fidelity and dosage are implementation components included in the studies most often, and typically measured quantitatively using self-report data. Responsiveness, in contrast, needs observational data from several informants as commitment to the program is crucial for both instructors and participants.

Durlak (2015) emphasizes that it is not possible to avoid adaptation in field studies and following dissemination. Some of the modifications can be beneficial (e.g., adding culturally relevant material contributing main program aim), but some – negative (e.g., selecting only certain type of activities or shortening the time of activity). It is crucial to document adaptations made during implementations, and to evaluate their value according the aims of the program and implementation context. Moreover, adaptation is typically measured qualitatively, allowing to provide more contextualized information about the implementation.

There are several highly valuable examples when psychometrically sound measures of fidelity are developed for certain programs (e.g., Abry et al., 2015), or intervention systems as School-Wide Positive Behavior Support (Horner et al., 2004). For example, in Norwegian PBIS program (Sørlie et al., 2015) implementation dosage was estimated by percentage of trained school staff, and quality of the implementation has been measured by asking teachers how do they implement support to positive behavior (e.g., "Expected student behavior is consequently encouraged and positively acknowledged"). This teacher behavior addresses one of the core component of the program, and scale composed from several items is useful for self-report or observation. Considering core components of each program procedures and measures should be developed for the monitoring of the implementation process both for faithful replication and evaluation of the possible effect on outcomes.

Nevertheless, evaluating the implementation of a wide spectrum of preventive and intervention programs provides empirical evidence on the key role of appropriate implementation in the success of programs. These factors were considered when the monitoring system of the Promoting Mental Health at Schools (PROMEHS) program was developed.

The Promoting Mental Health at Schools program

The Promoting Mental Health at Schools was developed within the Erasmus + Key Action 3 project co-funded by the European Commission. The project's timeline was from 2019 to August 2022, and it aimed to develop a comprehensive mental health curriculum, implement it, and evaluate its effectiveness. The consortium involved researchers, practitioners, and policymakers from seven European countries: Italy, Latvia, Portugal, Croatia, Romania, Greece, and Malta.

The PROMEHS theoretical framework includes three domains, namely, promoting social-emotional learning (SEL) and resilience and preventing social, emotional, and behavioral problems. This framework was described and substantiated by Cavioni et al. (2020).

The key features of the universal curriculum were based on principles of international research (CASEL, 2020), such as the whole-school approach, evidence-based content, multi-year handbooks, developmental perspectives, teacher training, etc. The capacity of this curriculum was built through teacher training and ongoing assistance, sustaining partnerships with policymakers, and parents' involvement.

The PROMEHS curriculum consists of seven handbooks. Two are for teachers with ready-to-use, step-by-step activity plans for leading pre-school and school students aged from 3 to 18 years. Two handbooks are for both pre-school/primary school students and middle/secondary school students with activities to carry out independently at home or together with their parents. The other three handbooks are for teachers to promote their own mental health, for parents to promote mental health at home, and for supplying recommendations to policymakers.

Since the curriculum was aimed at fostering students' SEL and resilience and preventing social, emotional, and behavioral problems, all these topics were covered in the offered activities. Each activity has the same structure, namely, defined learning outcomes, a clearly defined age group, and a step-by-step activity plan. The activity starts with a story, followed by a discussion, role-play, group work, or another learning strategy. An important part of the activities is reflection. At the end of every activity, a teacher is provided with a brief formative evaluation chart, tips on how to embed the goal into their everyday teaching practices, as well as culturally adapted further resources (lists of books, movies, videos).

The curriculum was implemented in Italy, Latvia, Portugal, Croatia, Romania, and Greece, whereas the University of Malta acted as the external evaluator and was not involved in the development and implementation of PROMEHS. The project's implementation and the evaluation of its effectiveness were carried out in four age groups of students from pre-school to secondary school level (3–6, 8–10, 11–13, and 14–16 years), including disadvantaged children.

The quasi-experimental research design with experimental and control conditions was implemented to evaluate the effectiveness of the program. An integral part of the development and implementation of the PROMEHS program was the monitoring system, which was built with the purposes of ensuring the fidelity and quality of its implementation and of finding out culture-specific practices to develop recommendations for practitioners and educational policymakers (see **Figure 1**). A detailed description of the curriculum and the whole project is available in Cefai et al., 2022a,b.

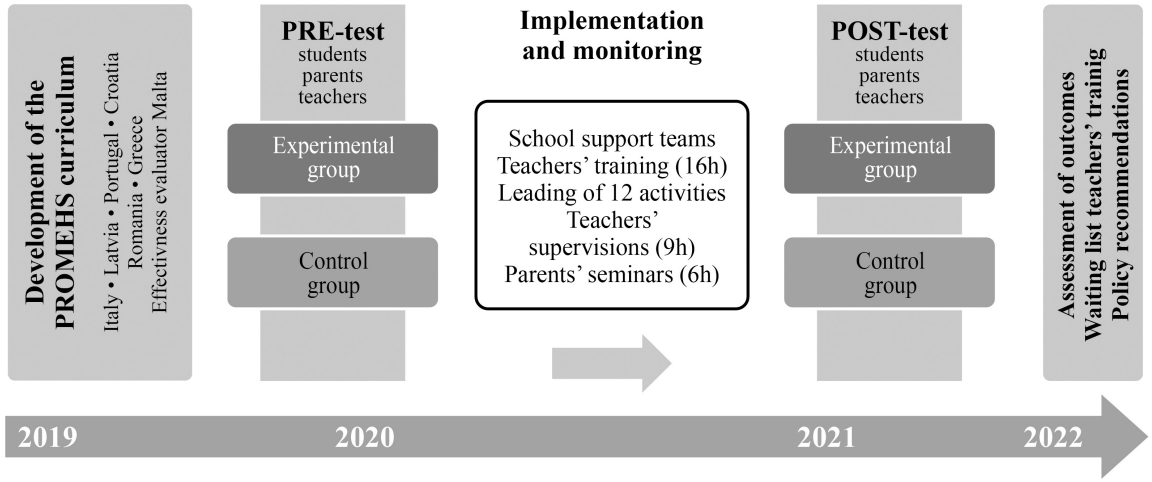


FIGURE 1
Promoting Mental Health at Schools project's timeline and design.

Framework of the Promoting Mental Health at Schools implementation monitoring system

The monitoring of the implementation can be done for diverse purposes, and decisions should always consider the balance between costs and added value. In this case, the purpose of the monitoring, as defined in the project proposal, was to evaluate the quality of the intervention's implementation (1) to ensure the fidelity and quality of its implementation and (2) to find out culture-specific practices for schools to develop recommendations for both practitioners and educational policymakers.

Five dimensions were used for this purpose: fidelity, dosage, quality, responsiveness, and adaptation (Dane and Schneider, 1998; Durlak and DuPre, 2008; Feely et al., 2018).

Fidelity characterizes the extent to which the implemented intervention corresponds to the originally intended program. The fidelity of the implementation of the PROMEHS program was supported by the provision of detailed materials on the content and procedures to be implemented. Both structural (the content to be delivered) and process components (how the content should be implemented) of the program's implementation were described in the PROMEHS materials. Comprehensive and detailed handbooks were developed for each age group, both for teachers and students (Grazzani et al., 2020a,b,c,d), teacher training and a series of supervisions were carried out, and activities for school leaders and parents' meetings were organized in line with the curricula with the aim to increase fidelity or adherence.

Dosage refers to how much of the intervention has been delivered. It has a high potential to be included in effectiveness studies, and therefore it was decided to monitor it as well. In the implementation of the PROMEHS, minimal exposure was defined as 12 activities proportionally covering all three parts of the program, namely SEL, promoting resilience, and preventing behavioral problems.

Quality refers to how well different program components have been implemented. The quality of implementation evaluates the competence of the program providers according to the content and manner of the intervention manual. Organizational factors (e.g., education or qualification requirements) are recognized as a useful way to increase the quality of the intervention. However, its combination with process evaluation is crucial, especially when using external observations (Feely et al., 2018).

Participant responsiveness refers to the degree to which the program stimulates the interest and engagement of participants. Most often, it is the responsiveness of the direct target group (e.g., students) that is measured (Durlak and DuPre, 2008). Considering that the success of the intervention is affected by the involvement of both school and family, a multi-informant approach was used, and all of them–teachers, students, and parents–were treated as participants.

Considering that the model for assessing the fidelity of the PROMEHS project's implementation was developed to provide information on how its implementation may vary across countries and to provide specific recommendations for its implementation in the future, fidelity is supplemented by adaptation assessments. Adaptation refers to changes made to the original program during its implementation (program modification). Previous research (Durlak and DuPre, 2008; Feely et al., 2018) suggests that adaptation should be evaluated separately (rather than as a failure to achieve fidelity) because it could make possible positive contributions to the outcome(s). Culture-specific adaptations can provide important insight into the best implementation practices crucial for the sustainability of the program (Forman et al., 2009) at the national and international levels.

Methodology

Research context and participants

The PROMEHS program was implemented in six European countries in the school year 2020/2021. Initially, it involved 10,209 students, but pre- and post-test evaluations were received from 4,501 participants in the experimental condition and 3,288 participants in the control group, where the evaluators were teachers. Both pre- and post-test parental evaluations were received in relation to the outcomes of 2,394 participants in the experimental group and 2,234 participants in the control group. Student self-reports at the two measure points were obtained from 1,845 students in the experimental group and 1,458 from the control group.

However, monitoring the data collection was not directly related to pre- and post-test data for the effectiveness study. The monitoring sample consisted of experimental condition participants, namely 2,534 students from primary and middle/secondary school (aged nine and older) and 2,868 parents, who provided feedback after their children's participation in the program activities (See **Table 1**).

During the project's implementation, 532 teachers were trained in total, of whom 421 filled out the final evaluation of the program, identifying the strengths and weaknesses of the materials and providing practitioners with their expertise for the further elaboration of the PROMEHS materials.

School support teams were organized in each country, with a range of members from three to eight per country. In sum, there were 29 members, all qualified professionals with specific knowledge and expertise as described in the quality requirements. They organized pre- and post-test data collection, managed teacher training and on-going supervisions, collected qualitative data from teachers, and contributed to developing

TABLE 1 Sample sizes for monitoring the implementation of Promoting Mental Health at Schools (PROMEHS).

Sample size in each informant group by country	Italy	Latvia	Romania	Croatia	Greece	Portugal	Total
School support team members	5	3	5	5	3	8	29
Teachers (training)	192	51	94	64	63	68	532
Teachers (3rd supervision)	140	54	93	34	45	55	421
Parents	296	728	704	273	250	617	2,868
Student (post-test)							
Primary school	154	209	235	26	139	346	1,109
Middle and high school	236	353	328	55	50	403	1,425

detailed recommendations for the further elaboration of the PROMEHS program's materials and its implementation in diverse contexts (e.g., remotely).

The number of participants was different between the countries due to different response rate (in groups of students and their parents), and the teachers' involvement (in some countries, more teachers participated in the PROMEHS than it was planned in the research protocol).

The monitoring system and measures

Several steps were taken for the development of the monitoring system as recommended in the literature (Feely et al., 2018): (1) defining the purpose and scope of the monitoring; (2) identifying the components for assessment; (3) developing the tools for assessment; (4) collecting data during the project's implementation; and (5) analyzing the data.

The monitoring system was developed by the first two authors of the paper in collaboration with project partners. Considering the purpose of the monitoring and principles of the program, a multi-component and multi-informant approach was chosen.

Detailed implementation procedures were developed following recommendations in literature (Domitrovich et al., 2008; Proctor et al., 2013; Fixsen et al., 2021). Essential components were identified based on the PROMEHS program and considering the importance of monitoring its implementation in all stages of the field trial, starting with the development of the school support team, providing teacher training and supervisions, followed by providing activities at schools and parents' meetings, as well as the management of the implementation. This approach was also based on the indicators supporting program sustainability (proposed by Elias, 2010), such as building a support system for teachers involving personnel outside the school's staff, providing on-going professional development for teachers, as well as integrating the program into the regular curriculum.

Following examples in the previous studies (e.g., Sørlie et al., 2015) and guidelines (e.g., Proctor et al., 2013) indicators for each monitoring dimension were identified, discussed, and selected. Items corresponding to each indicator were developed based on the balance between yes/no, Likert-type scale and open-ended questions. Item formulations were matched with the context in which different informants were expected to respond (e.g., quantitative scales for items about the clarity of the task after the training; open-ended questions for reflections about successes and difficulties experienced after each activity).

All procedures characterizing process components (how the program should be implemented) were discussed with project partners and translated into six national languages. All measures were piloted with the appropriate target audience, tested and corrected for clarity, discussed with partners, and translated into six national languages. Finally, monitoring data were collected during all stages of the field trial and analyzed before the results were presented to project partners.

The Ethics Committee for Humanities and Social Sciences Research Involving Human Participants of the University of Latvia granted permission for the research on 12 December 2019.

Results

The PROMEHS monitoring system resulting from the procedures and measures developed for monitoring purposes can be seen in **Table 2**.

The PROMEHS monitoring procedures are presented in **Table 3**.

The essential components of the monitoring were selected following curriculum and research protocols and were described in the framework section. Materials and guidelines were available for the school support team, sharing information on data collection and training, meeting with school leaders, teacher training and supervisions, activities at schools, and parents' meetings.

Aiming to evaluate the fidelity of the program's implementation, data were collected after teacher training, during activities at school, and during supervisions from both school support team members and teachers. Measures included categorical scales (e.g., online, on-site, or mixed training) as well as continuous scales with a Likert-type scale (e.g., the question in **Supplementary Annex 6** measuring the extent to which an activity from the handbook was implemented completely).

TABLE 2 Monitoring system for evaluating the implementation of Promoting Mental Health at Schools (PROMEHS).

Program components	Materials and procedures	Monitoring dimensions and indicators	Measures and informant	Informant
School support team	Development of the school support team Sharing procedures and materials for data collection and teacher training	Quality: qualification requirements; competence in teaching materials and procedures Fidelity: consequent implementation of the program components	**Supplementary Annex 1**. Table for keywords **Supplementary Annex 2**. Checklist about competence in materials and procedures **Supplementary Annex 11**. Checklist of activities	School support team members
Teacher training	Curriculum Handbooks for pre-school/primary school Handbooks for middle/high secondary school	Fidelity: 16 h training was organized (time, place, duration, number of participants); adherence to agenda Responsiveness: perceived Teachers' responsiveness and acceptance of the content	**Supplementary Annex 3**. Questionnaire of evaluation of teacher training	School support team members
		Quality: perceived usefulness of the training, sufficiency of information, understanding of the task to be performed, confidence in ability to carry out this program Quality: competence in teaching materials	**Supplementary Annex 4**. Questionnaire of evaluation of teacher training **Supplementary Annex 5**. Table for keywords	Teachers
Supervisions	Guidelines	Fidelity: 3 × 3 h supervisions were organized (time, place, duration, number of participants) Adaptation: best practices and changes made in the program	**Supplementary Annex 7**. Supervision summary	School support team members
Activities at schools	Handbooks for pre-school/primary school Handbooks for middle/high secondary school Handbook for teachers	Fidelity: program implemented as described in the manual Dosage: number of activities implemented Quality: observed evidence of students' competence, perceived effect on self-development in teaching SEL. Responsiveness: teachers' perception of the students' responsiveness; usefulness of handbook for teachers	**Supplementary Annex 6**. Teacher self-reflection form **Supplementary Annex 8**. Final evaluation form in 3rd supervision	Teachers
		Quality: evaluation of the teaching process Responsiveness: using students' handbook	**Supplementary Annex 9**. Student survey	Students
Meeting with school leaders	Guidelines for policymakers	Fidelity: meeting was organized	**Supplementary Annex 11**. Checklist of activities	School support team members
Parents' meetings	Curriculum Handbook for parents	Responsiveness: evaluation of the parents' meetings; evaluation of the handbook for parents and students' handbook	**Supplementary Annex 10**. Parent survey	Parents

It was planned that dosage would be measured during the implementation: each teacher should have filled in a self-reflection form (**Supplementary Annex 6**) after each activity and taken it with him/her to their supervision. However, the actual number of the implemented activities was reported by teachers at the post-test stage. Considering that testing the PROMEHS program's effectiveness took place in Europe during one of the waves of the COVID-19 pandemic, there were several threats to the filling-in of these forms. It can be assumed that some teachers gave up on implementing the program because of the stressful context of the COVID-19 pandemic and related epidemiological measures.

However, the sufficient variance of dosage, including significant deviations from the pre-planned length (min. 12 activities), provides the opportunity to test the dosage effect in relation to the effectiveness of the PROMEHS program.

The quality of implementation evaluates the skill and competence of the program providers according to the content and methods of the PROMEHS intervention manual. There were specific competence requirements for school support team members, and regular meetings related to testing, training, and supervisions were organized and reported. Several support materials were provided to strengthen their competence in PROMEHS materials (**Supplementary Annex 1**) and management of the field trial (**Supplementary Annex 11**).

The quality of evaluation addresses school support team members (self-reports) and teachers (self-reports and student reports). School support team members

TABLE 3 The Promoting Mental Health at Schools (PROMEHS) monitoring procedures.

(1). *Developing the school support team*
Description of competencies of the schools' support team members:
- appropriate qualification, desire to be a psychologist;
- do not work in the same school;
- familiar with the mental health concept and school environment;
- experience of working with groups;
- good knowledge of the PROMEHS materials;
- understanding of research principles and ethics.

(2). *Sharing procedures and materials of data collection with school support team members.*
Full information about data collection is provided, school support team members fill in the Checklist about competence in materials and procedures (**Supplementary Annex 2**).
School support team members organize an introductory visit to every school (experimental and control), where they
- discuss planned activities and the necessary conditions (e.g., collecting of informed consent forms from parents, the need for computers for surveys, the need for a specific number of students, clarifying the participant coding system, making an agreement for its storage in accordance with research ethics, etc.);
- inform/remind teachers to collect permission forms from parents for data collection;
- arrange a time for the other three meetings with the parents of the experimental group.

(3). *Sharing training procedures and materials with school support team members*
Full information (principles, agenda) about the teacher training is provided, school support team members fill in Table for keywords (School support team member) (**Supplementary Annex 1**) and Checklist about competence in materials and procedures (School support team member) (**Supplementary Annex 2**).

(4). *Meeting with parents*
School support team members organize an introductory meeting with parents to establish contact and introduce the project.
- During the introductory meeting, parents receive general information about the project as a whole, planned activities, and the opportunity to receive materials; give their agreement for participation and testing; and have the opportunity to answer questionnaires.
- No materials are distributed there yet!

The aim of the following meetings is to motivate parents in the experimental condition to participate in PROMEHS activities at home (using the student and parent handbooks) and share and discuss parenting practices in order to promote the mental health of their children.

(5). *Pre-test. Data collection in experimental and control schools*
Paper-pencil or electronic data collection (students', parents', and teachers' questionnaires). Student surveys are filled out in the presence of school support team members. Data collected from paper-pencil surveys must be filled into an online form (by a school support team member or researcher).

(6). *Training of teachers at the experimental schools*
School support team member leads the 16 h training for teachers according to the agenda. Table for keywords (teacher) (**Supplementary Annex 5**) can be used as support material for teachers to help them become more familiar with the material.
The evaluation will be done in written form at the very end of the teachers' training and in a reflective cycle. See questionnaire in Questionnaire of evaluation of teacher training (Teacher) (**Supplementary Annex 4**).
The aims of this evaluation are to:
- receive feedback about the quality of the training in terms of usefulness; and
- xmonitor teachers' readiness to implement the PROMEHS program.

Additionally, after school support team members collect filled-in questionnaires, two questions must be addressed in a reflective cycle:
What have I achieved during the training?
What questions remained unanswered?
After the evaluation, the school support team member reviews the responses (both questionnaires and reflective cycle) with the national team and makes a general analysis of the training fidelity, acceptance of agenda, and teacher responsiveness, as well as any adaptations of the program. See questionnaire in Questionnaire of evaluation of teacher training (School support team member) (**Supplementary Annex 3**).
Teachers are instructed to start their intervention immediately after the training for 12 weeks, with at least one activity per week. After the first activity in class, the student and parent handbooks are given to students.
After each PROMEHS activity in class, we ask the teachers to review and reflect on their practice individually using the Teacher self-reflection form (Teacher) (**Supplementary Annex 6**). The teachers should prepare for a supervision by making written notes after each activity.

(7). Supervisions of teachers (3 × 3 h) in the experimental schools
1st supervision (approximately 2–3 weeks after teacher training),
2nd supervision (approximately 4–6 weeks after the 1st supervision),
3rd supervision (approximately 4–6 weeks after the 2nd supervision).
- All supervisions have the same structure and content.
- Additionally, the 3rd supervision includes the final evaluation.
- Between supervisions, a support team member communicates with the school *via* e-mail or another platform.
- During a supervision, the school support team member makes notes according to guidelines in Supervision summary (School support team member) (**Supplementary Annex 7**).

Principles:
Emotional support: "Thank you for your involvement."
Plan for (rules of) the meeting: "We have met to discuss the situation, answer questions, and plan the next activities. This is not about control."
Confidentiality: "Outside this group, each person can only share personal information with others," "Let each participant express his/her opinion," "Each person will have an opportunity to speak," "Every participant is asked to speak from their own perspective," "If there appear to be some problems, we will support each other and share responsibilities to find a solution for your school."

(Continued)

TABLE 3 (Continued)

Space for reflection: How do I feel? What is my attitude? What are my personal concerns? What resources do I have?
During the supervision, the main questions are discussed together:
Success. How did we succeed during this time?
Challenges. What has been challenging?
Adaptation. If changes were made to the program, what were they and why were they made?
Continue to develop teachers' understanding of the PROMEHS approach to the promotion of mental health by answering questions about the content of handbooks.
During the supervision, the school support team member writes down specific observations on best practices and how the material has been adapted. After the supervision, a summary must be done. See Supervision summary (School support team member) (**Supplementary Annex 7**).
During the 3rd supervision, the usual content is supplemented by an evaluation. Teachers are asked to fill in Final evaluation form in 3rd supervision (Teacher) (**Supplementary Annex 8**) and comment with questions of their own choice.

(8). *Meetings with school leaders of the experimental schools*
Information for the administration about PROMEHS and how to support the intervention.

(9). *Meetings with parents of students of the experimental schools*
Responsiveness evaluation of the parents and students. Parent survey (Parents) (**Supplementary Annex 10**).

(10). *Data collection for monitoring the quality of the implementation*
If possible, the student survey should be carried out by school support team members among students who participated in the intervention. Use Student survey (Students) (**Supplementary Annex 9**) to evaluate how students felt and what the class environment was like.
After the last supervision, a meeting with all school support team members should be organized (for a reflection on the process/about themselves). Work on the final report, including a brief summary of quantitative data from the student survey, and on finding out the best practices and cultural adaptations is also done at this point.
As a result, a written report with specific initial recommendations should be developed

- to improve the teachers' training
- to improve the handbooks
- for educational policy

(11). *Post-test. Data collection in experimental and control schools*
Paper-pencil or electronic data collection (students', parents', and teachers' questionnaires). Student surveys are filled out in the presence of school support team members.
Data collected from paper-pencil surveys must be filled into an online form.
For support, it was recommended to use Checklist of activities (School support team member) (**Supplementary Annex 11**).

evaluated their own competence in teaching materials and procedures (**Supplementary Annex 2**) before starting on the implementation. Teachers evaluated their understanding of the task to be performed and their confidence in their ability to carry out this program, as well as their competence in the related teaching materials (**Supplementary Annex 4**). At the end of the intervention, students were asked to evaluate the manner in which the program was implemented (**Supplementary Annex 9**).

Considering the principles of the PROMEHS program emphasizing collaboration between school and family, teachers, students, and parents were all treated as participants, and their levels of responsiveness were measured. Teacher responsiveness was estimated after the teacher training and was evaluated by school support team members (**Supplementary Annex 3**). Students' responsiveness was evaluated by teachers after each activity using a special self-reflection form (**Supplementary Annex 6**). Teacher responsiveness, according to support materials for their own mental health, was assessed during the last supervision (**Supplementary Annex 8**). Responsiveness measures were included in the post-test survey: students (aged nine and older) were asked to evaluate the usefulness of the student handbooks (**Supplementary Annex 9**), and parents were asked to evaluate the usefulness of parents' meetings and the handbooks for parents and students as well (**Supplementary Annex 10**).

Adaptation refers to changes made to the original program during its implementation (program modification, reinvention). As a result, adaptation was integrated as an independent dimension with high value in the monitoring system in all stages of implementation, and qualitative data were collected. Adaptations were observed in several sources. Teachers filled out a self-reflection form (**Supplementary Annex 6**) after each activity and characterized what was changed and why, and they were also asked to describe their successes and any difficulties. This information gave a comprehensive picture of the adaptations made, reasons for these, the most successful practices, as well as activities where changes or updates would be welcomed. School support teams collected best practices and difficulties during supervisions (**Supplementary Annex 7**) and summarized them after the implementation to develop national-level recommendations for the implementation of the program.

Discussion

A program can be evaluated as effective if it is implemented as intended. The fidelity of the intervention can be significantly increased by developing materials on content, what to implement, and the manner in which it should be implemented. The PROMEHS program filled this requirement by providing comprehensive, ready-to-use handbooks for teachers, students, families, and policymakers. Added value is related to the

inclusion of content and process components in the monitoring system, where several materials can be used as tools to familiarize oneself with the content of the program while following the guidelines described in detailed procedures. Additionally, the monitoring measures include a checklist to ensure the implementation of all program components.

The PROMEHS monitoring system covers all the most relevant components of the intervention, starting with the development of the school support team, followed by teacher training, supervision, activities at schools, parents' meetings, and student feedback. Specific requirements for the school support team members are described in the monitoring procedures to ensure quality. Moreover, these prepared the professional continuing education course for pedagogues so that PROMEHS could be maintained sustainably.

Providing support for people involved in the implementation of the program was recognized as a crucial principle, and therefore materials in the form of checklists were included to make the monitoring system user-friendly and helpful, allowing participants to practice self-monitoring during the implementation. A checklist on the content of handbooks allowed implementers to use it both as a training task and as a piece of evidence on how familiar both the school support team and teachers were with the provided materials. Considering that the usefulness and user-friendliness of tools can increase readiness to use monitoring tools, further research is needed on the applicability of the monitoring system after the project. It can be assumed that teacher self-report forms (e.g., **Supplementary Annex 6**) can be used to strengthen their self-reflection skills; however, further research is needed before confirming such a recommendation.

The implementation of the program is not always compatible with an aim to explore factors affecting its success or failure. A research strategy combining monitoring data and pre- post-test data allows the testing of a hypothesis about possible mediating or moderating effects of implementation characteristics on program outcomes. It can be assumed that diverse informants can evaluate different qualities of the program's implementation, allowing key predictors of program efficacy and necessary support for program providers to be explored.

Both quantitative and qualitative (according to Dowling and Barry, 2020) data were collected for the assessment of fidelity and quality, responsiveness and dosage were measured quantitatively, and adaptation was evaluated exclusively using qualitative data. The reflections of teachers and observations of school support teams during supervisions provided an opportunity to explore nuanced and highly applicable experiences on how certain topics and activities were perceived in different countries, age groups, and backgrounds.

It is known that observational data are more reliable than self-reported data, and the reliability of measures can be strengthened by combining different data sources. It is important not to limit the monitoring only to activities in the main target group (students), since the intervention included activities focused on teachers, parents, and school-leaders as well. Direct observation was not included in the present monitoring system; however, this limitation was addressed by collecting multi-informant data from the program's implementers (teachers), students, and their parents, as well as from the support team members, who provided teacher training and on-going supervisions and parents' meetings. This strategy allows the implementation of the community engagement principle to be monitored as well, which is crucial to the sustainability of the program.

Implications, limitations, and conclusion

The findings of this study highlight the importance of including several aspects often generally described as fidelity but which, nevertheless, allow the implementation process of a program to be evaluated from different angles, namely, dosage, responsiveness, quality, and adaptation.

This study also emphasizes the role of monitoring every aspect of implementation regarding both its content and its procedure. Moreover, it emphasizes the importance of building scientifically sound and, at the same time, user-friendly monitoring procedures in order not to overwhelm participants with data collecting but rather to support them during the implementation process. This study contributes to the field providing an elaborated framework for monitoring of implementation of different interventions. This supports both researchers and practitioners in developing, implementing, assessing, and sustaining the best possible practice in the intervention.

The strengths of this monitoring system are its observation of both content and process with scientifically sound dimensions, thus covering the whole spectrum of implementation, its collection of qualitative and quantitative data, and its use of a multi-informant approach. PROMEHS implementation during the COVID-19 pandemic allowed to document adaptation related with remote learning and computer mediated instructions.

The system also has some limitations. First, no direct observation of the teacher's competence and interaction with students during the activities was available, limiting conclusions about the quality of the implementation of the program. Observation would be beneficial for providing more contextualized feedback and helping to develop teacher competence in instructing SEL. However, this can partly be offset with observations during supervisions when teachers interact with each other, which can also be used as an indicator of the manner in which they implement principles of the PROMEHS program. This limitation was partially neutralized by collecting evaluations from all groups of participants, direct observation of the responsiveness during teacher training was

done by school support teams, whereas responsiveness of the students was evaluated by their parents. In the future, it would be useful to add direct observation during class activities to estimate quality of the implementation, as well as responsiveness of the students. Second, there was no monitoring of the control group. One critical point that was emphasized in the literature was the necessity to control other possible interventions in the control group. However, the COVID-19 pandemic context, with the related social distancing and remote learning, provided an opportunity to overcome this limitation since, due to the restrictions of the pandemic, the control group did not receive any alternative interventions. This naturally alleviated the necessity to monitor it.

Data availability statement

The original contributions presented in this study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee for Humanities and Social Sciences Research Involving Human Participants of the University of Latvia on 12 December 2019. Written informed consent to participate in this study was provided by the participants or their legal guardian/next of kin.

Author contributions

BM: lead writer, arranging the research in Latvia, contributing to the development of the monitoring system, and collecting data. IS: contributing to the development of the monitoring system and to writing. IG: a key contribution to designing the research and revising the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This study was conducted within the EU-funded Erasmus + KA3 research project "Promoting Mental Health at Schools" (No. 606689-EPP-1-2018-2-IT-EPPKA3-PI-POLICY).

Acknowledgments

We recognize contribution of all partners of the PROMEHS consortium in facilitating the development of the monitoring system. Partners represent the University of Milano-Bicocca, Italy; University of Latvia, Latvia; University of Malta, Malta; University of Rijeka, Croatia; Stefan cel Mare University of Suceava, Romania; University of Lisbon, Portugal; University of Patras, Greece.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.1043001/full#supplementary-material

References

Abry, T., Hulleman, C. S., and Rimm-Kaufman, S. E. (2015). Using indices of fidelity to intervention core components to identify program active ingredients. *Am. J. Eval.* 36, 320–338. doi: 10.1177/1098214014557009

CASEL (2020). *Evidence-based social and emotional learning programs: CASEL criteria updates and rationale.* Available Online at: https://casel.org/wpcontent/uploads/2021/01/11_CASEL-Program-Criteria-Rationale.pdf [accessed August 20, 2022].

Cavioni, V., Grazzani, I., and Ornaghi, V. (2020). Mental health promotion in schools: A comprehensive theoretical framework. *Int. J. Emot. Educ.* 12, 65–82.

Cefai, C., Camilleri, L., Bartolo, P., Grazzani, I., Cavioni, V., Conte, E., et al. (2022a). "Evaluating the effectiveness of the PROMEHS programme un improving students' and teachers' social and emotional competence resilience and mental health," in *Proceedings of the seventh monograph in resilience and health, centre for resilience and socio-emotional health*, (Malta: University of Malta). doi: 10.13140/RG.2.2.26562.25288

Cefai, C., Camilleri, L., Bartolo, P., Grazzani, I., Cavioni, V., Conte, E., et al. (2022b). Thex effectiveness of a school-based, universal mental health programme in six European countries. *Front. Psychol.* 13:925614. doi: 10.3389/fpsyg.2022.925614

Dane, A. V., and Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clin. Psychol. Rev.* 18, 23–45. doi: 10.1016/s0272-7358(97)00043-3

Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., et al. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: A conceptual framework. *Adv. Sch. Ment. Health Promot.* 1, 6–28. doi: 10.1080/1754730X.2008.9715730

Dowling, K., and Barry, M. M. (2020). Evaluating the implementation quality of a social and emotional learning program: A mixed methods approach. *Int. J. Environ. Res. Public Health* 17:3249.

DuBois, D. L., Holloway, B. E., Valentine, J. C., and Cooper, H. (2002). Effectiveness of mentoring programs for youth: A meta-analytic review. *Am. J. Community Psychol.* 30, 157–197.

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

Durlak, J. A., and DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am. J. Community Psychol.* 41, 327–350.

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

Elias, M. (2010). Sustainability of social-emotional learning and related programs: Lessons from a field study. *Int. J. Emot. Educ.* 2, 17–33. doi: 10.1111/mono.12425

Feely, M., Seay, K. D., Lanier, P., Auslander, W., and Kohl, P. L. (2018). Measuring fidelity in research studies: A field guide to developing a comprehensive fidelity measurement system. *Child. Adolesc. Soc. Work J.* 35, 139–152.

Fixsen, A. A. M., Aijaz, M., Fixsen, D. L., Burks, E., and Schultes, M.-T. (2021). *Implementation frameworks: An analysis*. Chapel Hill, NC: Active Implementation Research Network.

Forman, S. G., Olin, S. S., Hoagwood, K. E., Crowe, M., and Saka, N. (2009). Evidence-based interventions in schools: Developers' views of implementation barriers and facilitators. *School Ment. Health* 1, 26–36.

Grazzani, I., Agliati, A., Conte, E., Cavioni, V., Gandellini, S., Ornaghi, V., et al. (2020a). *PROMEHS: The curriculum to promote mental health at school. Activities handbook for teacher, preschool and primary school (Ages 3-10)*. Milano: Mimesis Publishing.

Grazzani, I., Agliati, A., Conte, E., Cavioni, V., Gandellini, S., Ornaghi, V., et al. (2020b). *PROMEHS. The curriculum to promote mental health at school. Students handbook (Ages 3-10)*. Milano: Mimesis Publishing.

Grazzani, I., Agliati, A., Conte, E., Cavioni, V., Gandellini, S., Ornaghi, V., et al. (2020c). *PROMEHS. The curriculum to promote mental health at school. Activities handbook for teacher, secondary and high school (Ages 11-18)*. Milano: Mimesis Publishing.

Grazzani, I., Agliati, A., Conte, E., Cavioni, V., Gandellini, S., Ornaghi, V., et al. (2020d). *PROMEHS. The curriculum to promote mental health at school. Students handbook (Ages 11-18)*. Milano: Mimesis Publishing.

Horner, R. H., Todd, A. W., Lewis-Palmer, T., Irvin, L. K., Sugai, G., and Boland, J. B. (2004). The School-Wide Evaluation Tool (SET): A research instrument for assessing school-wide positive behavior support. *J. Posit. Behav. Interv.* 6, 3–12. doi: 10.1177/10983007040060010201

Lendrum, A., Humphrey, N., and Wigelsworth, M. (2013). Social and emotional aspects of learning (SEAL) for secondary schools: Implementation difficulties and their implications for school-based mental health promotion. *Child Adolesc. Ment. Health* 18, 158–164. doi: 10.1111/camh.12006

Mohr, D. C., Ho, J. C., Hart, T. L., Baron, K. G., Berendsen, M. A., Beckner, V., et al. (2014). Control condition design and implementation features in controlled trials: A meta-analysis of trials evaluating psychotherapy for depression. *Transl. Behav. Med.* 4, 407–423. doi: 10.1007/s13142-014-0262-3

Proctor, E. K., Powell, B. J., and McMillen, J. C. (2013). Implementation strategies: Recommendations for specifying and reporting. *Implement. Sci.* 8:139. doi: 10.1186/1748-5908-8-139

Rojas-Andrade, R., and Bahamondes, L. L. (2019). Is implementation fidelity important? A systematic review on school-based mental health programs. *Contemp. Sch. Psychol.* 23, 339–350. doi: 10.1007/s40688-018-0175-0

Sørlie, M.-A. (2021). Structural, cultural and instructional predictors essential to sustained implementation fidelity in schools: The School-Wide Positive Behavior Support Model (SWPBS). *Int. J. Educ. Res. Open* 2:100082. doi: 10.1016/j.ijedro.2021.100082

Sørlie, M.-A., Ogden, T., and Olseth, A. R. (2015). Preventing problem behavior in school through school-wide staff empowerment: Intervention outcomes. *World J. Educ. Res.* 2:117. doi: 10.22158/wjer.v2n2p117

Stith, S., Pruitt, I., Dees, J., Fronce, M., Green, N., Som, A., et al. (2006). Implementing community-based prevention programming: A review of the literature. *J. Prim. Prev.* 27, 599–617.

Wandersman, A., Duffy, J., Flaspohler, P., Noonan, R., Lubell, K., Stillman, L., et al. (2008). Bridging the gap between prevention research and practice: The interactive systems framework for dissemination and implementation. *Am. J. Community Psychol.* 41, 171–181.

Wilson, S. J., Lipsey, M. W., and Derzon, J. H. (2003). The effects of school-based intervention programs on aggressive behavior: A meta-analysis. *J. Consult. Clin. Psychol.* 71, 136–149.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 15 December 2022
DOI 10.3389/fpsyg.2022.1023505

OPEN ACCESS

EDITED BY
Andres Molano,
University of Los Andes,
Colombia

REVIEWED BY
Nicholas John Shaman,
University of Houston–Clear Lake,
United States
Shauna Tominey,
Oregon State University,
United States

*CORRESPONDENCE
Leah J. Hunter
ljh61@pitt.edu

This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

RECEIVED 19 August 2022
ACCEPTED 22 November 2022
PUBLISHED 15 December 2022

CITATION
Hunter LJ, Bayly BL, Bierman KL, Welsh JA and Gest JM (2022) Predicting school readiness program implementation in community-based childcare centers. *Front. Psychol.* 13:1023505. 10.3389/fpsyg.2022.1023505

Predicting school readiness program implementation in community-based childcare centers

Leah J. Hunter[1]*, Benjamin L. Bayly[2], Karen L. Bierman[3], Janet A. Welsh[4] and Julia M. Gest[3]

[1]Department of Psychology, University of Pittsburgh, Pittsburgh, PA, United States, [2]Department of Agricultural Economics, Sociology, and Education, College of Agricultural Sciences, The Pennsylvania State University, University Park, PA, United States, [3]Department of Psychology, The Pennsylvania State University, University Park, PA, United States, [4]Edna Bennett Pierce Prevention Research Center, College of Health and Human Development, The Pennsylvania State University, University Park, PA, United States

Introduction: Targeted curricular interventions can increase preschool program quality and boost children's academic and social–emotional readiness skills, but variable funding and weak organizational infrastructure in many community-based childcare centers may reduce the effective implementation of these programs.

Method: This study examined individual teacher and workplace predictors of the REDI program implementation, a targeted school readiness program that was adapted to support delivery in childcare centers. REDI was delivered by 63 teachers in 37 community-based childcare centers with center directors serving as local implementation coaches.

Results: Results showed that individual teacher factors (e.g., teaching skills and receptivity to intervention consultation) predicted the quality with which REDI activities and teaching strategies were delivered, and workplace factors were important predictors across multiple implementation indicators.

Discussion: Practice and policy implications for improving intervention implementation and corresponding program quality in childcare centers are highlighted.

KEYWORDS

school readiness, implementation, childcare center, social–emotional learning, literacy skills, preschool, teacher coaching

Introduction

High-quality early childhood education (ECE) promotes school readiness skills and fosters long-term school success, with heightened benefits for preschool children from low-income families (Yoshikawa et al., 2013; Phillips D. A. et al., 2017). Access to preschool programs has increased over the past two decades, as has the inclusion of social–emotional learning elements in early learning programs (Bryant et al., 2021), but wide variations in program quality remain a significant concern (Ackerman and

Sansanelli, 2010; Donoghue, 2017; Pianta and Hamre, 2020). Community-based childcare centers are especially variable, with average teaching quality levels significantly lower than those in publicly managed programs such as Head Start and school district prekindergarten programs (Burchinal et al., 2008; Dowsett et al., 2008; Hillemeier et al., 2013; Bassok et al., 2016a).

Research conducted during the past two decades suggests that preschool program quality can be enhanced by enriching classrooms with evidence-based curricula and providing teachers with corresponding professional development support and coaching (Yoshikawa et al., 2013; Phillips D. A. et al., 2017). However, this research has focused almost exclusively on publicly-funded Head Start and public prekindergarten programs (McCormick et al., 2015; pre-k; Phillips D. A. et al., 2017). Childcare centers have more variable organization and funding structures than publicly-funded programs, with fewer resources and regulatory supports, which may reduce their capacity to adopt new evidence-based programming (Ackerman and Sansanelli, 2010; Bassok et al., 2016a; Whitebook et al., 2018; McCormick et al., 2022). Indeed, initial efforts to "scale up" evidence-based preschool programs more broadly in community-based childcare settings have encountered significant implementation challenges (Baker et al., 2010; Yurdon et al., 2016). Research is needed to better understand the factors that facilitate or impede the effective implementation of evidence-based programming in childcare contexts to ensure these programs can be brought to scale successfully. The current mixed methods study addressed this issue by exploring teacher and workplace factors associated with the quality of implementation of an evidence-based school readiness program (the Research-based Developmentally Informed [REDI] program) in childcare classrooms.

The need to improve preschool programming in childcare centers

Childcare centers serve one-third of children attending preschool in the United States (NCES, 2020). Unlike Head Start or school district pre-K programs, childcare centers represent separate, diverse entities operating within a de-centralized system that lacks common standards for accreditation or operation (Ackerman and Sansanelli, 2010; Bassok et al., 2016a). Childcare centers operate under varied management structures, ranging from for-profit corporations and non-profit cooperatives to small, independently-owned and operated businesses (Ackerman et al., 2009). They are often under-resourced, with average teacher salaries and benefits well below those provided in publicly-funded programs (Whitebook et al., 2018; Johnson et al., 2019). Correspondingly, childcare teachers serving preschool children often have lower levels of formal education and training than teachers in publicly-funded programs, and typically experience higher levels of stress and job dissatisfaction (Bassok et al., 2016a; Whitebook et al., 2018). They leave their jobs at high rates and move to more well-funded positions when they can (Zaslow et al., 2010).

Not surprisingly, when compared on similar measures of observed preschool teaching quality, childcare centers show mean levels of emotional support and cognitive stimulation that are significantly lower than those documented in Head Start or school district pre-K classrooms (Dowsett et al., 2008; Hillemeier et al., 2013; Bassok et al., 2016b; McCormick et al., 2022). Evidence-based strategies that have proven effective at improving quality in publicly-funded preschool settings may also enhance the quality of childcare centers; however, these strategies are rarely studied in childcare contexts (McCormick et al., 2015; Phillips D. A. et al., 2017). leaving unanswered questions about the ways in which childcare teacher or workplace factors might affect implementation quality of the strategies.

Implementing evidence-based strategies that boost the school readiness of preschoolers

Current research suggests that the most effective strategies for improving preschool program quality and boosting child school readiness outcomes utilize two approaches (Yoshikawa et al., 2013; Phillips D. A. et al., 2017). First, effective intervention approaches provide teachers with professional development support and coaching in high-quality teaching practices designed to boost emotional support, enriched language use, and instructional quality (Hamre et al., 2012; Pianta et al., 2020). Second, some effective approaches also increase child learning opportunities in the classroom by enriching daily programming with manualized, skill-specific curriculum components that provide lesson plans and sequenced learning activities (Jenkins and Duncan, 2017; Nguyen et al., 2018). These curriculum components are typically domain-specific (e.g., focused on early literacy, mathematics, or social–emotional skills) and are especially effective for boosting child skills in the targeted domains relative to more global curricular approaches (Jenkins and Duncan, 2017). The ultimate goal of these two approaches is to elevate levels of social–emotional support and cognitive stimulation in the classroom, and thereby accelerate the pace of growth in school readiness skills (Jones and Bouffard, 2012; Maier et al., 2022).

Key markers of implementation quality for interventions that use both approaches include: (1) completing the sequenced lesson plans as written, reflecting *adherence* to intervention guidelines, and (2) using the prescribed teaching strategies while delivering lessons and interacting with children in the classroom, reflecting *quality* in program delivery and generalized use of the recommended teaching strategies (Gearing et al., 2011). A limited research base suggests that the predictors of implementation quality may vary depending upon the facet studied (e.g., curriculum delivery adherence or teaching strategy quality) as described in the next section.

Predictors of implementation quality in evidence-based preschool intervention

Domitrovich et al. (2008) proposed a multilevel framework to describe the determinants of school-based program implementation. Determinants included individual-level characteristics of the teachers who implement the intervention (such as teacher training and experience) and also workplace factors (such as school climate and administrative leadership) that provide a support system for the intervention. In the following sections, we review evidence regarding the association of teacher characteristics and workplace factors with adherence and quality of school readiness intervention implementation in preschool classrooms.[1]

Teacher characteristics

Professional background.

Teacher education has been fairly well-studied as a predictor of preschool program implementation quality. Two studies have linked teacher education to intervention adherence. Teachers with an early childhood education background conducted more Banking Time dyadic intervention sessions to target children's disruptive behavior compared to teachers without an early childhood specialization (Williford et al., 2015). The authors speculated that having a degree focused on early childhood increased uptake of the teacher-child relationship-focused intervention. In the second study, teachers with master's degrees used the BEST in CLASS behavior management strategies more often than teachers with high school or associate degrees (Sutherland et al., 2018). However, only the BEST in CLASS (and not the Banking Time) intervention documented links between teacher education and the quality with which the intervention was delivered, possibly due to the demands of the 2-tiered intervention (Sutherland et al., 2018). Teacher education was not consistently related to implementation adherence or quality in multiple interventions that included classroom curricular lessons and strategies, including the Bloom Language Curriculum (Phillips B. M. et al., 2017), Building Bridges (Baker et al., 2010), Second Step (Wenz-Gross and Upshur, 2012), and the Head Start REDI (*Re*search-based, *D*evelopmentally *I*nformed) program delivered in Head Start centers (Domitrovich et al., 2009). These findings suggest that teacher education levels are generally not predictive of implementation for interventions that include guided classroom curricula, but they may affect uptake of new teaching strategies in more intensive intervention programs that focus on student–teacher interaction quality.

1 See Supplementary material for a table listing a description of each intervention reviewed, the predictors of implementation included in the study, and their relation to implementation.

Teaching skills.

From a conceptual standpoint, foundational teaching skills, such as positive classroom management skills and proficiency in instructional support may foster high-quality preschool program implementation by reducing child disruptiveness and increasing student engagement. Further, teaching skills may accelerate a teacher's capacity to adopt new teaching strategies by allowing teachers to build upon their higher baseline levels of competence and confidence (Gage et al., 2015). Supporting this hypothesis, pre-intervention observations of teacher-student interaction quality significantly predicted the quality of delivery of the preschool Second Step curriculum (Wenz-Gross and Upshur, 2012), the BEST in CLASS intervention (Sutherland et al., 2018), the Bloom Language Curriculum (Phillips B. M. et al., 2017), and the Getting Ready for School program (Marti et al., 2018). Pre-intervention teacher-student interaction quality also predicted adherence (number of lessons taught) in the Second Step curriculum study (Wenz-Gross and Upshur, 2012), but was not related to adherence in the other studies.

Responsiveness to intervention

Researchers have suggested that teachers put more effort into delivering an intervention when they feel comfortable with the intervention approach and are open to consultation and feedback about their implementation quality (Domitrovich et al., 2008). Consistent with this expectation, positive attitudes toward the intervention (measured *via* pre-intervention teacher self-report) predicted the quality of teacher delivery of a language-literacy skills intervention (Zucker et al., 2013) and the Bloom Language Curriculum (Phillips B. M. et al., 2017). Similarly, both Domitrovich et al. (2009) and LoCasale-Crouch et al. (2016) found that teachers who were more responsive to and enthusiastic about the coaching they received showed higher levels of quality when using the teaching strategies that were a focus of the intervention. Teacher receptivity to the intervention also predicted adherence in delivery of the Bloom Language Curriculum (Phillips B. M. et al., 2017). Conversely, teacher concerns about the intervention predicted lower adherence in delivering the Building Bridges curriculum activities (Baker et al., 2010).

In summary, prior studies generally suggest little impact of teacher education on implementation adherence or quality, more consistent support for baseline teaching skills as a facilitator of implementation quality (and sometimes adherence), and consistent associations between teacher receptivity toward the intervention and both implementation quality and adherence. With few exceptions, the studies cited examined intervention implementation in Head Start or public pre-kindergarten contexts, leaving unknown questions about the value of these teacher characteristics as predictors of implementation in childcare settings.

Workplace factors

In contrast to teacher characteristics, workplace factors are rarely studied as predictors of preschool program implementation,

but they may be key to understanding challenges associated with diffusing evidence-based programs in under-resourced childcare centers characterized by variable and generally low levels of infrastructure support. In the conceptual framework outlined by Domitrovich et al. (2008), school-level factors may influence intervention implementation either directly by the degree to which the intervention is supported at the administrative level, or indirectly, though the impact of the workplace on teacher morale. Several features in this domain distinguish childcare centers from publicly-supported preschools: classroom resources, teacher job satisfaction, organizational learning support, and workplace challenges (Dennis and O'Connor, 2013).

Classroom resources

The early learning standards of the National Association for the Education of Young Children (2002) specify the importance of adequate classroom resources to support the implementation of high-quality early education practices. Child-care centers vary considerably in their access to these resources due to the limited and fragmented funding streams they rely on (Ma et al., 2021). We found only one prior study that examined classroom resources as a predictor of evidence-based program implementation. Wenz-Gross and Upshur (2012) assessed the classroom environment with the Early Childhood Environment Rating Scale–Revised. This composite rating reflected the classroom space and furnishings, books and communication supports, activity centers and materials, and program schedule. It supported teacher adherence to the delivery of the Second Step program but was unrelated to implementation quality (Wenz-Gross and Upshur, 2012). The authors speculated that being in a more well-resourced classroom reduced obstacles to intervention delivery and boosted teacher feelings of efficacy and motivation to invest in improved programming.

Job satisfaction

Prekindergarten teaching positions pay less, offer fewer benefits (including less time off), and provide teachers with fewer opportunities for professional development opportunities than similar positions in public schools (Whitebook et al., 2018; Johnson et al., 2019). Teachers in these settings view their jobs as lower status jobs (Morrissey et al., 2007) and often express higher levels of stress and job dissatisfaction than their counterparts working in public schools (Bassok et al., 2016a; Whitebook et al., 2018). Prior research suggests that when teachers feel more supported, satisfied, and effective at their jobs, they implement a new program more effectively, whereas job-related stress and burnout are associated with reduced implementation adherence and quality (Ransford et al., 2009; Baker et al., 2010).

Organizational learning

Research suggests that a key characteristic of high-quality ECE programs is a high level of support for staff professional development and program improvement efforts (Ehrlich et al., 2016). Referred to as organizational learning (Bryk et al., 1999), this construct reflects the attitudes and efforts made by school administrators and staff to increase competencies, explore innovations, and engage in activities that can enhance program quality. Whereas public schools and Head Start programs provide teachers with professional development opportunities and dedicated time, most childcare centers lack the financial and staffing resources to do so (Whitebook et al., 2018). Teacher perceptions of school-based professional development supports (e.g., provision of coaching) predicted implementation dose and quality of a new elementary school program (Ransford et al., 2009), suggesting that organizational learning may function similarly to support new preschool programming.

Workplace challenges

Conceptually, working in a well-run center characterized by predictable schedules, stable staffing, and strong collegial working relationships should increase teacher willingness and capacity to invest effort in new program implementation (Domitrovich et al., 2008). Center directors with the resources and administrative skills necessary to support the effective day-to-day management of the organization are well-positioned to provide the oversight and support needed for intervention implementation (Baker et al., 2010). However, community-based childcare center directors are often significantly under-resourced and belabored by the day-to-day challenges of recruiting and retaining high-quality teachers, attracting families, and monitoring and complying with state regulations. These kinds of workplace challenges are demoralizing and stressful for teachers and can interfere with their ability to provide consistent programming, as well as decrease their motivation to invest in new programming (Baker et al., 2010; Hunter and Bierman, 2020). Supporting this hypothesis, Baker et al. (2010) found that teacher perceptions of a supportive, collegial, and fair work climate predicted adherence, reflected in the number of Building Bridges intervention activities delivered.

Scaling school readiness: Predicting implementation of REDI in childcare centers

Originally evaluated in Head Start centers, the REDI program was recently adapted for use in childcare centers. REDI is an evidence-based, multi-component curricular enrichment program targeting social–emotional and early literacy skills. The foundation for REDI is a social–emotional curriculum, Preschool PATHS (Domitrovich et al., 2007), which includes scripted lessons targeting social–emotional skills. REDI added a daily interactive reading program that uses books linked to the PATHS lessons designed to support oral language skill development, along with a Sound Games program to promote phonological awareness and alphabet center activities to build print awareness. A randomized controlled trial of REDI in Head Start classrooms produced positive effects on teaching quality and child outcomes in both social–emotional and language-literacy domains (Bierman et al., 2008) with sustained child benefits through ninth grade (Bierman et al., 2021).

Adaptations to REDI were made to accommodate the less-centralized structure of childcare centers and facilitate program scalability. First, given that online PD can reduce the burden of training (Powell et al., 2010; Piasta et al., 2012), the REDI training sequence for childcare teachers was reduced to two face-to-face workshop days supplemented with four online learning modules that teachers could review at their convenience. In addition, recognizing the difficulties childcare centers face in accessing professional coaches, REDI used a novel model of PD support that trained center directors to serve as coaches for their teachers. Directors attended the teacher trainings and were also provided with a one-day workshop and three online modules demonstrating the REDI coaching model (for more detail, see Hunter and Bierman, 2020). Directors held regular meetings with teachers during the implementation year to provide supportive and corrective feedback. Directors were supported by REDI Consultants who visited centers once a month to provide technical assistance and answer questions.

The current study

The current study explored teacher characteristics and workplace factors that may have affected the implementation of REDI in childcare centers. Implementation outcomes included: (1) adherence, reflecting the number of REDI lessons and activities that were delivered, (2) quality of REDI curriculum delivery, reflecting the quality with which the lessons and activities were delivered, (3) quality of generalized teaching strategies, reflecting the overall use of REDI-prescribed teaching practices in the classroom, and (4) plans to sustain REDI implementation in the future. Based upon prior research linking teaching attributes to implementation quality (Phillips D. A. et al., 2017; Marti et al., 2018; Sutherland et al., 2018), it was hypothesized that *teacher characteristics,* especially baseline teaching skills and teacher receptivity to the intervention would predict implementation quality, including the quality of REDI curriculum implementation and the more generalized use of REDI teaching strategies in the classroom. Given that workplace factors may be especially relevant for program completion in childcare centers which are often under-resourced, it was hypothesized that *workplace factors*, including classroom resources, job satisfaction, organizational learning, and workplace challenges would predict implementation adherence, reflecting the amount of the REDI program that was delivered. We also explored the possibility that workplace factors would affect implementation quality. Finally, we explored the degree to which teacher characteristics and workplace factors might affect enthusiasm for and plans to continue REDI implementation in subsequent years.

Materials and methods

Design overview

During three successive years (2015–2017), licensed childcare centers serving preschool children in ten Pennsylvanian counties were sent emails describing the study. To be included, centers had to have: (1) at least one classroom that served at least five children of prekindergarten age, (2) a full-time director who could serve as a program coach, (3) an organized, regular daily schedule of activities (e.g., not a drop-in center or unstructured day care), and (4) not currently be using a formal curriculum-based social–emotional learning program. Each year enrolled childcare centers were stratified by county and size (number of classrooms) and then randomized at the center level to either the intervention or "usual practice" control condition. This study focused on the centers randomized to the intervention condition. Teachers provided information about their education and teaching experience, and classroom observations were conducted prior to intervention initiation to assess baseline teaching skills. A certified REDI trainer (the fifth author) provided intervention training to center directors and teachers in October and coordinated the intervention delivery. Classroom teachers implemented the intervention through April, with local coaching provided by their center directors. Regional REDI consultants (experienced educators trained in the REDI program and coaching process) visited centers twice monthly for the first 2 months and monthly thereafter. They met with the center director to discuss teacher progress and offer coaching support. They also observed REDI lessons and rated the quality of intervention delivery. Post-intervention classroom observations were collected in May. The guidelines for the ethical conduct of research developed by the American Psychological Association were followed throughout this study, and all procedures were approved by the University's Institutional Review Board.

Participants

Preschool teachers ($N = 63$) from 37 childcare centers provided data for the current study. Teachers were predominantly female (98%) and White (89%; 5% Biracial; 4% Black; 2% Latinx; < 1% Asian). They varied in age between 22 and 60 years of age ($M = 35.6$; SD = 10.7). A small sample of teachers co-taught (9.5%) and all teachers were in classrooms with at least five children who were eligible to start kindergarten in the following year. The 37 center directors were 100% female and predominantly White (90%; 7% Black; 3% multiracial). Directors' ages ranged from 25 to 65 years ($M = 41$ years, SD = 8.73), they had between 1 and 21 years of experience as directors ($M = 6$ years, SD = 6.30) and varied in education (23% Associate degree, 26% Bachelor's degree, 52% some graduate training or degree). Two directors were replaced during the study, one just before the intervention period began, and the other mid-intervention. The majority of centers only had one participating preschool classroom ($n = 30$; centers with two preschool classrooms $n = 7$).

Measures

Predictors of implementation included teacher characteristics and workplace factors. Measures of implementation included

adherence, quality of the REDI program delivery, quality of the generalized REDI teaching strategies, and plans for future REDI implementation.

Teacher characteristics

Teacher characteristics that served as predictors of program implementation included education, experience, baseline teaching practices, and receptivity to intervention.

Experience and education

Teachers self-reported the number of years they had taught in a preschool classroom ($M = 7.24$ years, SD = 5.96, range 1–24 years). Teachers also self-reported their highest level of education on an 8-point scale (1 = less than high school, 0%; 2 = high school diploma or GED, 1.8%; 3 = Some training beyond high school but not a degree, 19.3%; 4 = one-year vocational training certificate, 5.3%; 5 = two-year Associate's degree, 14.0%; 6 = four-year Bachelor's degree, 33.3%; 7 = some graduate coursework, 19.3%; 8 = graduate degree, 7.0%; $M = 5.44$, SD = 1.60).

Baseline teaching practices

The quality of teacher–student interactions was evaluated during pre-intervention observations using the Classroom Assessment Scoring System for Pre-K (CLASS Pre-K; Pianta et al., 2008a). Trained research staff who were naïve concerning the intervention and intervention/ control group center assignment observed teachers for four 20-min periods, rating teacher-student interactions after each period on the ten items of the CLASS Pre-K. Items were rated using a 7-point scale. Three items reflected teacher efforts to promote learning and support children's academic development (concept development, quality of feedback, and language modeling) and were averaged across the four observation periods to represent *Instructional Support* ($\alpha = 0.94$; $M = 2.69$; SD = 0.98). Four items reflected teacher efforts to promote prosocial behaviors and social–emotional development (positive climate, negative climate, teacher sensitivity, and regard for students' perspectives) and were averaged across the observation periods to represent *Emotional Support* ($\alpha = 0.79$; $M = 5.74$; SD = 0.83).

Observers also rated the quality of classroom language use at baseline using the Classroom Language and Literacy Environment Observation (CLEO; Holland Coviello, 2005). CLEO observations occurred separately from the CLASS observations but were conducted by the same observers and often on the same day. They involved 20-min sessions during book reading, free play, and snack/lunch time. During each period, observers coded all teacher utterances directed toward children, identifying teacher directives/commands, questions, and other comments/statements. A total *Non-directive Talk* score was calculated by summing all questions and comments/ statements across the three settings ($\alpha = 0.51$; $M = 22.42$; SD = 6.08). In addition, after each 20-min observation, research assistants used a 5-point scale (1 = *never*; 5 = *always*) to rate the quality of teacher's talk in areas of vocabulary use, elaboration, cognitive challenge, and decontextualized language. These scores were averaged across items and across the three settings to reflect *Richness of Talk* in the classroom ($\alpha = 0.91$; $M = 1.81$; SD = 0.69).

Preliminary analyses showed that these different dimensions of teaching practices showed moderate to high levels of inter-correlation ($r = 0.33$–0.78; $p < 0.05$) and similar patterns of association with implementation. Hence, an overall score reflecting *Positive Teaching Practices* was calculated by standardizing and averaging scores from the *Instructional Support* and *Emotional Support* dimensions of the CLASS Pre-K and the *Non-directive Talk* and *Richness of Talk* dimensions of the CLEO ($\alpha = 0.76$; $M = 0.00$; SD = 0.80; range = −2.21–1.85).

Receptivity to intervention

Center directors and REDI consultants each rated teachers' receptivity to intervention. Ratings measured the frequency of positive teacher responses (1 = *almost never*; 5 = *almost always*) during coaching sessions with directors and consultation sessions with REDI consultants. Director ($M = 4.72$; SD = 0.32; range = 3.67–5.00) and consultant ratings ($M = 4.43$; SD = 0.58; range = 2.43–5.00) were positively skewed and were within one point of one another 87% of the time. Director and consultant scores were averaged to create an overall rating of teacher intervention receptivity ($r = 0.25$, $p = 0.09$; $M = 4.59$; SD = 0.37; range = 3.19–5.00).

Workplace factors

Workplace factors included measures of classroom resources, teacher job satisfaction, organizational learning, and workplace challenges.

Classroom resources

Observers documented classroom resources using the CLEO Literacy Environment Inventory (LEI). They rated 16 items describing the number of books and writing materials in the classroom, and three items describing the degree to which literacy-related activities were displayed (e.g., "is there an area that is designated just for book reading?"; "how many varieties of paper are available for writing?"). Items were rated on a 3-point scale (0–2). We standardized and averaged the literacy environment and literacy activities scores to create an overall classroom resources variable ($\alpha = 0.76$; $M = 0.00$; SD = 0.92; range = −2.95–2.06).

Job satisfaction

Teachers rated their overall job satisfaction using an 11-item scale developed by Gill et al. (2007). Using a 5-point scale (1 = *very dissatisfied*; 5 = *very satisfied*) teachers rated their satisfaction with their salary and benefits, workload, role, and job responsibilities

(α = 0.85). Scores were averaged across the 11-items (M = 2.71; SD = 0.60; range = 1.00–4.00).

Organizational learning

Teachers completed a 7-item rating scale to describe their center's orientation toward innovation and professional development. Items reflected staff orientation toward and support for program improvement (e.g., "In this early childhood program teachers and other professional staff… are encouraged to stretch and grow; are continually learning and seeking new ideas; respect those who take the lead in program improvement efforts; Bryk et al., 1999). Items were rated on a 5-point scale (*strongly agree* to *strongly disagree*) and summed for a total score (M = 3.15; SD = 0.67; α = 0.90; range = 2.00–4.00).

Workplace challenges

At the end of the intervention year qualitative interviews were conducted with teachers to discuss their experiences with the REDI intervention (see Hunter and Bierman, 2020 for a full report of these interviews). Participants were asked several questions about their workplace and colleagues. Following recommendations from Creswell and Plano Clark (2011) and Campbell et al. (2013), the first author and a graduate student undertook an iterative process for qualitative coding by clustering quotes from the interviews into thematic categories, discussing discrepancies, and reaching codebook consensus. The final codes reflected workplace challenges in the areas of staffing, scheduling, and professional development/supervision, as well as personal stress/overwork (independent coding kappa = 0.72; see Hunter and Bierman, 2020). A variable was calculated representing the proportion of workplace challenges mentioned relative to overall comments made about the workplace (M = 0.60; SD = 0.17; range = 0.24–0.90).

REDI program implementation outcomes

Implementation of the REDI program included adherence (percentage of lessons and activities that were delivered) and program quality (quality of REDI teaching strategy use during REDI curriculum delivery and also generalized throughout the day). We also evaluated teacher and director plans to sustain REDI implementation in future years.

Program adherence

Adherence was measured using teacher reports of the REDI lessons completed during each week. Adherence scores were calculated for each component of REDI (i.e., Preschool PATHS lessons, interactive reading, Sounds Games, and alphabet center) and aggregated to reflect a teacher's overall adherence to the program delivery plan. Adherence was calculated at the individual teacher level; adherence rates from teachers who left centers mid-year before having the opportunity to fully implement REDI were excluded from analyses. The number of lessons delivered over the course of the year were summed and divided by the total number of REDI lessons to calculate overall adherence as a percentage of the program that was delivered per teacher (M = 73.15%; SD = 28.13%; range = 3.86–98.00%).

Quality of delivery

Two aspects of REDI program delivery quality were measured: the quality with which teachers delivered the components of REDI (i.e., quality of REDI curriculum delivery) and the quality with which they used REDI teaching strategies throughout the day (i.e., quality of REDI teaching strategy use). In both areas, quality of delivery was rated by trained REDI consultants who observed teachers regularly throughout the school year. Consultants visited centers twice per month during the first 2 months of the academic year, and once per month thereafter. At each visit, consultants made efforts to watch teachers delivering the various components of REDI (e.g., Preschool PATHS, interactive reading, sound games, and alphabet center). They rated curriculum delivery quality using a 7-point scale ranging from *poor* to *exemplary* implementation. Scores were averaged across all four components to create an overall score (α = 0.86; M = 5.23; SD = 1.06; range = 2.00–6.86).

REDI consultants also rated teachers on the quality with which they used REDI teaching strategies in generalized ways throughout the day. Specifically, REDI consultants rated teachers on their demonstration of each of 5 teaching strategies (positive classroom management, sensitivity and responsiveness, emotion communication and support, positive limit-setting, and richness of talk) using a 5-point scale. Sample items included: "teacher encourages children to communicate how they feel, particularly when they are upset. He/she validates the children's feelings when they are expressed" (emotion communication and support) and "the teacher is physically and mentally available to children in the setting" (sensitivity and responsiveness). An overall score for quality of REDI teaching strategy use was created by averaging scores over time across the five core REDI teaching strategies (α = 0.94; M = 4.22; SD = 0.54; range = 2.78–5.00).

Intentions for future REDI implementation

At the end of their intervention year, teachers and directors completed a 5-item scale describing their personal enthusiasm for the continued use of the REDI program in the future, and the degree to which they and their center colleagues value the program and support continued use in the future. Items were rated on a 5-point scale (from not at all to very much) and averaged to represent overall intentions for future program use (α = 0.86; M = 3.56; SD = 0.87; range = 2.00–5.00).

Plan for analysis

As an initial step, we accounted for missing data by conducting multiple imputation (MI) with ten iterations with all variables of interest included in the model using SPSS version 26. MI is a Monte Carol technique where missing data is replaced with

estimated values based on the available data and is preferable over single imputation methods of accounting for missing data (e.g., mean replacement; Graham, 2009). All analyses were conducted with imputed data.

We first conducted bivariate correlations with all variables of interest to gain a better understanding of associations between the independent and dependent variables. We then conducted a set of four multiple regressions predicting the four REDI implementation outcomes: adherence, quality of REDI curriculum delivery, quality of REDI teaching strategies used, and intentions for future REDI implementation. Predictors included teacher characteristics (teacher experience and education, positive teaching practices, and receptivity to intervention) and workplace factors (classroom resources, job satisfaction, organizational learning, and workplace challenges). Because a small number of teachers shared directors, regression analyses included robust standard errors to account for clustering (Hayes and Cai, 2007). While multilevel models were another option to account for the nested data, this approach could potentially produce biased estimates given the small sample size, number of clusters, and small intraclass correlations (i.e., < 0.001) in the current study (see Musca et al., 2011; McNeish and Stapleton, 2016). Regression analyses controlled for study cohort, county, and if the teacher had a co-teacher in the classroom with the rationale that co-teaching may have a positive impact on implementation (Shim et al., 2004). Finally, because the relatively high number of predictors in our regression models (i.e., 11) may have inflated the R^2 values (see Akossou and Palm, 2013), we conducted regression models for teacher characteristics (4 variables) and workplace factors (4 variables) separately to produce R^2 values specific to each of these constructs.

Results

Preliminary analyses

Bivariate correlations are presented in Table 1. Consider first the correlations evident among the teacher characteristics and workplace factors studied as predictors of implementation. More experienced teachers tended to be less receptive to the intervention and less satisfied with their jobs than less experienced teachers. Teachers in more well-resourced classrooms displayed more positive teaching practices at baseline than teachers in less well-resourced rooms. Job satisfaction, organizational learning, and workplace challenges were significantly inter-related, with job satisfaction higher in centers that supported organizational learning and faced fewer negative workplace challenges relative to positive workplace experiences.

Next, consider correlations linking these variables with the implementation outcome measures. Intervention adherence had only one significant relationship: teachers in centers that faced fewer workplace challenges showed higher adherence delivering the REDI lessons than teachers in more organizationally challenged centers. The quality of REDI curriculum delivery was associated with baseline positive teaching practices, teacher receptivity to the intervention, and classroom resources. The quality of REDI generalized teaching strategies tended to be negatively correlated with teacher experience and positively correlated with positive teaching practices, receptivity to the intervention, and classroom resources. Finally, intentions to continue REDI implementation in the future was significantly associated with organizational learning and fewer workplace challenges.

Regression analyses

Separate regression analyses that included the full set of teacher characteristics and workplace factors and control variables were conducted predicting each of the four implementation outcomes. Table 2 includes results from the regression analysis predicting program adherence. Teacher characteristics explained 1% of the variance and no teacher characteristics were uniquely associated with adherence. However, workplace challenges explained 42% of the variance and fewer REDI lessons were completed by teachers experiencing multiple workplace challenges compared with teachers who contended with fewer workplace challenges.

Table 3 includes regression results from the model predicting the quality of REDI curriculum delivery. Teacher characteristics predicted 45% and workplace factors predicted 41% of the variance in quality of REDI curriculum delivery and there were multiple unique predictors. Positive baseline teaching practices and receptivity to intervention were positively associated with the quality of REDI curriculum delivery, indicating that teachers who were more skilled and receptive to consultation delivered the REDI lessons with higher levels of quality compared to less skilled teachers and teachers who were less receptive to consultation. In addition, classroom resources were positively associated with curriculum delivery quality; teachers with more classroom resources were able to implement REDI lessons with higher quality.

Teacher characteristics predicted 40% and workplace factors predicted 18% of the variance in quality of generalized use of REDI teaching strategies in the classroom (Table 4). Positive baseline teaching practices and receptivity to the intervention were both significant, unique predictors of generalized REDI teaching strategy use. Teachers who were more skilled and more receptive to coaching used REDI teaching strategies with higher quality throughout the day compared to less skilled teachers and those who were less open to REDI coaching.

The results from the regression predicting intentions for future REDI implementation are presented in Table 5. Overall, teacher characteristics explained 13% and workplace factors explained 27% of the variance in intentions for future REDI use. The only significant unique predictor in this regression model was workplace challenges; teachers and

TABLE 1 Correlations among all variables.

	1	2	3	4	5	6	7	8	9	10	11
Teacher Experience	–										
Teacher Education	−0.02	–									
Baseline Teaching Practices	−0.00	0.23	–								
Receptivity to Intervention	−0.35*	−0.02	0.10	–							
Classroom Resources	0.03	0.19	0.50*	−0.07	–						
Job Satisfaction	−0.26†	0.01	−0.16	0.21	−0.12	–					
Organizational Learning	−0.11	−0.01	−0.13	0.29	−0.08	0.52*	–				
Workplace Challenges	0.04	−0.01	0.18	−0.14	−0.09	−0.47*	−0.39*	–			
Program Adherence	0.01	0.09	−0.07	−0.09	0.26	0.27	0.06	−0.56*	–		
Quality of Curriculum Delivery	−0.22	0.22	0.47*	0.50**	0.54***	0.15	0.20	−0.23	0.20	–	
Quality of Teaching Strategies	−0.25†	0.16	0.30†	0.61***	0.25†	0.26	0.26	−0.23	0.02	0.86***	–
Intentions for Future REDI Use	−0.02	−0.15	−0.22	0.22	0.03	0.19	0.38*	−0.51*	0.32	0.31	0.31†

$^{\dagger}p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.00$.

TABLE 2 Regression analyses predicting program adherence.

Variable	*B*	SE *B*	*β*	*t*	*p*
Teacher characteristics					
Teacher experience	−0.18	0.67	−0.07	0.27	0.79
Teacher education	−1.39	2.20	−0.05	0.63	0.53
Baseline teaching practices	−3.11	4.90	−0.08	0.64	0.53
Receptivity to intervention	−9.56	10.71	−0.11	0.89	0.37
Workplace factors					
Classroom resources	6.19	5.33	0.17	1.16	0.25
Job satisfaction	6.05	9.08	0.03	0.67	0.51
Organizational learning	−6.14	7.76	−0.06	0.79	0.43
Workplace challenges	**−57.43**	**23.05**	**−0.32**	**2.49**	**0.02**

Adjusted R^2 for only teacher characteristics = 0.01; adjusted R^2 for only workplace factors = 0.42.

TABLE 3 Regression analyses predicting quality of REDI curriculum delivery.

Variable	*B*	SE *B*	*β*	*t*	*p*
Teacher characteristics					
Teacher experience	0.00	0.02	0.00	0.07	0.95
Teacher education	0.07	0.09	0.12	0.81	0.42
Baseline teaching practices	**0.29**	**0.17**	**0.36**	**1.72**	**0.02**
Receptivity to intervention	**1.47**	**0.38**	**0.55**	**3.81**	**<0.001**
Workplace factors					
Classroom resources	**0.44**	**0.18**	**0.40**	**2.44**	**0.02**
Job satisfaction	0.03	0.27	0.03	0.12	0.91
Organizational learning	0.07	0.23	−0.01	0.31	0.76
Workplace challenges	−0.39	0.81	−0.09	0.48	0.63

Adjusted R^2 for only teacher characteristics = 0.45; adjusted R^2 for only workplace factors = 0.41.

directors in centers characterized by fewer organizational challenges felt more positively about continuing to use REDI in the future than those in centers that faced more organizational challenges.

Discussion

Evidence-based programs that enrich preschool classroom curricula and provide teachers with professional development

TABLE 4 Regression analyses predicting quality of REDI teaching strategies.

Variable	*B*	SE *B*	*β*	*t*	*p*
Teacher characteristics					
Teacher experience	0.01	0.01	0.10	0.77	0.44
Teacher education	0.01	0.05	0.02	0.29	0.77
Baseline teaching practices	**0.15**	**0.10**	**0.35**	**2.26**	**0.02**
Receptivity to intervention	**0.89**	**0.25**	**0.62**	**3.60**	**<0.001**
Workplace factors					
Classroom resources	0.12	0.10	0.20	1.20	0.24
Job satisfaction	0.17	0.14	0.21	1.19	0.24
Organizational learning	−0.04	0.13	−0.11	0.29	0.77
Workplace challenges	−0.18	0.46	−0.09	0.39	0.70

Adjusted R^2 for only teacher characteristics = 0.40; adjusted R^2 for only workplace factors = 0.18.

TABLE 5 Regression analyses predicting intentions for future REDI implementation.

Variable	*B*	SE *B*	*β*	*t*	*p*
Teacher characteristics					
Teacher experience	0.01	0.03	0.05	0.43	0.67
Teacher education	−0.13	0.11	−0.22	1.18	0.24
Baseline teaching practices	−0.04	0.30	0.02	0.12	0.91
Receptivity to intervention	0.40	0.47	0.16	0.85	0.40
Workplace context					
Classroom resources	−0.01	0.25	−0.03	0.03	0.98
Job satisfaction	−0.38	0.44	−0.25	0.86	0.39
Organizational learning	0.29	0.33	0.18	0.89	0.38
Workplace challenges	**−2.44**	**1.04**	**−0.50**	**2.35**	**0.03**

Adjusted R^2 for only teacher characteristics = 0.13; adjusted R^2 for only workplace factors = 0.27.

supports can improve classroom quality (Yoshikawa et al., 2013; Phillips D. A. et al., 2017). Such programming could be especially helpful for community-based childcare centers, which are often underserved by evidence-based supports (Larose et al., 2020) and yet are increasingly expected move the mark on readiness skills across literacy, numeracy, and social–emotional domains (Markowitz et al., 2018). In this study, we examined predictors of implementation for an evidence-based school readiness program, REDI, that targets early literacy and social–emotional skills and was adapted to support delivery in community-based childcare centers.

Results supported the first hypothesis that baseline teaching skills and teacher receptivity to the intervention would support high-quality program implementation but not adherence. Baseline teaching practices and receptivity to intervention emerged as important teacher predictors of the quality of REDI lesson delivery and REDI teaching strategy use. In contrast, none of the teacher characteristics studied were significantly associated with intervention adherence.

Study findings also supported the second hypotheses that workplace factors would support intervention adherence. Workplace factors were significantly associated with intervention adherence defined as the proportion of REDI activities completed ($R^2 = 0.42$). Interestingly, workplace factors were also significantly associated with quality of curricular delivery and intentions to implement REDI again in the future. Associations with program adherence and implementation quality are discussed further in the following sections.

Predicting adherence to the REDI program

Curriculum adherence is one key ingredient to enriching classrooms in effective school readiness programs like REDI (Jenkins and Duncan, 2017; Nguyen et al., 2018); study findings suggest that the workplace challenges that characterize some childcare centers can undermine efforts to implement evidence-based programming. When asked to describe their workplace in an open-ended interview, childcare teachers who focused on organizational challenges at their centers such as insufficient funding, enrollment instability, staffing/turnover concerns, low-quality programming, and unsupportive or ineffective administration were likely to show reduced adherence (see Hunter and Bierman, 2020 for details about these interviews). These kinds of workplace challenges are amplified in childcare centers relative to publicly-funded programs because childcare centers often do not have consistent funding mechanisms or

incentives to promote staff longevity and workforce professional development (Whitebook et al., 2018). On the other hand, teachers who reported fewer challenges and more well-resourced and effective workplaces implemented more of the REDI lessons. Teachers who faced higher levels of workplace challenges described unpredictable and stressful workdays along with less administrative and collegial support, which affected their morale and ability to prioritize delivering a new program. The fact that workplace challenges singularly predicted adherence in a model that also included individual teacher characteristics suggests that teachers were able to complete REDI lessons regardless of their background and training, a finding that has important policy implications (e.g., focusing quality standards beyond on educational attainment alone).

Predicting quality of REDI implementation

Two facets of implementation quality were assessed by the REDI consultants: the quality with which teachers delivered the REDI curriculum components and the quality with which they utilized the REDI-prescribed teaching strategies throughout the day. Quality of REDI curriculum delivery ratings ($M = 5.23$) were higher on average than quality of generalized teaching strategy use ($M = 4.22$), likely because the guided lesson plans provided a helpful scaffold to support high-quality delivery (Phillips D. A. et al., 2017). However, the two aspects of implementation quality were highly related ($r = 0.86$). Predictors of both of these dimensions of implementation quality included two teacher characteristics—baseline levels of positive teaching skills and a teacher's receptivity to the intervention consultation and coaching. Strong teaching skills can enable preschool teachers to more easily adopt new strategies, given their classrooms are already well-managed (Phillips D. A. et al., 2017), and learning new skills is not overwhelming. Receptivity to the intervention may have directly enhanced delivery quality by increasing teacher responsiveness to director input and feedback and by promoting efforts to adopt the recommended teaching strategies. Alternatively (or in addition), this finding could indicate other indirect influences on implementation quality such as a positive director–teacher working relationship (Baker et al., 2010), administrative support for the intervention (Langley et al., 2010; Wanless and Domitrovich, 2015), or stronger teacher "buy-in" (Locke et al., 2019). Interestingly, teacher experience was negatively correlated with receptivity to the intervention, suggesting that more experienced teachers were less responsive to the coaching process. Some past research suggests that more experienced teachers generally rate new programs as less acceptable than newer teachers (Witt et al., 1984; Ghaith and Yaghi, 1997), possibly because they feel settled and comfortable with their established practices.

In addition to these teacher characteristics, workplace factors, specifically classroom resources, also predicted the quality of REDI lesson delivery. Presumably, better-resourced classrooms provide teachers with access to higher quality learning materials which they can utilize to structure effective learning experiences.

Intentions for future REDI implementation

Intentions to use REDI in the future is not a direct measure of implementation quality but represents an important implementation outcome indicating perceptions of a REDI's sustainability. Preschool administrators and teachers make a significant investment when they initiate new evidence-based programming into classrooms. Prior research suggests that initial implementation experiences play an important role in determining whether preschools leverage this initial investment with sustained program use (Bierman et al., 2013). In this study, workplace factors emerged as the critical predictors of intentions for future REDI program implementation. Organizational learning and workplace challenges each predicted future intentions to use REDI, with workplace challenges contributing unique variance in the regression model. In this study, childcare center workplace challenges emerged as a barrier to implementation adherence, quality of curriculum delivery and teaching strategy use, as well as intentions for future REDI use. These findings suggest that improving childcare program quality will require addressing the day-to-day management struggles faced by under-resourced childcare centers (Whitebook et al., 2018) which represent critical barriers to the effective implementation of enriched curricula and uptake of professional development training.

Implications for evidence-based interventions in childcare centers

Early childhood stakeholders including program administrators, researchers, and policy-makers are concerned with bringing effective school readiness interventions to scale, which necessarily requires an understanding of their implementation determinants. As school-based SEL program implementation ramps up in a variety of learning contexts (Bryant et al., 2021), questions remain about how school readiness programs like REDI are being delivered in practice. To that end, this study has several implications for scaling evidence-based programs in childcare centers serving preschool children. First, workplace factors require more attention in research and more emphasis in policy-related efforts to improve childcare. In the previous study of REDI implemented in the Head Start context, workplace climate perceptions were not related to implementation quality (Domitrovich et al., 2019). In contrast to Head Start, the characteristics of childcare center workplaces may be especially variable given the range of organizational infrastructures that support them, including private owners, nonprofit cooperatives,

and faith-based organizations (Ackerman and Sansanelli, 2010; Bassok et al., 2016a). Workplace variation also exists as a function of the financial stressors and unstable teacher and student base in many childcare centers, as well as the diverse cultural expectations in different communities served (Ackerman and Sansanelli, 2010). Attention paid to the workplace environment in centers may help assess readiness for implementation (e.g., Wanless and Domitrovich, 2015), a factor that itself may need to be a target of intervention in some cases. Policy efforts aimed at expanding childcare center opportunities for preschool children (e.g., Child Care and Development Block Grants; Guarino, 2021) should also note that over-burdened and under-resourced centers may require more structural assistance (e.g., incentives to retain staff; expanded professional development; classroom resources) to assist teachers with implementing high-quality preschool programming.

In addition, this study has implications for improving and scaling professional development efforts for childcare center teachers. The professional development model evaluated in the original Head Start REDI study (Bierman et al., 2008) was adapted in this study to provide a more cost-effective and scalable model for childcare centers. Innovations included the use of online learning modules to reduce time spent in in-person workshops which are challenging for childcare centers to finance, and the use of center directors as coaches for their teachers. To help teachers deliver the REDI curriculum with high levels of quality and use the generalized teaching strategies throughout the day, directors met with them routinely for goal-setting, problem-solving, and mentoring and also observed their implementation of REDI lessons to provide feedback. Teachers' receptivity to this consultative process emerged as an important predictor of implementation, over and above other teacher characteristics such as education and experience. This result supports the expectation that engagement with and receptiveness to coaching is highly relevant for implementation outcomes, which can improve classroom quality. Indeed, a prior study of MyTeachingPartner (Pianta et al., 2008b) showed that teacher responsiveness to intervention (coach ratings of engagement) mediated the association between perceived intervention quality (teacher ratings) and changes in teacher-child interactions (LoCasale-Crouch et al., 2016). In the current study, center directors reported that although it was difficult to find time for coaching, they considered it a valuable process (Hunter and Bierman, 2020). As such, an emphasis on building infrastructure to support coach-mentors (directors or other qualified staff) in childcare centers seems warranted and would be improved by boosting teachers' readiness to engage with such coaching.

Finally, past research examining teacher professional experience with successful program implementation has yielded mixed results. In reviewing the literature, we found that teachers with advanced education implemented some preschool interventions with higher quality than teachers with less formal education (Williford et al., 2015; Sutherland et al., 2018) whereas teacher education levels showed no association with implementation quality in a number of other studies (Baker et al., 2010; Wenz-Gross and Upshur, 2012; Phillips B. M. et al., 2017). Associations between teacher education levels and implementation adherence and quality were also non-significant in the present study. At the same time, current policy efforts target regulating structural features of preschool programs such as teacher education levels in an effort to equalize early childhood care across sectors (Child Trends, 2019), although empirical relations among such structural features to child outcomes are weak (Early et al., 2007; Farran and Hofer, 2011; Perlman et al., 2017). Given that evidence-based programming and professional development can strengthen process quality (i.e., teacher-child interactions), which is more consistently related to lasting child outcomes (Phillips D. A. et al., 2017), emphasizing program curricular enhancements and PD may be a more effective and cost-efficient target for policy efforts than structural regulatory control. However, with the relatively sparse research base and mixed results, more research is needed to determine which facets of program implementation are assisted by teacher education levels or other structural characteristics of childcare centers, especially amid calls to standardize the training, educational requirements, and wages of pre-K through third grade teachers (Whitebook et al., 2018).

Limitations and conclusions

This study represents an initial evaluation of individual and workplace factors related to teachers' implementation of an evidence-based curricular enhancement intervention for preschool children in childcare centers. Additional research examining predictors of implementation quality for these kinds of program enhancements could inform the challenge of scaling up high-quality early education and reducing disparities in quality often experienced by children living in under-served communities and settings (Phillips B. M. et al., 2017; Nguyen et al., 2018). Given that access to early childhood programs has increased dramatically in recent years but quality remains inconsistent and often low (Friedman-Krauss et al., 2014; Barnett et al., 2017; Child Trends, 2019), research targeting processes that can improve the scale-up of high-quality programming across the many settings where young children are served is vital.

This study adds to limited previous work examining preschool interventions in childcare centers (Larose et al., 2020) and extends it by measuring both qualitative and quantitative aspects of the workplace in addition to multiple measures of implementation. However, several limitations are noted. First, the childcare centers in this sample were all located in Pennsylvania, and although they represented diverse geographic and socio-economic areas in the state, the generalizability of the findings to centers in states with different structures supporting and regulating early care is unknown. Second, this study followed recommendations from conceptual models to measure multiple factors at the individual (teacher) implementer level as well as the contextual organizational (workplace) level to understand implementation (Domitrovich et al., 2008). However, it is possible that other latent characteristics of either the teachers, directors, or interactions in the workplace not measured in this study could explain further variance in the outcomes assessed. Further, this study did not include director buy-in as a potential

influence on either the workplace climate or implementation (although other studies have considered it a workplace factor; Domitrovich et al., 2009). Future research validating this work, and potentially understanding prediction by individual aspects of the workplace (e.g., administrative leadership; Domitrovich et al., 2019) or evaluating the mechanisms through which individual and workplace factors impact child outcomes, is necessary.

Additional studies might also address some of the measurement limitations of this study. These included the narrow focus on the literacy materials in the classrooms to represent classroom resources and the fact that REDI consultants rated teacher receptivity to intervention and contributed to ratings of delivery quality, thereby possibly increasing the associations of those two variables. Finally, caution should be taken with the R^2 values from the regression models. While we attempted to limit the risk of overinflating the R^2 by calculating adjusted R^2 for teacher characteristics and workplace factors separately, our R^2 values were higher than what has been reported in a limited number of implementation studies (e.g., Ransford et al., 2009 reported R^2 between 0.10 and 0.16 for implementation quality). It is possible that our sample size over-inflated the R^2 (Karch, 2020) and more research is needed to better understand the amount of variance explained in implementation by teacher characteristics and workplace factors.

Although they are often overlooked in large-scale research and policy efforts, the community-based childcare professionals in this study showed positive interest and enthusiasm for the REDI program and PD model, and many were especially interested in continuing to implement the social–emotional learning aspects of REDI (Hunter and Bierman, 2020). Teachers also achieved high levels of implementation adherence and quality despite structural and workplace challenges. As such, the current study demonstrates both the viability and importance of attending to the unique characteristics of childcare centers in designing school readiness interventions and implementation supports that bolster program competence and child skills. Considerations for evidence-based policy and practice recommendations that assist childcare center professionals in making sustainable improvements like those studied here remain a priority.

Data availability statement

The datasets presented in this article are not readily available because the authors are bound by ethical obligations that could be violated upon sharing of the original dataset. They will make the data supporting the conclusions of this article available to qualified researchers upon request, with restrictions as needed to preserve the confidentiality of participating teachers, directors, and centers. Requests to access the datasets should be directed to kb2@psu.edu.

Ethics statement

The studies involving human participants were reviewed and approved by The Human Research Protection Program (HRPP) at Penn State. The participants provided their written informed consent to participate in this study.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This work was supported by the National Institute of Child Health and Human Development under Grant number HD079410; the Institute of Education Sciences under Grant number R305B090007; and a dissertation grant from the Society for the Study of School Psychology.

Acknowledgments

We appreciate the cooperation of our partners in this project: the teachers, directors, and other program personnel of the participating childcare centers.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author disclaimer

The views expressed in this article are ours and do not necessarily represent the granting agencies.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.1023505/full#supplementary-material

References

Ackerman, D. J., Barnett, W. S., Hawkinson, L. E., Brown, K., and McGonigle, E. A. (2009). *Providing Preschool Education for all Four-year-olds: Lessons From Six State Journeys*. New Brunswick, NJ.

Ackerman, D. J., and Sansanelli, R. A. (2010). The source of child care center preschool learning and program standards: implications for potential early learning challenge fund grantees. *Early Child. Res. Pract.* 12, 1–14.

Akossou, A. Y. J., and Palm, R. (2013). Impact of data structure on the estimators R-square and adjusted R-square in linear regression. *Int. J. Math. Comput.* 20, 84–93.

Baker, C. N., Kupersmidt, J. B., Voegler-Lee, M. E., Arnold, D. H., and Willoughby, M. T. (2010). Predicting teacher participation in a classroom-based, integrated preventive intervention for preschoolers. *Early Child. Res. Q.* 25, 270–283. doi: 10.1016/j.ecresq.2009.09.005

Barnett, W. S., Friedman-Krauss, A. H., Weisenfeld, G. G., Horowitz, M., Kasmin, R., and Squires, J. H. (2017). *The state of preschool 2016: State preschool yearbook*. New Brunswick, NJ.

Bassok, D., Fitzpatrick, M., Greenberg, E., and Loeb, S. (2016a). Within- and between-sector quality differences in early childhood education and care. *Child Dev.* 87, 1627–1645. doi: 10.1111/cdev.12551

Bassok, D., Latham, S., and Rorem, A. (2016b). Is kindergarten the new first grade? *AERA Open* 2, 1–31. doi: 10.1177/2332858415616358

Bierman, K. L., Domitrovich, C. E., Nix, R. L., Gest, S. D., Welsh, J. A., Greenberg, M. T., et al. (2008). Promoting academic and social-emotional school readiness: the head start REDI program. *Child Dev.* 79, 1802–1817. doi: 10.1111/j.1467-8624.2008.01227.x

Bierman, K. L., Heinrichs, B. S., Welsh, J. A., and Nix, R. L. (2021). Reducing adolescent psychopathology in socioeconomically disadvantaged children with a preschool intervention: a randomized-controlled trial. *Am. J. Psychiatr.* 178, 305–312. doi: 10.1176/appi.ajp.2020.20030343

Bierman, K. L., Sanford DeRousie, R. M., Heinrichs, B., Domitrovich, C. E., Greenberg, M. T., Gill, S. (2013). Sustaining high quality teaching and evidence-based curricula: Follow-up assessment of teachers in the REDI project. *Early Educ. Dev.* 24, 1194–1213. doi: 10.1080/10409289.2013.755457

Bryant, G., Mainelli, A., Crowley, S., Glennen, C., and Edzie, K. (2021). Finding your place 2021: social emotional learning takes center stage in K-12. Available at: https://tytonpartners.com/library/post-pandemic-schooling-emphasizing-social-emotional-learning/.

Bryk, A., Camburn, E., and Louis, K. S. (1999). Professional community in Chicago elementary schools: facilitating factors and organizational consequences. *Educ. Adm. Q.* 35, 751–781. doi: 10.1177/0013161X99355004

Burchinal, M., Nelson, L., Carlson, M., and Brooks-Gunn, J. (2008). Neighborhood characteristics and child care type and quality. *Early Educ. Dev.* 19, 702–725. doi: 10.1080/10409280802375273

Child Trends (2019). Policies, initiatives, and resources to support the ECE workforce. Available at: https://www.childtrends.org/project/fcd-ece-workforce#resources

Creswell, J. W., and Plano Clark, V. L. (2011). *Designing and Conducting Mixed Methods Research. 2nd Edn*. Thousand Oaks, CA: SAGE Publications, Inc.

Campbell, J. L., Quincy, C., Osserman, J., and Pedersen, O. K. (2013). Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociol. Methods Res.* 42, 294–320. doi: 10.1177/0049124113500475

Dennis, S. E., and O'Connor, E. (2013). Reexamining quality in early childhood education: exploring the relationship between the organizational climate and the classroom. *J. Res. Child. Educ.* 27, 74–92. doi: 10.1080/02568543.2012.739589

Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., et al. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: a conceptual framework. *Adv. School Ment. Health Promot.* 1, 6–28. doi: 10.1080/1754730X.2008.9715730

Domitrovich, C. E., Cortes, R. C., and Greenberg, M. T. (2007). Improving young children's social and emotional competence: a randomized trial of the preschool "PATHS" curriculum. *J. Prim. Prev.* 28, 67–91. doi: 10.1007/s10935-007-0081-0

Domitrovich, C. E., Gest, S. D., Gill, S., Jones, D., and DeRousie, R. S. (2009). Individual factors associated with professional development training outcomes of the head start REDI program. *Early Educ. Dev.* 20, 402–430. doi: 10.1080/10409280802680854

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

Donoghue, E. A. (2017). Quality early education and child care from birth to kindergarten. *Pediatrics* 140, 1–8. doi: 10.1542/peds.2017-1488

Dowsett, C. J., Huston, A. C., Imes, A. E., and Gennetian, L. (2008). Structural and process features in three types of child care for children from high and low income families. *Early Child. Res. Q.* 23, 69–93. doi: 10.1016/j.ecresq.2007.06.003

Early, D. M., Maxwell, K. L., Burchinal, M., Alva, S., Bender, R. H., Bryant, D., et al. (2007). Teachers' education, classroom quality, and young children's academic skills: results from seven studies of preschool programs. *Child Dev.* 78, 558–580. doi: 10.1111/j.1467-8624.2007.01014.x

Ehrlich, S. B., Pacchiano, D. M., Stein, A. G., and Luppescu, S. (2016). Essential organizational supports for early education: The development of a new survey tool to measure organizational conditions.

Farran, D. C., and Hofer, K. (2011). "Evaluating the quality of early childhood education programs," in *Handbook of Research on the Education of Young Children*. eds. O. B. Saracho and B. Spodek (New York, NY: Routledge/Taylor & Francis), 426–437.

Friedman-Krauss, A. H., Raver, C. C., Morris, P. A., and Jones, S. M. (2014). The role of classroom-level child behavior problems in predicting preschool teacher stress and classroom emotional climate. *Early Educ. Dev.* 25, 530–552. doi: 10.1080/10409289.2013.817030

Gage, N. A., MacSuga-Gage, A. S., Prykanowski, D., Coyne, M., and Scott, T. M. (2015). Investigating the collateral effects of behavior management on early literacy skills. *Educ. Treat. Child.* 38, 523–540. doi: 10.1353/etc.2015.0018

Gearing, R. E., El-Bassel, N., Ghesquiere, A., Baldwin, S., Gillies, J., and Ngeow, E. (2011). Major ingredients of fidelity: a review and scientific guide to improving quality of intervention research implementation. *Clin. Psychol. Rev.* 31, 79–88. doi: 10.1016/j.cpr.2010.09.007

Ghaith, G., and Yaghi, H. (1997). Relationships among experience, teacher efficacy, and attitudes toward the implementation of instructional innovation. *Teach. Teach. Educ.* 13, 451–458. doi: 10.1016/S0742-051X(96)00045-5

Gill, S., Greenberg, M. T., Moon, C., and Margraf, P. (2007). Home visitor competence, burnout, support, and client engagement. *J. Hum. Behav. Soc. Environ.* 15, 23–44. doi: 10.1300/J137v15n01_02

Graham, J. (2009). Missing data analysis: making it work in the real world. *Annu. Rev. Psychol.* 60, 549–576. doi: 10.1146/annurev.psych.58.110405.085530

Guarino, A. (2021). Child care and early learning in the senate version of the American rescue plan. Available at: https://www.ffyf.org/child-care-and-early-learning-in-the-american-rescue-plan-passed-by-the-senate/

Hamre, B. K., Pianta, R. C., Burchinal, M., Field, S., LoCasale-Crouch, J., Downer, J. T., et al. (2012). A course on effective teacher-child interactions: effects on teacher beliefs, knowledge, and observed practice. *Am. Educ. Res. J.* 49, 88–123. doi: 10.3102/0002831211434596

Hayes, A. F., and Cai, L. (2007). Using heteroskedasticity-consistent standard error estimators in OLS regression: an introduction and software implementation. *Behav. Res. Methods* 39, 709–722. doi: 10.3758/bf03192961

Hillemeier, M. M., Morgan, P. L., Farkas, G., and Maczuga, S. A. (2013). Quality disparities in child care for at-risk children: comparing head start and non-head start settings. *Matern. Child Health J.* 17, 180–188. doi: 10.1007/s10995-012-0961-7

Holland Coviello, R. (2005). Language and literacy environment quality in early childhood classrooms: Exploration of measurement strategies and relations with children's development.

Hunter, L. J., and Bierman, K. L. (2020). Implementing a school readiness intervention in community-based childcare centers: director and teacher perceptions. *Early Educ. Dev.* 32, 1153–1173. doi: 10.1080/10409289.2020.1833162

Jenkins, J. M., and Duncan, G. J. (2017). "Do pre-kindergarten curricula matter?" in *The Current State of Scientific Knowledge on Pre-kindergarten Effects*. eds. D. Phillips and K. Dodge, and Pre-Kindergarten Task Force, 37–44.

Johnson, A. D., Martin, A., and Schochet, O. N. (2019). How do early care and education workforce and classroom characteristics differ between subsidized centers and available center-based alternatives for low-income children? *Child Youth Serv. Rev.* 107:104567. doi: 10.1016/j.childyouth.2019.104567

Jones, S. M., and Bouffard, S. M. (2012). Social and emotional learning in schools: from programs to strategies. *Soc. Policy Rep.* 26, 1–33. doi: 10.1002/j.2379-3988.2012.tb00073.x

Karch, J. (2020). Improving on adjusted R-squared. Collabra. *Collabra Psychol.* 6. doi: 10.31234/osf.io/v8dz5

Langley, A. K., Nadeem, E., Kataoka, S. H., Stein, B. D., and Jaycox, L. H. (2010). Evidence-based mental health programs in schools: barriers and facilitators of

successful implementation. *Sch. Ment. Heal.* 2, 105–113. doi: 10.1007/s12310-010-9038-1

Larose, M.-P., Ouellet-Morin, I., Vergunst, F., Vitaro, F., Girard, A. E., Tremblay, R., et al. (2020). Examining the impact of a social skills training program on preschoolers' social behaviors: a cluster-randomized controlled trial in child care centers. *BMC Psychol.* 8:39. doi: 10.1186/s40359-020-00408-2

LoCasale-Crouch, J., DeCoster, J., Cabell, S. Q., Pianta, R. C., Hamre, B. K., Downer, J. T., et al. (2016). Unpacking intervention effects: teacher responsiveness as a mediator of perceived intervention quality and change in teaching practice. *Early Child. Res. Q.* 36, 201–209. doi: 10.1016/j.ecresq.2015.12.022

Locke, J., Lee, K., Cook, C. R., Frederick, L., Vázquez-Colón, C., Ehrhart, M. G., et al. (2019). Understanding the organizational implementation context of schools: a qualitative study of school district administrators, principals, and teachers. *Sch. Ment. Heal.* 11, 379–399. doi: 10.1007/s12310-018-9292-1

Ma, X., Shen, J., Brandi, K., Lu, X., and Baron, M. (2021). Effects of classroom practices and funding streams on early childhood learning outcomes. *Early Educ Dev.* 32, 323–342. doi: 10.1080/10409289.2020.1742039

Maier, M. F., McCornmick, M. P., Xia, S., Hsueh, J., Weiland, C., Morales, A., et al. (2022). Content-rich instruction and cognitive demand in prek: using systematic observations to predict child gains. *Early Child. Res. Q.* 60, 96–109. doi: 10.1016/j.ecresq.2021.12.010

Markowitz, A. J., Bassok, D., and Hamre, B. (2018). Leveraging developmental insights to improve early childhood education. *Child Dev. Perspect.* 12, 87–92. doi: 10.1111/cdep.12266

Marti, M., Melvin, S., Noble, K. G., and Duch, H. (2018). Intervention fidelity of getting ready for school: associations with classroom and teacher characteristics and preschooler's school readiness skills. *Early Child. Res. Q.* 44, 55–71. doi: 10.1016/j.ecresq.2018.02.010

McCormick, M. P., Cappella, E., O'Connor, E. E., and McClowry, S. G. (2015). Social-emotional learning and academic achievement: using causal methods to explore classroom-; evel mechanisms. *AERA Open* 1:233285841560395. doi: 10.1177/2332858415603959

McCormick, M. P., Mattera, S. K., Maier, M. F., Xia, S., Jacob, R., and Morris, P. A. (2022). Different settings, different patterns of impacts: effects of a pre-K math intervention in a mixed-delivery system. *Early Child. Res. Q.* 58, 136–154. doi: 10.1016/j.ecresq.2021.08.005

McNeish, D. M., and Stapleton, L. M. (2016). The effect of small sample size on two-level model estimates: a review and illustration. *Educ. Psychol. Rev.* 28, 295–314. doi: 10.1007/s10648-014-9287-x

Morrissey, T. W., Lekies, K. S., and Cochran, M. M. (2007). Implementing New York's universal pre-kindergarten program: an exploratory study of systemic impacts. *Early Educ. Dev.* 18, 573–596. doi: 10.1080/10409280701681649

Musca, S. C., Kamiejski, R., Nugier, A., Méot, A., Er-Rafiy, A., and Brauer, M. (2011). Data with hierarchical structure: impact of intraclass correlation and sample size on type-I error. *Front. Psychol.* 2:74. doi: 10.3389/fpsyg.2011.00074

National Association for the Education of Young Children (2002). Early learning standards: creating the conditions for success. Available at: https://www.naeyc.org/sites/default/files/globally-shared/downloads/PDFs/resources/position-statements/position_statement.pdf

NCES (2020). Enrollment rates of young children. Available at: https://nces.ed.gov/programs/coe/indicator/cfa.

Nguyen, T., Jenkins, J. M., and Auger Whitaker, A. (2018). Are content-specific curricula differentially effective in head start or state prekindergarten classrooms? *AERA Open* 4:233285841878428. doi: 10.1177/2332858418784283

Perlman, M., Fletcher, B., Falenchuk, O., Brunsek, A., McMullen, E., and Shah, P. S. (2017). Child-staff ratios in early childhood education and care settings and child outcomes: a systematic review and meta-analysis. *PLoS One* 12:e0170256. doi: 10.1371/journal.pone.0170256

Phillips, B. M., Ingrole, S. A., Burris, P. W., and Tabulda, G. (2017). Investigating predictors of fidelity of implementation for a preschool vocabulary and language curriculum. *Early Child Dev. Care* 187, 542–553. doi: 10.1080/03004430.2016.1251428

Phillips, D. A., Lipsey, M. W., and Dodge, K. A.Pre-Kindergarten Task Force (2017). The current state of scientific knowledge on pre-kindergarten effects.

Pianta, R., and Hamre, B. (2020). "Improving quality and impact through workforce development and implementation systems," in *Getting It Right: Using Implementation Research to Improve Outcomes in Early Care and Education* (New York, NY: Foundation for Child Development), 109–130.

Pianta, R. C., la Paro, K. M., and Hamre, B. K. (2008a). *Classroom Assessment Scoring System, Pre-k*. Baltimore, MD: Paul H Brookes Publishing Co.

Pianta, R. C., Mashburn, A. J., Downer, J. T., Hamre, B. K., and Justice, L. (2008b). Effects of web-mediated professional development resources on teacher–child interactions in pre-kindergarten classrooms. *Early Child. Res. Q.* 23, 431–451. doi: 10.1016/j.ecresq.2008.02.001

Pianta, R. C., Whittaker, J. E., Vitiello, V., Ruzek, E., Ansari, A., Hofkens, T., et al. (2020). Children's school readiness skills across the pre-K year: associations with teacher-student interactions, teacher practices, and exposure to academic content. *J. Appl. Dev. Psychol.* 66:101084. doi: 10.1016/j.appdev.2019.101084

Piasta, S. B., Justice, L. M., Cabell, S. Q., Wiggins, A. K., Turnbull, K. P., and Curenton, S. M. (2012). Impact of professional development on preschool teachers' conversational responsivity and children's linguistic productivity and complexity. *Early Child. Res. Q.* 27, 387–400. doi: 10.1016/j.ecresq.2012.01.001

Powell, D. R., Diamond, K. E., Burchinal, M. R., and Koehler, M. J. (2010). Effects of an early literacy professional development intervention on head start teachers and children. *J. Educ. Psychol.* 102, 299–312. doi: 10.1037/a0017763

Ransford, C. R., Greenberg, M. T., Domitrovich, C. E., Small, M., and Jocobson, L. (2009). The role of teachers' psychological experiences and perceptions of curriculum supports on implementation of a social and emotional learning curriculum. *Sch. Psychol. Rev.* 38, 510–532.

Shim, J., Hestenes, L., and Cassidy, D. (2004). Teacher structure and child care quality in preschool classrooms. *J. Res. Child. Educ.* 19, 143–157. doi: 10.1080/02568540409595061

Sutherland, K. S., Conroy, M. A., McLeod, B. D., Algina, J., and Kunemund, R. L. (2018). Factors associated with teacher delivery of a classroom-based tier 2 prevention program. *Prev. Sci.* 19, 186–196. doi: 10.1007/s11121-017-0832-y

Wanless, S. B., and Domitrovich, C. E. (2015). Readiness to implement school-based social-emotional learning interventions: using research on factors related to implementation to maximize quality. *Prev. Sci.* 16, 1037–1043. doi: 10.1007/s11121-015-0612-5

Wenz-Gross, M., and Upshur, C. (2012). Implementing a primary prevention social skills intervention in urban preschools: factors associated with quality and fidelity. *Early Educ. Dev.* 23, 427–450. doi: 10.1080/10409289.2011.589043

Whitebook, M., McLean, C., Austin, L. J. E., and Edwards, B. (2018). *Early Childhood Workforce Index–2018*. Berkeley, CA.

Williford, A. P., Wolcott, C. S., Whittaker, J. V., and Locasale-Crouch, J. (2015). Program and teacher characteristics predicting the implementation of banking time with preschoolers who display disruptive behaviors. *Prev. Sci.* 16, 1054–1063. doi: 10.1007/s11121-015-0544-0

Witt, J. C., Moe, G., Gutkin, T. B., and Andrews, L. (1984). The effect of saying the same thing in different ways: the problem of language and jargon in school-based consultation. *J. Sch. Psychol.* 22, 361–367. doi: 10.1016/0022-4405(84)90023-2

Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M. R., Espinosa, L. M., Gormley, W. T., et al. (2013). Investing in our future: The evidence base on preschool education.

Yurdon, M., Weiland, C., and Sachs, J. (2016). BPS K1DS: Piloting the Boston public schools' prekindergarten model in community-based organizations.

Zaslow, M., Tout, K., Whittaker, J. V., and Lavelle, B. (2010). Toward the identification of features of effective professional development for early childhood educators.

Zucker, T. A., Solari, E. J., Landry, S. H., and Swank, P. R. (2013). Effects of a brief tiered language intervention for prekindergartners at risk. *Early Educ. Dev.* 24, 366–392. doi: 10.1080/10409289.2012.664763

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 16 December 2022
DOI 10.3389/feduc.2022.1044835

OPEN ACCESS

EDITED BY
Celene Domitrovich,
Georgetown University,
United States

REVIEWED BY
Brianna Devlin,
University of Oregon,
United States
Milan Kubiatko,
J. E. Purkyne University, Czechia
Nicholas Ialongo,
Johns Hopkins University,
United States

*CORRESPONDENCE
Karen L. Thierry
kthierry@cfchildren.org

SPECIALTY SECTION
This article was submitted to Educational Psychology, a section of the journal Frontiers in Education

RECEIVED 15 September 2022
ACCEPTED 22 November 2022
PUBLISHED 16 December 2022

CITATION
Thierry KL, Page A, Currie C, Posamentier J, Liu Y, Choi J, Randall H, Rajanbabu P, Kim TE and Widen SC (2022) How are schools implementing a universal social–emotional learning program? Macro- and school-level factors associated with implementation approach. *Front. Educ.* 7:1044835. doi: 10.3389/feduc.2022.1044835

How are schools implementing a universal social–emotional learning program? Macro- and school-level factors associated with implementation approach

Karen L. Thierry*, Amanda Page, Cailin Currie, Jordan Posamentier, Yangyang Liu, Jane Choi, Haleigh Randall, Poojashree Rajanbabu, Tia E. Kim and Sherri C. Widen

Committee for Children, Seattle, WA, United States

Introduction: For universal SEL programs to contribute to positive learning environments, all school staff must be involved in implementing the program (CASEL, 2020). The first aim of the current study was to examined school/district- and macro-level factors associated with two approaches to SEL program implementation observed in schools: (1) classroom teachers as instructors of SEL lessons (i.e., teacher-facilitated) or (2) school counselors as instructors of SEL lessons (i.e., counselor-facilitated). A second aim was to examine the SEL provider's perception of the context of counselor-facilitated implementation in schools.

Methods: Public elementary and middle schools in the U.S. (N = 6,657), that adopted the Second Step digital program in the 2021-22 school year, were identified as utilizing teacher- or counselor-facilitated implementation using usage records. Predictor variables, namely support for SEL (i.e., state plans to utilize federal funding for SEL programs or access to systemic SEL consultation) and state adoption of stand-along K-12 SEL standards/competencies, were obtained from publicly available data sources. To evaluate the second aim, interviews were conducted with Second Step client support staff (N = 5), each representing hundreds of schools utilizing a counselor-facilitated implementation approach.

Results: A Generalized Linear Mixed Model analysis indicated that schools in states with support for SEL (i.e., plans to utilize federal funding for SEL programs or access to systemic SEL consultation) were more likely to use teacher-facilitated implementation than schools without support (OR = 1.64, p < .01, CI = 1.15 – 2.34). Schools in states that were early adopters of stand-alone K-12 SEL standards/competencies tended to be more likely than those without K-12 SEL standards/competencies to use teacher-facilitated implementation (OR = 1.70, p = .06, CI = 1.00 – 2.95). A qualitative study involving iInterviews with Second Step staff who support hundreds of schools utilizing counselor-facilitated implementation identified other potential factors motivating counselors as facilitators, including low SEL buy-in and

limited staff capacity. Although this approach has challenges, it could be a pathway to teacher-facilitated implementation over time.

Discussion: Taken together, findings indicate promising strategies for the promotion of more schoolwide use of SEL programs.

KEYWORDS

social–emotional learning, program implementation, educational policy, learning standards, educators

Introduction

Students learn best in the context of safe and supportive relationships with their peers, teachers, and other school staff (Klem and Connell, 2004; MacNeil et al., 2009). These interactions promote social and emotional skill development throughout childhood and adolescence which is critical for students' academic achievement and life success (Moffitt et al., 2011; Jones et al., 2015). A positive learning environment enables a systemic approach to social–emotional learning (SEL), in which all students are actively involved in practicing social and emotional skills (Mahoney et al., 2021). In a consensus report authored by a nonpartisan and multi-disciplinary team of educators, researchers, policymakers, business, and military leaders, the National Commission on Social, Emotional, and Academic Development (2018) affirmed the importance of the social–emotional environment for learning in schools. To foster these skills in students, the Commission laid out several recommendations for districts and schools to consider when integrating SEL into schoolwide practices. One of these recommendations is the adoption of an evidence-based, universal SEL program for the explicit instruction of social–emotional skills with regular opportunities to integrate these skills with academic content and throughout school-wide experiences. The current study examined whether state policies and support mechanisms designed to promote SEL in schools are associated with a more schoolwide approach to the implementation of a universal SEL program.

For SEL programs to contribute to positive learning environments, all school staff must be involved in implementing the program, especially the teachers who have the most frequent and direct interaction with students (CASEL, 2020b). In efficacy studies of universal SEL programs, teachers are typically the implementers of the curriculum, whereby they receive training to support high-quality delivery of the lessons and integration of the skills throughout the school day (e.g., during academic instruction, in the hallways, cafeteria, and playground). In a meta-analysis of the effects of universal SEL programs in schools, Durlak et al. (2011) found that over half of the interventions were administered by classroom teachers. The other half of the interventions were delivered by non-school personnel (e.g., program developer staff, research staff). They also found that teacher-facilitated SEL programs had the broadest effects on students, resulting in greater improvements in every outcome measured examined.

Outside of efficacy studies, when SEL programs are implemented at scale in schools, teachers may not be the program implementers. Rather, anecdotal evidence indicates that school counselors are often tasked with delivering universal SEL programs to students. In this instance, a school counselor may push into classrooms across grade levels to teach lessons to students on a regular basis. They may often, but not always, do so without any active involvement from teachers in the SEL content. Other models could involve counselors teaching SEL in most classrooms, with a few teachers self-selecting to implement the program in their classroom. This latter model could represent cases where counselors are attempting to transition the program to being facilitated by teachers schoolwide. The choice of school counselors as implementers of SEL programs aligns with their role in supporting students' social and emotional development. However, a reliance on school counselors as facilitators of SEL programs could present barriers to schoolwide SEL. For instance, counselors typically lack the authority to require that teachers also engage in the SEL program, which makes teachers less likely to integrate SEL skills into their academic instruction with students.

If teachers are the primary implementers of SEL curricula, then they should be more likely to acquire and use SEL language and skills themselves, and in turn, reinforce these skills with students beyond the lesson instruction period. When teachers model SEL skills and build positive, respectful, and empathetic relationships with students, greater improvements in students' own SEL skills and academic performance have been found (Mashburn et al., 2008; Burchinal et al., 2010; Downer et al., 2012). Counselors can still play a supportive role in teacher-facilitated implementation (e.g., providing teachers with foundational training on social and emotional skill development, reinforcing student learning *via* use of shared strategies, etc.), but the key difference is that counselors are not tasked with delivering the program to all students. Given evidence showing the critical role of the classroom teacher in supporting students' SEL competencies, research is needed to understand factors that contribute to decisions to utilize counselors versus teachers as facilitators of SEL programs, beyond general differences in professional training.

Potential factors associated with social–emotional learning implementation approach

Teacher- vs. counselor-facilitated SEL implementation approaches are likely impacted by various socio-ecological factors, or the context in which programs are located (Atkins et al., 1998). Based on a socio-ecological conceptualization, Domitrovich et al. (2008) proposed a multi-level quality of implementation framework for SEL programs, in which the following three levels of systems can affect program implementation in schools: (1) individual level, involving characteristics of those delivering the program (e.g., professional and psychological characteristics of staff), (2) school/district level, such as having enough staff capacity to support high-quality program implementation, and (3) macro level, which involves state and federal funding and policies to support a program.

As previously discussed, an individual-level characteristic of school counselors that distinguishes them from teachers is that counselors' roles are more explicitly tied to children's social and emotional well-being. However, given that differences in counselors' and teachers' professional training is a stable characteristic, we would not expect this factor *per se* to account for any variation in leaders' decision-making around teacher- or counselor-facilitated implementation. Other individual-level factors that could impact implementation approach is the extent of teacher buy-in for SEL and self-efficacy for teaching SEL. Motivation to implement a new program is a well-documented factor associated with uptake of programs across a range of settings (Atkins et al., 2008; Aarons and Sommerfeld, 2012). Studies also show that these individual-level factors are heavily influenced by school/district- and macro-level factors (Scaccia et al., 2015), specifically the availability of district- and state-level resources and policies that support teachers vs. counselors as SEL program facilitators along with structural characteristics of schools that make it more or less feasible for teachers vs. counselors to serve as SEL program facilitators.

Support for social–emotional learning programming

A school- and district-level factor likely to be associated with implementation approach is the extent to which schools have direct strategic programming support for systemic SEL. During the last decade, the Collaborative for Academic, Social, and Emotional Learning (CASEL) has supported the Collaborating Districts Initiative in systemic SEL improvement. To date, this effort, includes 20 large school districts across 15 states in the country. One recent evaluation indicated that districts and schools participating in the Collaborating District Initiative had more indicators of systemic SEL (e.g., use of a universal SEL curriculum, supportive school and classroom climate, supportive schoolwide discipline practices) than districts and schools not participating in the Collaborating Districts Initiative (Schwartz et al., 2022). Although the use of counselors versus teachers as implementers was not assessed in this study, the indicators of systemic SEL would suggest that these schools were also more likely to be utilizing teacher-facilitated implementation of SEL programs.

At the macro-level, the use of federal or state funds for schoolwide SEL program implementation may also be a correlate of teacher- versus counselor-facilitated implementation approaches. Of course, educational funding availability or allotment can vary widely across states. In response to the COVID-19 pandemic, the federal government provided a $189 billion distribution of COVID-relief funding for school districts, known as the Elementary and Secondary Schools Emergency Relief (ESSER) fund. Based on initial reviews of the spending plans of approximately 100 large, urban school districts, about 43% reported plans to invest in social–emotional support for students (Dusseault and Pillow, 2021). Similar findings were obtained from a survey of hundreds of superintendents about their plans to use COVID relief funds (AASA, 2022). Given that teacher-facilitated implementation is more costly than counselor-facilitated (e.g., more staff to train, higher curricula costs, etc.), schools with funding for SEL programming may be more likely to utilize teacher-facilitated implementation. The AASA report also indicated regional differences in SEL spending plans, with superintendents from rural districts being less likely than those from suburban and urban districts to indicate plans to invest in SEL programs (AASA, 2022). These regional differences in use of funding for SEL instruction could be tied to SEL implementation approach in rural, urban, and suburban schools. No extant studies have examined how funding utilization is tied to SEL implementation approach.

Social–emotional learning policy

Also at the macro level, schoolwide SEL is often supported by national and state policies that prioritize SEL alongside core subject areas (reading, math, science). One policy that may be correlated with implementation approach is whether states have adopted stand-alone SEL standards or competencies across grade levels. To facilitate state education agency support for SEL implementation in school districts, CASEL also initiated the Collaborating States Initiative in 2016 (involving 30 states). A core outcome of this effort has been an increase in the number of states with K-12 SEL standards/competencies, from four states in 2016 to 2027 states in 2022 (Dusenbury et al., 2020), the majority of which are located in the Central, Northeastern, and West Coast regions of the country (Dermody and Dusenbury, 2022). Fewer states in Gulf Coast, Southeast, and Mountain Plains have adopted K-12 SEL standards or competencies. The existence of stand-alone SEL standards/competencies across grade levels, especially for those who adopted SEL standards early on, may signal to educators the importance of teaching these skills alongside academic skills (e.g., reading, math). In this context, district and school leaders may be more likely to view teachers (rather than counselors) as key implementers of universal SEL curricula (CASEL, 2020a,b). No extant studies have examined whether this specific policy is in fact associated with SEL implementation approach.

Structural characteristics of schools

Certain structural characteristics at the school level could make teacher- or counselor-facilitated implementation approaches more or less feasible. For instance, the choice of counselors as lesson facilitators may be a more feasible option in elementary grades than in middle school grades due to the larger enrollment and departmentalization of classes in middle school grades, which would make it more difficult for school counselors in middle schools to teach SEL lessons. Relatedly, because schools in rural areas and towns tend to have lower student enrollment than schools in urban and suburban areas, counselor-led approaches may be more feasible in rural areas and small towns.

Study aims and hypotheses

Aim one

The first aim of this study was to describe the frequency of teacher- versus counselor-facilitated implementation of a widely used SEL program (i.e., Second Step) in K-8 schools in the United States and to identify school/district- and macro-level predictors of implementation approach. The following variables were examined as primary predictors of implementation approach in schools: (1) Access to support for SEL programs (i.e., district participation in CASEL's Collaborating Districts Initiative and/or utilization of ESSER funding for SEL programs) and (2) State adoption of stand-alone SEL standards/competencies. Covariates included school-level structural/demographic variables (i.e., grade levels served, number of students enrolled, school locale or urbanicity, student race/ethnicity, and student participation in free/reduced-price lunch program). In addition, given that some states mandate that schools have counselors, we also included this variable as a covariate, as it may increase the likelihood of schools using a counselor-led approach. Finally, we included state government party control (Democratic, Republican, or divided between the two) as a covariate, given the influence of this factor on policies and funding.

We hypothesized that controlling for the covariates, schools with access to support for SEL programming would be more likely to use a teacher-facilitated implementation approach (and less likely to use counselor-facilitated) than schools in districts without these types of SEL support. We also predicted that schools in in states with stand-alone K-12 SEL standards/competencies would be more likely to use teacher-facilitated implementation, especially those that were early adopters of SEL standards/competencies.

Aim two

A second aim was to examine the SEL provider's perception of the context of counselor-facilitated implementation in schools. Given their direct interaction with both school/district leaders and counselors across a diverse range of education settings (e.g., urban, suburban, rural areas), Second Step staff could provide a unique view of multi-level factors (individual, school, district, macro) associated with the decision to have counselors facilitate the SEL program. Most existing implementation studies of SEL programs focus on schools utilizing teacher-facilitated implementation and what teachers need to be successful in their SEL instruction. Hence, there is a dearth of implementation studies that have explored what, if any, challenges counselors encounter when facilitating universal SEL programs. From a systemic SEL implementation perspective, it might be difficult for counselors to ensure that SEL is integrated into classroom and schoolwide practices given that organizationally, counselors have no authority over teachers' and other staff practices. Last, we sought to identify any promising practices associated with counselor-facilitated implementation, particularly related to whether this approach could be a pathway to teacher-facilitated implementation over time and if so, what types of support could help to facilitate that transition.

Materials and methods

Participants

The aims were examined in the context of a research-based SEL program, known as Second Step, which is used in schools across all 50 states. The digital curriculum includes lessons from kindergarten through eighth grade that teach social–emotional skills aligned with the CASEL framework of SEL competencies (i.e., self-awareness, self-management, social awareness, relationship skills, and responsible decision-making).

Aim one: Factors associated with implementation approach

Inclusionary/exclusionary criteria

A total of 7,918 sites (including public, private, parochial schools and non-school organizations) were identified as using the Second Step K-8 digital program during the 2021–22 school year. In the current study, the sample was restricted to only public-school users of the program, which made up the majority of schools using the program (87.1%). Schools outside the United States were also excluded from the study sample (0.9% of all schools). Teachers and counselors were the most frequent users of the program (based on job title entries), making up 92.7% of users. Users with other job titles (e.g., Principal, Assistant Principal, Support Staff, and Specialist), which made up less than 1% of all users, were excluded from the study sample. When these exclusionary criteria were applied, a total of 6,866 schools remained.

Study sample characteristics

When using the Second Step digital program, users are required to create "classes" representing unique groups of students for whom the lessons are being delivered. At the class level, a total of 109,629 classes from kindergarten

through grade 8 were registered by counselors or teachers in the Second Step digital program system during the 2021–2022 school year. These classes were located in 6,866 public elementary and middle schools and 1,820 districts. About 80% ($n = 87,223$) of the classes were facilitated by teachers and 20% ($n = 22,406$) were facilitated by counselors. Table 1 shows site-level demographic information, including the race/ethnicity of students, the percentage students qualifying for free and reduced-price lunch, average student enrollment, and representation of schools among regions of the U. S. User-level demographic information is not captured by the program.

Aim two: Interviews with Second Step support staff

Five Second Step client support staff who help district and school leaders in their implementation of the Second Step program participated in interviews with research staff, in which they reflected on their experience working with schools that utilized predominantly counselor-led implementation. These staff supported schools in primarily the Gulf Coast and Mountain Plains regions of the United States (years of experience working in this capacity ranged from 5 to 10 years). All staff provided informed consent for their participation.

TABLE 1 Descriptive statistics school- and state-level covariates

Variable	*M*	*SD*	Range
Race/Ethnicity			
African American	13.2%	20.1	0–100
American Indian	1.2%	7.4	0–100
Asian	6.2%	11.7	0–98
Latinx	32.3%	29.8	0–100
White	46.8%	32.8	0–100
Free/Reduced lunch	55.7%	29.1	0–100
Number students enrolled	575	285	13–4,000
	N	**%**	
School grade span			
Elementary (K-5)	4,040	58.8	
Middle School (6–8)	2,826	41.2	
School locale			
Rural	936	13.6	
Suburban	2,298	33.5	
Town	647	9.4	
Urban	2,588	37.7	
State counselor mandate			
Yes	1,594	23.3	
No	5,257	76.7	
Political Party Trifecta			
Democratic	3,234	47.1	
Republican	1,939	28.2	
Divided	1,674	24.5	

Procedure

To address aim one, at the end of the 2021–22 school year, user- and class-level data were extracted from the Learning Management System (LMS). Data for the predictor variables and covariates were obtained from publicly available datasets or reports (see description below). Schools were defined as having teacher-facilitated implementation of Second Step if 50% or more of registered classes were facilitated by teachers and were otherwise defined as having counselor-facilitated implementation. To address aim two, during summer 2022, individual Second Step client support staff members participated in virtual semi-structured interviews with research staff.

Aim one measures

Predictor variables

Social–emotional learning support factors

District participation in CASEL's Collaborating Districts Initiative was obtained from the most recent report of district involvement in the initiative (CASEL, 2021). A binary indicator of participation was used, with 0 indicating that a school was not in a participating district and 1 indicating that it was in a participating district. In the current sample, a small percentage of schools were identified as participants of the Collaborating Districts Initiative (i.e., 239 schools or 3.5% of all schools in sample).

State ESSER spending plans for the 2021–22 school year were captured from a report created by an independent policy organization (FutureEd, 2022). FutureEd reviewed state education agency plans submitted to the United States Department of Education which delineated the types of programs and resources that states planned to prioritize to support students' learning recovery during the pandemic. Most of the allocated funding (90%) went directly to schools and districts. Based on these data, research staff identified those states that had plans to use funds for SEL programming. About 32% ($n = 2,197$) of schools in the sample were in states where ESSER funds were flagged for SEL programming. A binary indicator was also created for this support type (0 = no plan to use funds for SEL programming, 1 = plans to use funds for SEL programming).

For each school, a combined SEL support measure was obtained by summing scores across the two types of SEL support (range = 0 to 2), with 0 indicating no support, 1 indicating either type of support, and 2 indicating both types of support.

Macro-level policy

States with freestanding K-12 SEL standards/competencies were identified from the 2022 CASEL State Scan (Dermody and Dusenbury, 2022). States were assigned as either not having SEL standards/competencies or as having SEL

standards/competencies. Those with SEL standards/competencies were further classified according to adoption timeframe as follows: (1) early (adopted 2015 or earlier), (2) mid (adopted between 2016 and 2018), or late (adopted between 2019 and 2021).

Covariates

Type of Second Step program utilized (elementary or middle school) was obtained from Second Step LMS records. School-level student demographics (i.e., student enrollment, percentage of students qualifying for free or reduced-price lunch, and percentage of students by race/ethnicity) along with school locale (rural, town, suburban, and urban) were obtained via a data lease from Market Data Retrieval (i.e., MDR Education), which provides validated demographic information aggregated at the building level, capturing 100% of elementary and secondary schools in the United States. The building-level data in the Second Step LMS were matched with the MDR database. The source of the MDR school locale data is the National Center for Education Statistics (NCES) classification system which is based on a school's physical address. Source of demographic information (student race/ethnicity, participation in free/reduced-price lunch) is also the NCES. For the student race/ethnicity covariate, a measure for the percentage of students who were Black, Indigenous, and People of Color (BIPOC) was created by summing the individual race/ethnicity percentages representing BIPOC students (i.e., Asian, African American, American Indian/Alaskan Native, Latinx). School enrollment data was obtained from state enrollment reports.

States with mandates for counselors in K-8 schools were identified from a report produced by the American School Counselors' Association (American School Counselors' Association, 2022). State government political trifecta information was obtained from Ballotpedia (2022), which categorized states as Democratic or Republican if one of the parties held the governorship, a majority in the state senate, and a majority in the state house. States were categorized as being divided if neither party had trifecta control.

Criterion variable

An examination of the distribution of teacher- and counselor-registered classes in schools implementing Second Step indicated a bimodal distribution. Although some schools (i.e., 6.1% of the sample) had an approximately equal combination of both counselors and teachers facilitating lessons, classes in most schools (71.2%) were facilitated entirely by either counselors or teachers. As a result, we created a binary criterion or outcome variable for each school. Schools were defined as having teacher-led implementation of Second Step if 50% or more of registered classes were facilitated by teachers (assigned score of 1) and were otherwise defined as having counselor-facilitated implementation (assigned score of 0).

Aim two: Interviews with Second Step support staff

Virtual interviews (*via* Zoom) were conducted with selected Second Step client support staff who each represented hundreds of schools utilizing a counselor-facilitated implementation approach. Each interview was conducted by two research staff (one served as the primary interviewer; the other took notes). Individual interviews lasted approximately 30–45 min, and with the permission of the interviewee, all sessions were recorded. The primary interviewer asked a series of questions about their experience supporting schools utilizing counselor-facilitated implementation of Second Step. Although all interviewees received the same initial prescribed questions (see Appendix), there was more flexibility in the framing of follow-up questions, given the variability of interviewee's responses to the questions.

Recordings of interview sessions were transcribed. For each session, a thematic analysis of responses was conducted using the grounded theory method of coding (Corbin and Strauss, 2014). A single coder, trained in qualitative analysis methods, identified themes directly from the participant responses to each question. A second coder independently coded 20% of the interviews. Any discrepancies or disagreement between the coders were resolved through discussion. Inter-rater reliability was high, ranging from 90 to 100%. Using the Domitrovich et al. (2008) model, the primary coder and a third coder with experience utilizing the model in research studies, independently categorized themes for each question according to level of the system represented (i.e., macro-, district-, school-, or individual). Any discrepancies in coding were resolved through discussion.

Data analytic plan

Because schools were nested within states, potential dependencies in the outcome measure within states was examined using the intra-class correlation (ICC). Although schools could also be nested within districts, on average there were less than five schools per district, with a substantial portion having only one school per district. As a result, this level was not included in the analysis (Brauer and Curtain, 2018). The ICC for the effect of clustering of schools within states was 0.13 indicating intra-state dependencies in the data. As a result, a Generalized Linear Mixed Model (GLMM) was used to account for these dependencies, which was run using the GENLINMIXED procedure in version 27 of SPSS. Given the inclusion of a binary outcome measure (i.e., school-level implementation approach with teacher-facilitated coded as 1 and counselor-facilitated coded as 0), a logit link function was used. State was included as a random effect in the model. The following fixed effects variables were examined as primary predictors of implementation approach: (1) Access to support for SEL programs (i.e., district participation in CASEL's

Collaborating Districts Initiative and/or utilization of ESSER funding for SEL programs) and (2) State adoption of stand-alone SEL standards/competencies (early, mid, late adopters). The following fixed effects school-level structural/demographic variables were included as covariates: (1) School grade levels served (elementary vs. middle school), (2) Number of students enrolled in school, (3) School locale/urbanicity (rural, urban, suburban, town), (4) School percentage of students who are BIPOC, (5) School percentage of students in free/reduced-price lunch program, (6) State-level school counselor mandates, and (7) State-level political party trifecta (democratic, republican, divided).

Descriptive statistics for each of the predictors/covariates and the criterion measure is shown in Table 2. Bivariate correlations were used to examine potential collinearity among continuous covariate/predictor variables, which was confirmed using collinearity diagnostics (i.e., large variance inflation factor (VIF) coefficients). Chi-square tests and Cramer's V were used to examine potential dependencies among categorical predictor variables.

Results

A total of 209 schools were missing demographic data, reducing the total number of schools included in the analysis to 6,657. Schools with missing data did not differ from the rest of sample on predictor or outcome variables. The correlation between the percentage of students qualifying for free or reduced-price lunch and the percentage of students who identified as BIPOC was in the higher range ($r = 0.66$). However, follow-up collinearity diagnostics were acceptable (i.e., tolerance >0.1 and VIF < 5), and both variables were retained in the GLMM analysis. Dependence between categorical predictors and covariates were examined using chi-square tests and Cramer's V as the measure of the strength of association between variables. All relationships were in the small to medium range (i.e., Cramer's V < 0.30).

TABLE 2 Descriptive statistics for predictor variables and criterion variable.

Variable	*N*	%
Support for SEL		
0 (neither type of support indicated)	4,513	66.6
1 (ESSER funding or CDI participant)	2,091	30.9
2 (Both funding and CDI participant)	172	2.5
State K-12 SEL standards		
Early adopter	959	14.0
Mid adopter	1,247	18.2
Later adopter	2,335	34.0
No	2,310	33.6
Implementation approach		
Counselor-led	1,649	24.0
Teacher-led	5,217	76.0

Aim one: Factors associated with implementation approach

Based on the cut score criteria, the majority of schools (76%) were identified as having teacher-led implementation of Second Step. About a quarter of the schools (24%) were identified as having counselor-led implementation. Higher percentages of the counselor-led approach were in schools in the Southeast (39.1%), Gulf Coast (36.5%), and Mountain Plains (32.1%), compared to the West Coast (12.9%), Central (19.8%), and Northeast regions (25.8%).

The results of the logistic GLMM are summarized in Table 3. The model correctly classified 73.5% of the cases. Results of the fixed effects estimates indicated that having support for SEL (i.e., plans to use ESSER funds for SEL programming or from a district participating in CASEL's Collaborating Districts Initiative) increased the likelihood of schools being teacher-facilitated (OR = 1.64, $p < 0.01$). In addition, schools in states that were early adopters of stand-alone K-12 SEL standards/competencies tended to be more likely to use teacher-facilitated implementation of Second Step than schools in states without K-12 SEL standards/competencies, although this association was marginally significant (OR = 1.70, $p = 0.06$). No significant increase in the likelihood of teacher-facilitated implementation was found for mid-adopters (OR = 1.11, $p = 0.73$) and late adopters (OR = 0.94, $p = 0.81$).

Schools in states with a Democratic trifecta (OR = 1.70, $p < 0.05$) were more likely to use teacher-facilitated implementation than those with a Republican trifecta. When examining school demographic and structural factors, schools in rural areas (OR = 0.63, $p < 0.01$) and schools in towns (OR = 0.73, $p < 0.05$) were less likely than schools in urban areas to have teachers facilitate Second Step programming. Elementary schools were less likely to use teacher-facilitated implementation than middle schools (OR = 0.59, $p < 0.01$). Schools in states with a mandate that schools have counselors were less likely to be teacher-facilitated compared to schools in states without a counselor mandate (OR = 0.56, $p < 0.001$). Student demographic characteristics had no significant association with implementation approach, but a significant association, in the direction of increased likelihood of being teacher-facilitated, was found for number of students enrolled (OR = 1.001, $p < 0.01$).

Aim two: Interviews with Second Step support staff

Themes identified from the Second Step support staff interviewers are described according to the macro-, district-,

TABLE 3 Results of generalized linear mixed model analysis: fixed effects.

Parameter	*F*	Coefficient (SE)	OR	Value of *p*	CI
SEL support (ref = none)					
Either type	7.393**	0.495 (0.182)	1.640	0.007	1.148–2.342
Both types	2.858	−0.966 (0.571)	0.381	0.091	0.124–1.167
SEL K-12 standards (ref = none)					
Early adopters	3.508[a]	0.529 (0.282)	1.697	0.061	0.976–2.952
Mid adopters	0.115	0.104 (0.306)	1.109	0.734	0.609–2.020
Late adopters	0.056	−0.060 (0.255)	0.942	0.813	0.571–1.553
Locale/Urbanicity (ref = urban)					
Rural	9.542**	−0.467 (0.151)	0.627	0.002	0.466–0.843
Suburban	2.126	−0.157 (0.108)	0.855	0.145	0.693–1.055
Town	4.382*	−0.320 (0.153)	0.726	0.036	0.538–0.980
Program					
(0 = Elementary, 1 = Middle)	12.786***	−0.531 (0.148)	0.588	0.000	0.440–0.787
State party trifecta (ref = Republican)					
Democratic	4.380*	0.535 (0.256)	1.707	0.036	1.034–2.817
Divided	2.535	0.445 (0.279)	1.560	0.111	0.902–2.697
State counselor mandate					
(0 = none, 1 = state mandate)	7.833**	−0.588 (0.210)	0.556	0.005	0.368–0.839
Student enrollment	7.507**	0.001 (0.000)	1.001	0.006	1.000–1.001
Percent free/reduced lunch	0.256	0.001 (0.002)	1.001	0.613	0.997–1.006
Percent BIPOC	1.599	0.004 (0.003)	1.004	0.206	0.998–1.009

[a]$p = 0.06$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.
SE, standard error; OR, odds ratio; CI, confidence Interval.

school-, and individual-level factors that staff perceived to be associated with decisions to assign counselors as facilitators of the Second Step program. All staff interviewed reported challenges with counselor-facilitated implementation. Themes for challenges were organized according to the same system-level factors that gave rise to decisions to assign counselors as facilitators. See Table 4 for a summary of each system-level factor related to implementation approach decision-making as well as implementation challenges encountered by counselors.

Macro-level factors

Macro-level factors that appeared to steer schools in the direction of counselor-facilitated implementation included policies related to SEL and interpretation of national guidelines related to school counselors' roles and responsibilities. Four of the five participants stated that in states where new stand-alone K-12 SEL standards or competencies were adopted, districts subsequently purchased the Second Step program given the need to comply with new SEL standards teaching requirements. However, in this context, staff indicated that district and school leaders likely perceived counselor-led facilitation of SEL programs as the most efficient way to immediately comply with the new SEL standards.

In addition, two participants suggested that national guidelines for the professional standards and competencies of school counselors may be interpreted by district and school leaders as meaning that SEL programming should be the primary responsibility of counselors. According to the American School Counselor Association (ASCA) professional standards and competencies, counselors are responsible for identifying evidence-based curricula to support student "mindsets and behaviors" along with plans for ensuring effective implementation of instruction (American School Counselors' Association, 2022).

District-level factor: Budgeting for social–emotional learning within school systems

Participants also noted budgeting as another factor associated with counselor-facilitated implementation. In some of the schools supported by Second Step staff, they observed that SEL programming is budgeted under the counseling department. Two participants mentioned that once a precedent is set in a district or school that SEL resides within counseling vs. general education, it becomes difficult to engage stakeholders in exploring a more collaborative approach to implementing SEL programming. As one participant explained, "SEL is budgeted to counseling teams so that becomes the lane it lives in."

School-individual-level factor: Low social–emotional learning buy-in

Impact on decision-making

All participants mentioned low SEL buy-in among school leaders and teachers as a reason that schools choose

TABLE 4 Summary of Second Step staff interviews: Factors related to SEL program decision-making and implementation challenges experienced by counselors.

System level	Factor	Impact on counselor-facilitated decision-making	Implementation challenges experienced by counselors
Macro	SEL-related policies	• Need to comply with new SEL state standards • Interpreting the American School Counselor Association (ASCA) model as guidance that SEL should fall only within the realm of counselor responsibilities	• Less support for schoolwide implementation from district and school staff
District	SEL budgeting	• Placing budget for SEL within counseling departments	
School/Individual	Low SEL buy-in among teachers/ leader	• Lack of understanding among leaders/teachers of what SEL is and how it benefits students drives perceptions of SEL as more in the realm of counselors' roles and responsibilities.	• Little influence on building capacity for systemic SEL (i.e., lack of shared language/understanding of SEL) and thus few opportunities for each adult in school building to consistently reinforce SEL skills in students outside of lesson instruction time
	Limited staff capacity	• Teacher turnover/burnout: Rather than adding new initiatives for overburdened teachers, school leaders tend to assign SEL program facilitation to counselors.	• Counselor burnout due to lack of support from leaders/teachers and competing priorities (i.e., managing SEL implementation along with counselors' other job responsibilities) • Class-wide instruction not aligned with counselors' expertise (e.g., lack of training in pedagogy, classroom management)

counselor-facilitated implementation over teacher-facilitated implementation. Three of the five participants elaborated that a lack of understanding of what SEL is and how it benefits students makes educators more likely to assign SEL to the domain of counselors (or other support staff, such as social workers and school psychologists).

Implementation challenges

In the context of low buy-in among teachers and administrators, four of the five participants indicated that counselors feel isolated because they receive little implementation support from school leaders and teachers. Three participants mentioned that teachers often leave the classroom during the counselor-facilitated lesson instruction time, even though some counselors indicate that it would be helpful to have teachers' instructional support during this time. Although counselors might have more formal training in children's social and emotional development, one participant indicated that serving as an instructor of an SEL curricula is in many ways not aligned with counselors' expertise. Specifically, counseling staff often lack professional training and experience related to effective pedagogical practices and classroom management.

Additionally, three participants mentioned that when principals are not bought in, it leads to issues with schoolwide reinforcement of the program. As one participant explained, "If you do not have principals and classroom teachers guiding and supporting as well, there is no one to think of the big picture outcomes – no one directing the reinforcement."

School-individual-level factor: Limited staff capacity

Impact on decision making

Three of the five participants stated their clients rely on counselors to own and implement the Second Step program because teachers do not have the time or capacity to take on SEL programming given their existing workloads. In the context of the pandemic, in which many schools experienced high rates of teacher turn-over, overwhelm, and burnout, school leaders tended to refrain from asking teachers to take on any additional duties or initiatives.

Implementation challenges

Unfortunately, with limited support from leaders and teachers, counselors implementing SEL programs end up experiencing high levels of burnout themselves. Three participants highlighted that being solely responsible for implementing a schoolwide, universal SEL program can hinder a counselor's ability to fulfill their other important responsibilities (e.g., diagnostic testing, working directly with individual students). As a result of these competing priorities, counselors may be unable to implement the SEL program in a high-quality manner (e.g., limited engagement in program planning and adherence to program requirements). Alternatively, they might deprioritize other important duties to accommodate the SEL program. As a result, while trying to protect teachers from taking on the additional work of implementing an SEL program, schools and districts can thus inadvertently saddle counselors with more than they can handle.

Pathway toward schoolwide social–emotional learning

Because of these various challenges associated with counselor-facilitated implementation, students may be less likely to benefit from SEL instruction. However, Second Step staff also described some promising practices observed in counselor-facilitated implementation that could promote a pathway toward more schoolwide SEL. All five participants indicated observing two main forms of counselor-facilitated implementation in schools: (1) one in

which only counselors served as implementers of the program and (2) a second "transitional" form in which implementation begins with only counselors as facilitators but over time, teachers gradually take ownership of lesson facilitation as they begin to feel more comfortable with the SEL program content.

To support counselors in transitioning implementation to teachers, all five participants endorsed the need to adapt program resources or create case-specific resources for districts and schools. Two participants indicated that increasing teacher and leader buy-in for SEL programming was a critical step toward the transition to teacher-facilitated implementation. To increase principal buy-in, one participant recommended emphasizing that teacher-facilitated implementation can support broader schoolwide improvements (e.g., establishing consistent SEL language, addressing behavioral challenges). Strategies for gaining teacher buy-in included identifying a few teachers in the building who show greater interest in the SEL program and inviting them to observe counselors teach SEL lessons or co-teach lessons with counselors. In this type of diffusion of innovation model, these teachers would serve as early adopters of the program who can help counselors champion the program among other teachers.

In addition to creating opportunities for teacher buy-in via an early adopter model, participants also indicated that SEL program developers should pre-adapt implementation and lesson planning resources and provide them directly within the program, which would reduce the burden on counselors who often need to adapt the resources themselves. Counselors could then have more time to follow up with teachers on the SEL skills covered in class, which would make teachers more likely to reinforce the skills. For example, one participant recommended creating multi-grade bundles of planning materials in addition to a flexible pacing guide so that counselors tasked with facilitating lessons across multiple grades would need to spend less time compiling resources needed for implementation and more time engaging teachers and leaders to gain buy-in.

Finally, providing easy-to-use and easy-to-access materials that counselors can directly share with school leaders and classroom teachers was another suggestion made by interview participants (e.g., brief unit overview videos, tools to support use of SEL vocabulary in the classroom). These resources would support reinforcement of SEL knowledge and skills learned during SEL lessons and encourage teachers to learn more about the language and strategies their students are being taught.

Discussion

A goal of SEL programs is to help support generalization of social–emotional skills so that SEL becomes integrated into the fabric of the school community. Previous studies and best practice recommendations suggest that teachers are integral to supporting generalization of SEL skills and improvement of students' social–emotional competencies (Mashburn et al., 2008; Burchinal et al., 2010; Durlak et al., 2011; Downer et al., 2012; CASEL, 2020a). In the current study, teacher-facilitated implementation of the Second Step program was overall more frequently utilized than counselor-facilitated, although the use of counselor-facilitated was higher in specific regions of the United States (i.e., Gulf Coast, Southeast, and Mountain Plains).

Controlling for several robust covariates, having some type of support for SEL, whether that be access to federal funding for SEL programming (*via* ESSER funds) or direct support for systemic SEL (*via* CASEL's Collaborating Districts Initiative), was associated with a 64% increase in the likelihood of schools using teacher-facilitated implementation. The most pervasive type of support that could be documented in the current study was funding for SEL programming based on state plans for using ESSER funds. Of course, access to SEL funding was based on school *plans* for spending ESSER funds. It is possible that some schools delayed use of funds. In fact, recent reports of ESSER fund expenditures indicated significant variation in school districts' actual spending relative to what was planned (DiMarco and Jordan, 2022; Edunomics, 2022). Given that the study sample included schools that actually purchased the Second Step program, they may have been more likely to utilize ESSER funds for the SEL program. Regarding consultative support for systemic SEL, only a small number of schools in the study sample had access to direct programming support by way of CASEL's Collaborating Districts Initiative. Studies of the Collaborating Districts Initiative have in fact shown that participating schools have more positive indicators of systemic SEL than non-participating schools (Kendziora and Osher, 2016; Schwartz et al., 2022). In the current study, the strength of the association between SEL support and implementation approach may have been higher if we were able to capture other sources of SEL support at the local or state level to which schools may have accessed. In fact, CASEL has developed a collection of accessible resources that schools and districts can utilize to help support their strategic planning for systemic SEL. The recommendations involve multi-year efforts in which district and school leaders engage all community stakeholders (school staff, students, parents) in supporting and sustaining SEL in schools. Utilization of a teacher-facilitated SEL program does not mean that these conditions for systemic SEL fully exist, but it is, nevertheless, a critical component of the work (CASEL, 2020a; Mahoney et al., 2021).

Regarding state adoption of stand-alone SEL standards/competencies, we found a marginal association between this factor and implementation approach, but only for schools in states that were early adopters of SEL standards/competencies. That is, the likelihood of schools being teacher-facilitated tended to increase in states that were early adopters of SEL standards/competencies. Currently, 54% of states have stand-alone K-12 SEL standards/competencies, the majority of which are located in the Central, Northeastern, and West Coast regions of the country (Dermody and Dusenbury, 2022). A recent evaluation of the role of partisan politics in state adoption of K-12 SEL standards/competencies indicated that there were as many Democratic states with K-12 SEL standards/competences as there were Republican states (Committee for Children, 2020). For those

states in which there is interest in adopting SEL standards, some recommendations include reaching out to neighboring states that do have SEL standards so that they can share their experience with the adoption process, for instance, communicating how the standards were successfully positioned or framed (e.g., as a way to improve school safety or student health) as well as how the standards have positively impacted school learning environments (Committee for Children, 2020).

Having a more recent adoption of SEL standards/competencies (i.e., mid- and late adopters) was not associated with any increased likelihood of teacher-facilitated implementation. Perhaps over time, more schools in later adopting states will extend implementation to teachers, particularly if there are efforts to engage in strategic planning around schoolwide SEL. This idea aligns with Second Step support staff interviews. In states that were more recent adopters of SEL standards/competencies, interview participants indicated that schools seemed to comply with the new standards by initially charging counselors with the role of teaching SEL lessons. Thus, they perceived a *negative* association between having K-12 SEL standards/competencies and teacher-facilitated implementation in the schools that they supported. This perception may be a result of the fact that the respondents primarily worked with schools in the Gulf Coast and Mountain Plains regions, which tend to be more recent adopters of SEL standards/competencies.

Importantly, participants also indicated supporting schools that were, in fact, interested in transitioning from counselor-facilitated to teacher-facilitated implementation. Toward this end, they suggested that a critical step is gaining buy-in for SEL among leaders and teachers. One example of a buy-in strategy described by support staff was based on a diffusion of innovation approach whereby counselors engage a few highly motivated teachers in implementation by inviting them to co-teaching lessons or observe counselors teach lessons. Over time, additional teachers may take up the program as they hear feedback from their colleagues about the ease of implementation and observe positive effects on students. SEL program developers can also help support this process by providing implementation and lesson planning resources for counselors directly within the program along with resources to support classroom and schoolwide reinforcement of SEL skills.

Other potential factors that might make schools more likely to use counselor-facilitated implementation were uncovered from the qualitative study with Second Step support staff. At the district−/school-level, Second Step support staff pointed out that counselor-facilitated implementation might be more likely in districts in which the budget for SEL programming falls under the counseling or student services departments. This factor was not captured in the aim one analysis. However, the analysis did include state-level counselor mandates as a covariate, which as hypothesized, decreased the likelihood of schools utilizing teacher-facilitated implementation. Perhaps schools in states with counselor mandates have larger counseling departmental budgets, which might allow them to have greater discretion in using funds for universal SEL programs. Some urban school systems have intentionally situated SEL within their teaching and learning departments (e.g., Atlanta Public Schools), as a signal that SEL should be prioritized equally alongside other academic subject areas.

Additional factors at the individual and school-level that Second Step staff suggested were related to decisions to utilize counselor-facilitated implementation included: (1) low SEL buy-in among leaders and teachers and (2) limited staff capacity. Lack of buy-in from leaders is especially concerning as support from school leadership is a consistent predictor of successful SEL program implementation (Elias et al., 2000; Durlak and Dupre, 2008). In addition, assigning SEL to counselors could be an attempt to reduce the burden on already overwhelmed teachers. The interviews further indicated that district and school administrators may assume that compared to counselors, teachers do not have as great a level of expertise in social–emotional development, and some may also interpret ASCA guidelines to mean that SEL should be strictly in the lane of counselors.

Overall, Second Step staff agreed that counselor-facilitated implementation has its challenges, including counselor burnout, little reinforcement of students' SEL skills outside of direct instruction, and lack of alignment with counselors' skills as a classroom instructor. Research shows that these types of challenges are associated with lower quality of program implementation which, in turn, makes students less likely to benefit from the program (Domitrovich et al., 2008; Durlak and DuPre, 2008). Some of the challenges identified are not specific to counselors – even when teachers are the facilitators of SEL lessons, they may face similar challenges in regard to burnout and difficulty reinforcing skills outside of SEL lesson instruction time (Ransford et al., 2009; Durlak, 2016). One challenge mentioned that may be of particular concern for the quality of counselor-facilitated implementation is the lack of training that counselors receive in classroom management. Good classroom management is a core pedagogical competency needed for effective instruction and student learning. Poor classroom management is associated with student problem behaviors and low student engagement (Korpershoek et al., 2016). Furthermore, in SEL implementation studies, teachers' efficacy for classroom management positively predicted fidelity of program implementation, as measured by dosage of lessons completed (Rimm-Kaufman and Sawyer, 2004; Thierry et al., 2020). If counselors are charged with implementing lessons in classrooms, it would be important to provide them with professional learning in classroom management.

Study limitations and directions for future research

The current study identified factors that may contribute to more schoolwide SEL, with support for SEL emerging as a significant factor associated with teacher-facilitated implementation. No previous studies have examined the frequency of counselor-facilitated implementation of SEL programs in general or how different school/district- and macro-level factors are associated with implementation approach. However, several limitations of the study

should be mentioned. Although we controlled for school-level demographic characteristics, we did not have access to other variables that could make a difference in implementation approach, including utilization of other sources of support for systemic SEL, how SEL is situated within the organizational hierarchy of districts (e.g., within Teaching and Learning, Student Services, etc.), and direct indicators of systemic SEL (e.g., SEL integration with academic instruction, supportive disciplinary policies and practices, positive school and classroom climate).

Also, the only outcome measure examined was type of implementation approach. We did not have access to reliable data capturing the fidelity with which the program was implemented (i.e., completion of digital lessons). In addition, because the outcome measure was bimodally distributed and we lacked school-level data on the specific context of implementation approach being utilized, we were unable to capture schools that may have been in a more transitional phase, as described in the qualitative portion of the study. Last, the qualitative portion of the study focused on the perspective of only five support staff representing one SEL provider (i.e., Second Step) and did not directly capture the voices of counselors, administrators, teachers, and other support staff in schools.

These limitations could be addressed in future studies by including these additional predictor and outcome variables, perhaps using a longitudinal cohort design. For instance, schools in states that are recent adopters of K-12 SEL standards/competencies could be studied over time to more closely examine decision-making processes involved in implementation approaches, especially related to explicit strategies for those that utilize counselor-facilitated approaches and strategies in any transitions to teacher-facilitated implementation. Additionally, the inclusion of the voices of all stakeholders (school leaders, counselors, support staff, teachers) would allow for a deeper understanding of how factors within school systems, particularly the individual- and school-related factors highlighted by Second Step support staff, affect implementation decision-making and subsequent quality of program implementation.

Data availability statement

The datasets presented in this article are not readily available because we do not have approval to share the datasets outside the organization. Requests to access the datasets should be directed to KT, kthierry@cfchildren.org.

Ethics statement

The studies involving human participants were reviewed and approved by Pearl IRB. The participants provided their written informed consent to participate in this study.

Author contributions

KT contributed to conception and design of study, statistical analysis, and writing of the manuscript. AP contributed to conception and design of the study and wrote sections of the manuscript. CC contributed to design of the study, data collection, and wrote sections of the manuscript. JP contributed to acquisition of data and interpretation of results. YL performed statistical analysis. JC contributed to design of the study and wrote sections of the manuscript. HR organized the database, helped with data collection, and performed the qualitative analysis. PR helped with data collection and qualitative analysis. TK and SW contributed to manuscript revision. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Aarons, G. A., and Sommerfeld, D. H. (2012). Leadership, innovation climate, and attitudes toward evidence-based practice during a statewide implementation. *J. Am. Acad. Child Adolesc. Psychiatry* 51, 423–431. doi: 10.1016/j.jaac.2012.01.018

AASA (2022). School district spending of American rescue plan funding: part II. *School Superint. Assoc.* https://aasa.org/uploadedFiles/ARP-Survey-Findings-Part2.pdf

American School Counselors' Association. (2022). State school counseling mandates and legislation. Available at: https://www.schoolcounselor.org/About-School-Counseling/State-Requirements-Programs/State-School-Counseling-Mandates-Legislation

Atkins, M. S., Frazier, S. L., Leathers, S. J., Graczyk, P. A., Talbott, E., Jakobsons, L., et al. (2008). Teacher key opinion leaders and mental health consultation in low-income urban schools. *J. Consult. Clin. Psychol.* 76, 905–908. doi: 10.1037/a0013036

Atkins, M. S., McKay, M. M., Arvanitis, P., London, L., Madison, S., Costigan, C., et al. (1998). An ecological model for school-based mental health services for urban

low-income aggressive children. *J. Behav. Health Serv. Res.* 25, 64–75. doi: 10.1007/BF02287501

Ballotpedia (2022). State government trifectas. Available at: https://ballotpedia.org/State_government_trifectas

Brauer, M., and Curtain, J. J. (2018). Linear mixed effects models and the analysis of non-independent data: a unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychol. Methods* 23, 389–411. doi: 10.1037/met0000159

Burchinal, M., Vandergrift, N., Pianta, R. C., and Mashburn, A. J. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. *Early Child. Res. Q.* 25, 166–176. doi: 10.1016/j.ecresq.2009.10.004

CASEL. (2020a). *CASEL's district resource center*. Chicago, IL: Author. Available at: https://drc.casel.org/

CASEL. (2020b). *The CASEL guide to schoolwide social and emotional learning*. Chicago, IL: Author. Available at: https://schoolguide.caselorg/

CASEL (2021). *2011-2021: 10 Years of Social and Emotional Learning in U.S. School Districts: Elements for Long-Term Sustainability of SEL*. Available at: https://casel.org/cdi-ten-year-report/

Committee for Children. (2020). Politics may not impede SEL development. Available at: https://www.cfchildren.org/wp-content/uploads/policy-advocacy/politics-may-not-impede-sel-development-brief.pdf

Corbin, J., and Strauss, A. (2014). Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. *3rd Edn*. Thousand Oaks, CA: Sage Publications.

Dermody, C., and Dusenbury, L. (2022). *Social and emotional learning state scorecard scan*. Chicago: Collaborative for Academic, Social and Emotional Learning.

DiMarco, B., and Jordan, P. W. (2022). *Financial trends in local schools' COVID-aid spending* FutureEd. Georgetown University, Washington DC. Available at: https://www.future-ed.org/financial-trends-in-local-schools-covid-aid-spending/.

Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., et al. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: a conceptual framework. *Adv. School Ment. Health Promot.* 1, 6–28. doi: 10.1080/1754730x.2008.9715730

Downer, J. T., Lopez, M. L., Grimm, K. J., Hamagami, A., Pianta, R. C., and Howes, C. (2012). Observations of teacher–child interactions in classrooms serving Latinos and dual language learners: applicability of the classroom assessment scoring system in diverse settings. *Early Child. Res. Q.* 27, 21–32. doi: 10.1016/j.ecresq.2011.07.005

Durlak, J. A. (2016). Programme implementation in social and emotional learning: basic issues and research findings. *Camb. J. Educ.* 46, 333–345. doi: 10.1080/0305764X.2016.1142504

Durlak, J. A., and DuPre, E. P. (2008). Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am. J. Community Psychol.* 41, 327–350. doi: 10.1007/s10464-008-9165-0

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

Dusenbury, L., Yoder, N., Dermody, C., and Weissberg, R. (2020). An Examination of K-12 SEL Learning Competencies/Standards in 18 States.. *Frameworks Briefs*. Available at: https://casel.org/cdi-ten-year-report/

Dusseault, B., and Pillow, T. (2021). First look at ESSER priorities: districts are placing their bets on what they know: CRPE. Available at: https://crpe.org/first-look-at-esser-priorities-districts-are-placing-their-bets-on-what-they-know/

Edunomics (2022). *ESSER expenditure dashboard* Edunomics Lab. Georgetown University, Washington DC. Available at: https://edunomicslab.org/esser-spending/.

Elias, M. J., Bruene-Butler, L., Blum, L., and Schuyler, T. (2000). Voices from the field: identifying and overcoming roadblocks to carrying out programs in social and emotional learning/emotional intelligence. *J. Educ. Psychol. Consult.* 11, 253–272. doi: 10.1207/S1532768XJEPC1102_06

FutureEd (2022). How states are using federal funds for learning recovery. Available at: https://www.future-ed.org/how-states-are-using-federal-funds-for-learning-recovery/

Jones, D. E., Greenberg, M., and Crowley, M. (2015). Early social emotional functioning and public health: the relationship between kindergarten social competence and future wellness. *J. Public Health* 105, 2283–2290. doi: 10.2105/AJPH.2015.302630

Kendziora, K., and Osher, D. (2016). Promoting children's and adolescents' social and emotional development: district adaptations of a theory of action. *J. Clin. Child Adolesc. Psychol.* 45, 797–811. doi: 10.1080/15374416.2016.1197834

Klem, A. M., and Connell, J. P. (2004). Relationships matter: linking teacher support to student engagement and achievement. *J. Sch. Health* 74, 262–273. doi: 10.1111/j.1746-1561.2004.tb08283.x

Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., and Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Rev. Educ. Res.* 86, 643–680. doi: 10.3102/0034654315626799

MacNeil, A. J., Prater, D. L., and Busch, S. (2009). The effects of school culture and climate on student achievement. *Int. J. Leadersh. Educ.* 12, 73–84. doi: 10.1080/13603120701576241

Mahoney, J. L., Weissberg, R. P., Greenberg, M. T., Dusenbury, L., Jagers, R. J., Niemi, K., et al. (2021). Systemic social and emotional learning: promoting educational success for all preschool to high school students. *Am. Psychol.* 76, 1128–1142. doi: 10.1037/amp0000701

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., et al. (2008). Measures of classroom quality in pre-kindergarten and children's development of academic, language, and social skills. *Child Dev.* 79, 732–749. doi: 10.1111/j.1467-8624.2008.01154.x

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., et al. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proc. Natl. Acad. Sci. U. S. A.* 108, 2693–2698. doi: 10.1073/pnas.1010076108

National Commission on Social, Emotional, and Academic Development. (2018). *From a nation at risk to a nation at hope*. Washington, DC: The Aspen Institute.

Ransford, C. R., Greenberg, M. T., Domitrovich, C. E., Small, M., and Jacobson, L. (2009). The role of teachers' psychological experiences and perceptions of curriculum supports on the implementation of a social and emotional learning curriculum. *Sch. Psychol. Rev.* 38, 510–532.

Rimm-Kaufman, S. E., and Sawyer, B. E. (2004). Primary-grade teachers' self-efficacy beliefs, attitudes toward teaching, and discipline and teaching practice priorities in relation to the "responsive classroom," approach. *Elem. Sch. J.* 104, 321–341. doi: 10.1086/499756

Scaccia, J. P., Cook, B. S., Lamont, A., Wandersman, A., Castellow, J., Katz, J., et al. (2015). A practical implementation science heuristic for organizational readiness: R = MC2. *J. Community Psychol.* 43, 484–501. doi: 10.1002/jcop.21698

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

Thierry, K. L., Vincent, R. L., and Norris, K. (2020). Teacher-level predictors of the fidelity of implementation of a social emotional learning curriculum. *Early Educ. Dev.* 33, 92–106. doi: 10.1080/10409289.2020.1849896

Appendix

Interview questions:

1. What does counselor-facilitated implementation mean in your context?
2. Why do districts/schools choose counselor-facilitated implementation over teacher-facilitated implementation?
3. Do counselors experience any challenges in implementation that are distinct from those experienced when teachers are the implementers? If yes, what are those distinct challenges?
   a. Have you heard of counselor-specific challenges around school-level outcomes (like creating a shared language or common strategies that all educators can support?)
4. How are we (i.e., Second Step) currently able to effectively support counselors through these challenges? Please be specific about services and/or resources that currently exist in the platform and those created by client-facing staff.
5. Do we (i.e., Second Step) provide counselors with any guidance or resources to support transitioning implementation to teacher-facilitated? Please describe the guidance or resources provided.

frontiers | Frontiers in Education

TYPE Curriculum, Instruction, and Pedagogy
PUBLISHED 20 December 2022
DOI 10.3389/feduc.2022.1020204

OPEN ACCESS

EDITED BY
Celene Domitrovich,
Georgetown University,
United States

REVIEWED BY
Jennifer LoCasale-Crouch,
Virginia Commonwealth University,
United States
Yuli Utanto,
State University of Semarang, Indonesia

*CORRESPONDENCE
Barbara Wilder-Smith
bwildersmith@toolsofthemind.org

†These authors share first authorship

SPECIALTY SECTION
This article was submitted to Educational Psychology, a section of the journal Frontiers in Education

RECEIVED 15 August 2022
ACCEPTED 30 November 2022
PUBLISHED 20 December 2022

CITATION
Spacciapoli M, Viana M, Saunders Wilder O, Sullivan J, McCallum T and Wilder-Smith B (2022) An equitable and scalable approach to track fidelity of implementation in partnership with teachers.
Front. Educ. 7:1020204.
doi: 10.3389/feduc.2022.1020204

An equitable and scalable approach to track fidelity of implementation in partnership with teachers

Megan Spacciapoli†, Mara Viana†, Oliver Saunders Wilder†, Jillian Sullivan, Tia McCallum and Barbara Wilder-Smith*

Tools of the Mind, Shirley, MA, United States

Social and emotional learning (SEL) is a key focus of early childhood education. A significant body of research demonstrates the critical impact of fidelity of SEL curriculum and intervention implementation on child outcomes; however, few widely-used curricula regularly assess classroom-level implementation fidelity outside of the context of research or correlate fidelity with key areas of children's development of these skills. Fidelity measurement often focuses on easy-to-quantify variables such as classroom environment or lesson plan adherence, and is a periodic snapshot used as a moderator or co-variate when assessing child outcomes, rather than being intentionally leveraged as a systematic, ongoing process to evaluate and support implementation. In this paper, we present a novel approach to capturing fidelity data as a core component of professional development. We outline our findings from a pilot of our approach using short, teacher-recorded videos submitted across the school year as a vehicle for capturing and sharing real-time data related to professional learning, implementation, and curricula impact, as well as a framework for building equitable partnerships with teachers. Results from the initial pilot of this approach in several hundred classrooms across the US demonstrate feasibility and utility and suggest that teacher-recorded videos can offer a scalable means to collect continuous samples of fidelity data, providing a richer view of professional learning, while simultaneously creating the opportunity to provide ongoing feedback and engage teachers in partnership in reflecting on practice and its impact on children's development. We developed and piloted an approach where teachers record and upload videos of teaching practices and children engaging with their peers in specific classroom activities *via* a mobile application. Each video submission has a focal activity and associated set of indicators which are shared with teachers in advance to create an equitable feedback system in which both curriculum staff and teachers engage in reflecting on children's interactions and the application of the curricular approach in their classroom. Videos are viewed and coded on these sets of indicators by both the teachers and curricular coaches who provide targeted feedback in an interactive exchange on a dashboard accessible by teachers and their curricular coach.

KEYWORDS

implementation fidelity, video coaching, partnership with teachers, SEL, professional development model, educational technology, Tools of the Mind, early childhood

Introduction

Social and emotional learning is a critical component of early childhood education and a target of early childhood interventions and curricula. This key area of development is emphasized in the early childhood years because the social and emotional skills learned early in life serve as a foundation for future social interactions, emotion regulation, and are linked to improved life outcomes in school and beyond (Moffitt et al., 2011; Jones et al., 2015; Taylor et al., 2017). A significant body of research shows that when a curriculum is implemented with high fidelity, there is a positive impact on children's academic and social–emotional outcomes (Durlak and DuPre, 2008; Hamre et al., 2010; Durlak et al., 2011; Sklad et al., 2012; Quinn and Kim, 2017).

While extant research shows a strong relationship between fidelity and child outcomes, the practice of measuring fidelity to a curriculum remains highly variable, with practices ranging from ratings by program experts *via* in-classroom observation, videotaping, interviews, activity checklists and/or record reviews (Mowbray et al., 2003; Bickman et al., 2009). There is a need for continued work in the field to better operationalize fidelity in a way that provides teachers ongoing support, an equitable experience, and enables high-level implementation (e.g., Pianta, 2005; Landry et al., 2006; Pianta et al., 2008).

Measuring fidelity of implementation is a vital factor in examining the impact of interventions and curricula (de Leeuw et al., 2020), but measuring implementation in a supportive, cost-effective and scalable way has not always been feasible (Bickman et al., 2009). Traditional observation-based implementation fidelity data collection requires observers to be highly trained in the curriculum and fidelity measure, deemed reliable on the measure, and complete in-person classroom observations or video data collection multiple times across a school year. This approach requires training and travel time, both of which involve additional costs. While these observation-based methods of measuring fidelity are the "gold standard," executing them in person is time and labor-intensive, resulting in higher costs and challenges to scale (Huntley, 2009; Barton et al., 2017). In addition, teachers can be wary of outside observers in their classrooms, assuming evaluation or fearing punitive outcomes as a result of what is (or is not) observed (Shernoff et al., 2017). These approaches can often leave teachers in the dark in terms of what is being observed, how they are doing in implementing the intervention or curriculum and creates an inequitable relationship with teachers.

Adding an additional layer of challenge, fidelity measurement can be more complicated in interventions and curricula where practice may involve responsive differentiation of instruction and multiple activities. A window of opportunity to address these challenges coincides with the fact that educators implementing new curricula and interventions need professional development to support implementation of the curriculum to fidelity (Pianta et al., 2008).

In response to these challenges, Tools of the Mind developed TREE (Teachers Reaching Educational Excellence), a novel, scalable approach that re-conceptualizes fidelity measurement as part of on-going professional development. In this paper, we describe our approach and our experience piloting TREE in classrooms learning the Tools of the Mind curriculum during the 2021–2022 school year to illustrate the feasibility and promise of using teacher-recorded video data to measure fidelity of implementation and provide individualized professional development to support continuous learning and mastery. The use of video to provide a "window into practice" provides teachers the opportunity to reflect on their practice in a way that is not possible while in the midst of teaching (Clarke and Hollingsworth, 2002; Borko et al., 2008; Zhang et al., 2011; Marsh and Mitchell, 2014), and a cost-effective and wider-reaching solution for curricular coaches to work in tandem with teachers to review, reflect and support what is happening in the classroom. It is also a more cost-effective approach in comparison to higher cost repeated in-person coaching visits (Dede et al., 2009; Barton et al., 2017).

Fidelity of implementation has been defined in a range of ways, including adherence to an intervention or program as designed by the developers, "strict adherence" to methods and theory, "completeness" of implementation, or the quality of program delivery (Dusenbury et al., 2003; Carroll et al., 2007). A definition that emphasizes adherence to methods and theory is most appropriate for an intervention or curriculum such as Tools of the Mind (Tools), that views teachers as central in its theory of change and highlights the role of educators and their professional learning as a critical factor in fidelity of implementation. Defining fidelity of implementation as the skillful application of methods and theory (Dusenbury et al., 2003) opens the gateway to a continuous approach to measuring fidelity of implementation. A continuous approach is well-designed to capture the developmental trajectory of professional learning as teachers are applying new knowledge, teaching practices and strategies in new activities.

Early childhood educators implementing a new intervention or curriculum that requires adherence to methods and theory need to develop a solid understanding of and practice applying the methods and theory as well as an understanding of the key components (de Leeuw et al., 2020). Professional development and support in implementing the intervention or curriculum to fidelity would be a natural component of interventions and curricula that build teacher understanding to apply an approach and methods. Early childhood interventions and curricula rely on teachers as the key factor in the equation to impact child outcomes; fidelity measurement in turn, needs to capture the developmental trajectory of professional learning as teachers are applying new knowledge, teaching practices and strategies in new activities in a scalable, reliable and feasible way.

This paper offers a case study of an approach that roots fidelity measurement in ongoing teacher professional development to better support teachers as they learn, adopt, and utilize a curriculum. This approach engages teachers in the opportunity to reflect on practice and receive individualized feedback at the same time as a stream of real-time data capturing implementation,

teacher learning, and a developmental trajectory of teachers' fidelity of implementation is collected.

There are five overarching ideas that may prove useful to others in the field that come out of the pilot of TREE discussed in this paper: conceptualizing fidelity of implementation as a developmental journey; integrating fidelity measurement and responsive support of teacher learning; teachers benefit when fidelity of implementation measurement and tracking is an equitable partnership; and teacher-recorded video data provides a feasible and cost-effective approach that offers significant long-term benefits for learning and research.

This approach can offer a transformative model of fidelity measurement with application to a variety of SEL interventions and curricula that have teachers at the center of their theory of change, and value creating an equitable relationship with teachers, while offering a cost-effective and scalable approach to fidelity measurement.

Pedagogical framework

Tools of the Mind curriculum

While TREE was developed in the context of a specific curriculum (Tools of the Mind), we believe that the principles and methodology are broadly applicable to the field and can be used as a case study for classroom-based curricular interventions that are working to improve child outcomes in early childhood education. Tools of the Mind (Tools) is a comprehensive, research-based, early childhood curriculum rooted in Vygotskian theory and designed to embed self-regulation and social–emotional development in comprehensive PreK & K curricula, that has been identified as a CASEL-SELect program [Collaborative for Academic, Social, and Emotional Learning (CASEL), 2022]. From its very start, Tools was developed in partnership with teachers.

Over the last 30 years Tools has grown from a university-sponsored research project to become an independent nonprofit working with school and community-based partners across the country. Tools mission is to empower teachers with the tools they need to build inclusive classroom communities, leverage playful learning and individualize instruction to enable every child to reach their full potential. Tools has a decades-long track record of using research to continuously improve its program and approach to measuring fidelity of implementation, learning from participation in multiple research studies (e.g., Diamond et al., 2007, 2019; Barnett et al., 2008; Blair and Raver, 2014).

Tools' underlying theory and understanding of development extends to our design of teacher professional development. Over the course of learning Tools across their first year, teachers internalize concepts, theory and neuroscience research that applies to practice; 'learning by doing' by taking concepts taught in core workshops, and applying them in their classroom, reflecting on their impact on children through recording and sharing video on TREE. As teachers progress along their unique developmental trajectories, they are incorporating new approaches to teaching, implementing curricular strategies, and learning through and alongside the children in their classroom. We believe this continuous process of learning provides a unique opportunity to impact teacher and child development through providing responsive, customized scaffolding.

TREE background and motivation

The TREE approach to individualized teacher professional development and measuring fidelity came out of years of developing and using different approaches of capturing fidelity of implementation of the Tools program. Over the past two decades, multiple variations of fidelity measurement tools were used, including one that focused on all of the steps and their sequence for each Tools activity. The lesson from these earlier iterations revealed that the definition of fidelity of implementation, and how it is communicated to both teachers and researchers has an impact on implementation and child outcomes. This has led us to simplify our curriculum manuals and completely redesign our professional development, engaging in iterative research on teacher and child impacts through our data share partnerships with programs implementing Tools. It also led to a clearer conceptualization of fidelity of implementation and a new approach to capturing and measuring it through focusing on a smaller set of indicators that evolve over time. This allows Tools teachers to focus on the teaching practices and child interactions and actions that matter most–the "critical components" that are vital to ensuring the program is being implemented (Century et al., 2010; Stains and Vickrey, 2017). It also provides teachers transparency and an implementation road map with clear markers to focus on over time, and individualized support to enable teachers to gain a deeper understanding of the Tools approach, and how to apply it in their classrooms. In concert, identifying core components of the program and providing them clearly to teachers, strengthens fidelity of implementation (de Leeuw et al., 2020).

Prior to the creation of TREE, Tools provided coaching and captured fidelity of implementation through in-person technical assistance visits to schools. These visits would take place at one or two points during the year. The logistics of in-person visits meant that Tools' curricular coaches rarely were able to visit every classroom in a school to observe key teaching practices and activities in action and were unable to individually engage in 1:1 coaching with every teacher. In designing TREE, a key motivation was to find an alternative approach that would allow for more immediate and consistent feedback to every teacher, while also capturing and measuring fidelity of implementation over time.

As a curriculum that has from its origin been co-constructed with teachers, equity in our partnership with teachers is a core value. We believe that the measurement and tracking of fidelity of implementation can and should be inclusive of teachers as collaborators. Historically in the majority of our research

experiences, teachers were blind to what was being looked for in in-person fidelity measurement observations and video-recordings. They did not have the opportunity to learn about the impact on children until the end of year, or sometimes several years later, and were often blind to how their fidelity of implementation was scored. Engaging teachers in actively observing the impact of their teaching practices on children and connecting how they implement the curriculum to what they observe in children's ongoing development has tremendous value for them.

Building on the design of Tools' curriculum and professional development, we approached the creation of our fidelity instrument through a developmental lens. Rather than starting with a notion of 'perfect implementation' of a teacher with significant experience, we began by creating a developmental trajectory of Tools teachers' growing mastery applying Tools approach, and implementation of the curriculum to create fidelity indicators over time based on our prior data from in-person technical assistance visits. We also reviewed previously used fidelity assessment measures, distilling the 'key steps' in Tools activities, as well as creating indicators to capture teachers' responsive increase of challenge level and modification of teaching practices to support children as they develop across the year.

TREE learning environment and pilot learning objectives

Developing an integrated model of professional development and fidelity measurement

Tools' model of professional development consists of a foundation of live interactive workshops spread across the year. In workshops, teachers learn Tools' theory, how to implement Tools activities, and strengthen their knowledge of child development, while developing new teaching practices and strategies to scaffold and individualize instruction and build inclusive classroom cultures.

These workshops are paired with small-group Tools-facilitated Professional Learning Community (PLC) live virtual sessions focused on team building and co-construction between teachers. These sessions provide teachers with a cohort of peers to engage in discussion based on the groups' collective needs and challenges as they progress across the year.

The last component is our individualized professional development on TREE. TREE provides teachers with transparent information about the anticipated trajectory of implementation with a roadmap with clear benchmarks for implementation over time. The TREE model integrates professional learning with an approach to capturing the trajectory of fidelity of implementation in early childhood classrooms. Conceptualized over a 2-year period, segmented into the 1st and 2nd years of professional learning, with a subsequent 3rd year option to become certified as masterful Endorsed Tools teachers, this model builds a bridge between theory and practice. TREE engages teachers in closely observing and capturing videos of their teaching practices, children's actions and interactions, then viewing the video to reflect on practice and the indicators they are focused on in a given month. TREE creates an environment in which the fidelity of implementation in the classroom is captured while simultaneously supporting the individual trajectory of implementation development for each teacher.

Pilot study

In the pilot study of TREE we had the goal to assess the feasibility and usability of teacher-recorded video and identify if it is a cost-effective approach to capturing fidelity of implementation in an embedded system of individualized professional development.

Our primary learning objectives include:

- Is teacher-recorded video feasible and useful for fidelity assessment?
- Will teachers find the TREE approach supportive?
- Will our curricular coaches find TREE feasible and useful?
- Will this approach be cost-effective and scalable?

Participants

A total of 301 PreK and Kindergarten teachers from 65 different school districts across the United States took part in the initial pilot of TREE during the 2021–2022 school year. Teachers had between 0 and 34 years of teaching experience, with a median of 7 years of experience in the classroom. Participating school districts varied in both socioeconomic and racial/ethnic diversity, as measured by publicly available data of the percentage of students eligible for free or reduced-price lunch and Ethnic Diversity Index data. School districts in the pilot had a median of 43% (IQR = 33%) of students eligible for free or reduced-price lunch, and a median ethnic diversity index score of 0.59 (IQR = 0.28) as shown in Table 1.

Pedagogical framework of TREE

Anatomy of TREE cycles

In TREE, there are series of video cycles in each year teachers are learning and mastering Tools' approach, with select focus areas

TABLE 1 Participant demographics.

Characteristic	N = 301[1]
Mobile App Installed	249 (83%)
Years of Teaching Experience	7 (0, 15)
School District - % Free/Reduced Lunch	43 (18, 46)
School District - Diversity Score	0.56 (0.36, 0.62)

[1]n (%); Median (IQR).

that capture children's engagement in learning activities, interactions, and teachers' use of strategies in Tools activities designed to support self-regulation and social skill development as well as academic skills. The teaching practices and strategies taught in professional development become distilled into select actionable and measurable indicators in each video cycle. The indicators in a given activity are selected because they capture child behavioral impacts of fidelity of implementation of teaching practices or capture the teaching practices in action (see Figure 1). These indicators are chosen because they are generalizable teaching and learning practices that are applied in multiple Tools activities across the day, so teacher and/or child mastery in one activity extends across the day and correlates with fidelity of implementation.

Cycles are conceptualized as an equitable feedback system between the teacher, their curricular coach, and their peer PLC group as displayed in Figure 2. The format of each cycle consists of two video submissions, punctuated by a narrative feedback loop with the curricular coach tagged to key moments in the video, and a facilitated PLC meeting with their cohort.

Recording and submitting video

For each focus area, teachers record a short (≤8 min) video of children's actions and interactions in an activity as well as teacher interactions with children. Teachers are reminded frequently that the aim is not perfection, but instead to capture current development, as an opportunity for reflection and learning, to propel teachers' and children's development. Prior to submitting their video, the teacher reviews the indicators associated with the focus area and identifies those that are observed in their video. This encourages the teacher to reflect on their own practice and children's interactions using the same measurement tool that their curricular coach will use. This performs a central function in the equitable feedback system by providing a core set of achievable goals that support teachers' intrinsic motivation, ownership, developing understanding, reflective practice, and growth. Over time, TREE provides teachers a visualization of their integration of Tools approach and implementation of the curriculum in their classroom, and continuous growth across the year (see Figure 3), and they can look back in year 2 at where their children were, and what they were doing at a similar time in the previous year.

Feedback cycle

Following the submission of a video and self-identification of indicators, teachers receive feedback from their coach through an interactive interface that allows for the coach to tie written comments to time-stamped sections of the teacher's video submission. The coach may ask questions and comment on observed areas of focus to identify strengths and make suggestions about next potential steps based on teachers' and children's current levels of development. Teachers provide feedback to coaches about what is most applicable and helpful, as well as ask and respond to questions. The platform allows teachers to reflect on the coaches' feedback, rewatch key areas of their video, and decide what strategies they want to focus on to strengthen their implementation, and children's outcomes.

Professional learning community (PLC) meetings and shared learning

After receiving feedback from their coach, the teachers participate in facilitated PLC cohorts where they can bring questions and discuss topics relevant to the current area of focus to deepen their understanding of core teaching practices and the curriculum alongside their dedicated coach and colleagues. This time is used to deepen understanding of specific concepts, brainstorm strategies for common classroom scenarios, provide support, and build camaraderie.

Visualization of growth

Following their PLC meeting, teachers record and submit a second video of the same focus area and indicators, applying their strengthening understanding of the theory and child impact goals from their reflective practice, coaching and PLC discussions. Teachers view their video, identify the indicators they observe, write questions or notes to their coach, and their curricular coach provides feedback and captures the teacher's current fidelity of implementation by checking the indicators observed in the video. When the coach has completed this, the teacher's TREE dashboard animates to highlight growth in the form of a tree representing each video focus. Each tree starts from a seed that germinates, sprouts, and matures into a sapling, followed by a mature tree that blossoms and bears fruit. The teacher then begins the next cycle, represented by a new tree with a different focus area and indicators tied to that tree's growth. The tree visual mediator follows the teacher throughout the year, and over the course of a two-year period of professional development, after which teachers have the opportunity to become endorsed.

The curricular coach TREE experience

Our curricular coaches have their own TREE dashboards, which enables them to see the development of fidelity of implementation across all their classrooms, as well as displaying data at a school and individual classroom level. This data is used to inform customization of PLCs and workshops to meet teachers where they are and support ongoing learning and implementation.

For the curricular coaches providing dedicated support to teachers, TREE provides a modality for capturing fidelity of implementation, interaction and feedback that bypasses many challenges associated with traditional models of technical assistance. TREE provides a glimpse into the classroom, as one Coach reflected, "at exactly the right moment," and equips coaches with an interface that supports their workflows. Dashboard features, like the time-stamp functionality and categories of feedback, make engaging with the content easy, meaningful, and systematized. Coaches can reflect on the similarities and differences between their and teachers' observation of indicators

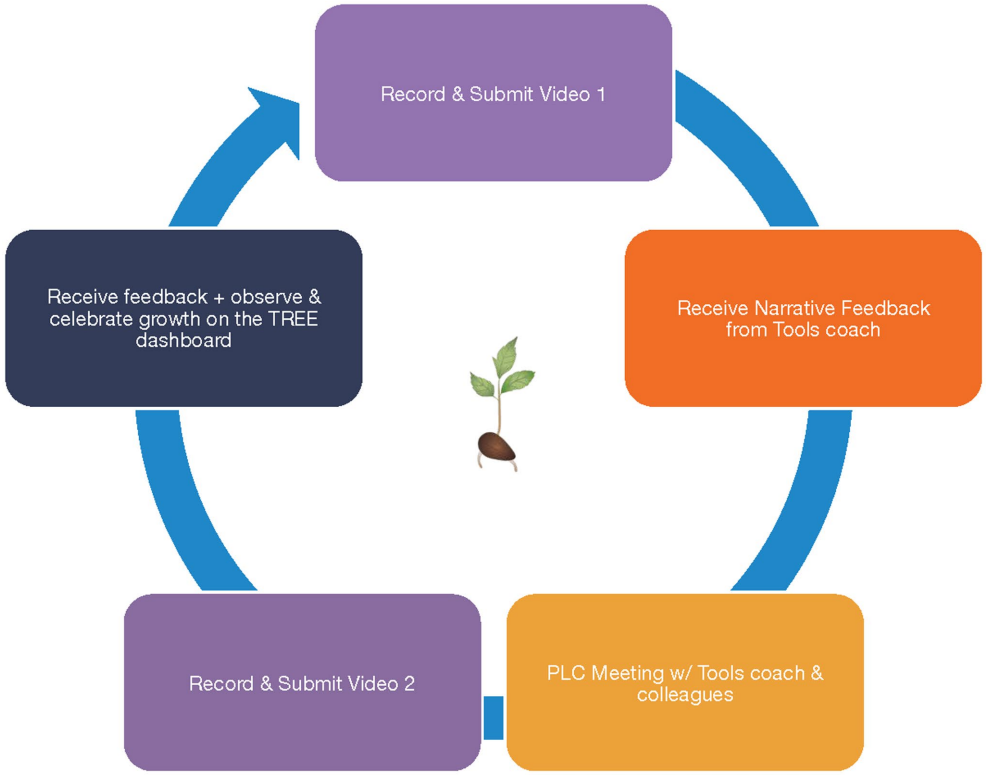


FIGURE 1
Tools TREE cycle.

and apply this awareness in their coaching interaction. Coaches can rewatch a video or rewind it to better observe or hear something – something not possible in live in-person visits. Capturing fidelity of implementation involves checking a short list of indicators, the same short list of indicators across all the classrooms they are coaching in a given month, building their capacity and experience, and making fidelity of implementation capturing more accurate.

For many coaches, TREE provides, for the first time, the ability to see the impact of their scaffolding – with visibility into whether more indicators are captured in the second video of the same activity in each cycle, and what new strategies are adopted by teachers to strengthen application of Tools approach and curriculum implementation.

TREE training process for curricular coaches

Our curricular coaches are employees of Tools of the Mind, and former Tools of the Mind teachers or coaches. Initial training for coaches starting to use TREE was over the course of 3 days. This training introduced the video focuses and indicators, how to download the app and navigate the platform and engaged coaches in shared practice sessions with video. Following this, coaches had monthly 2-h professional development sessions focused on each set of indicators and how to identify them, and how to apply Tools theory to support teacher development, with a focus on identifying the next steps that teachers will be able to independently apply after coaching to support their continuous growth in implementing Tools.

This training has enabled Tools to onboard new team members, oversee and support development of reliability in fidelity measurement, and monitor and learn about the effectiveness of coaching strategies. This combination of data visibility and shared learning has been an effective way to support team growth.

Results

Participants access to technology

All participating teachers had access to high-speed internet, and 83% of participating teachers chose to install and use the TREE mobile app available for Apple iOS and Android mobile devices. The remaining 17% utilized our web browser-based alternative for uploading video.

A total of 1,140 videos were recorded and submitted by teachers during our TREE pilot in the 2021–2022 school year. Submitted videos had a median duration of approximately 7.3 min (Median Duration = 440 s; IQR = 392 s), and median file size for uploaded videos was 491 MB (IQR = 572 MB). The majority (90%)

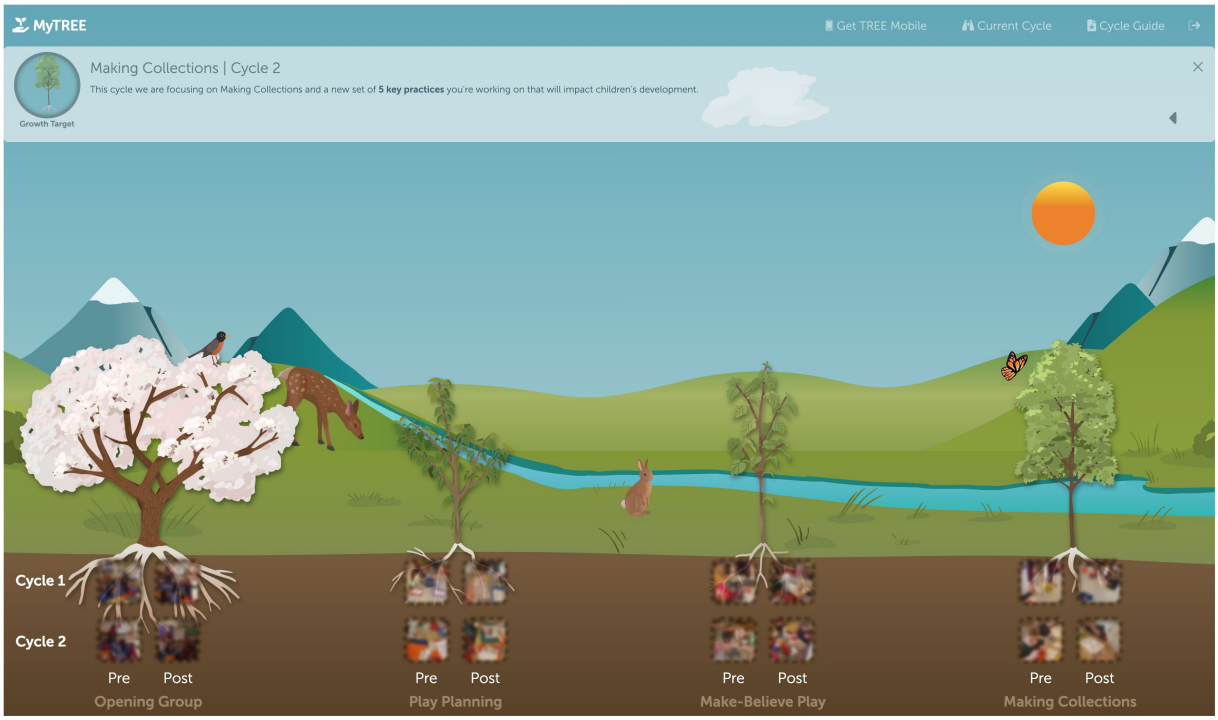


FIGURE 2
Teacher TREE growth visualization.

TABLE 2 TREE video descriptive statistics.

Characteristic	N = 1,400[1]
File Size (MB)	**491 (258, 830)**
Video Duration (seconds)	**440 (261, 653)**
Review Time (mins)	**17 (10, 30)**

[1]Median (IQR).

of videos submitted by teachers showed evidence of targeted classroom practices, and few teachers reported difficulty with recording videos (Table 2).

Feasibility and utility

Overall, internal product feedback and data usage from both teachers and curricular coaches indicated that the TREE model was convenient, user-friendly, and helpful in daily teaching practice. 100% of teacher-recorded videos were correctly formatted and submitted.

Of the teachers responding to a feedback survey ($n = 81$), 91% of teachers in their first year of learning Tools felt supported by their experience with TREE. One teacher relayed that they *"felt extremely motivated after each video feedback. The suggested ideas were always so beneficial. […] It was truly a valuable support to have while learning a new curriculum,"* and another wrote *"It is so worth it. The insight you get back not only improves your teaching strategies but the students' learning."*

In terms of user experience, our curricular coaches reported that teachers had few difficulties submitting videos and using the TREE interface, one coach writing "*even my teachers that were reluctant in using the mobile app thought it was easy to use*," and another said, "*teachers seemed to like the ease of it*." A key concern when introducing any new technology is the amount of technical support required to ensure user success in adopting and using it. We provided users with support *via* email and our website during business hours, and an analysis of support request tickets shows 35 unique TREE support requests across the year, the majority of which occurred at the launch of TREE rollout as teachers learned how to upload video. All were successfully resolved. There were no long-term technology barriers that created challenges to using TREE.

The TREE platform was also highly regarded by our coaches, who valued the ability to link feedback to specific moments in the video and look back on past video and feedback histories. A coach shared that TREE *"...allows me to prompt the teacher – look at minute x and this child – what do you notice?."* This pilot highlights the feasibility of partnership with teachers for the recording and sharing of video in their classroom using mobile devices for measuring fidelity and individualized professional development, as well as the feasibility for coaches who are reliably capturing fidelity of implementation and providing individualized support.

TOOLS OF THE MIND

Example Indicators for Share the News Activity	
Cycle 1 Indicator (September)	**Categories**
The teacher emphasizes "make sure everyone has a partner" on the Share the News mediator and "thinks out loud" while looking to ensure every child has a partner.	• **Building the Foundation**: Culture of peer scaffolding • **Building the Foundation**: Regulated classroom culture
Cycle 2 Indicator (January)	**Categories**
Children independently take responsibility for partnering and resolving partnership issues, or the teacher scaffolds children to support each other in ensuring everyone has a partner.	• **Evidence**: Culture of peer scaffolding • **Evidence**: Regulated classroom culture • **Tools Core Practices**: Increasing the challenge

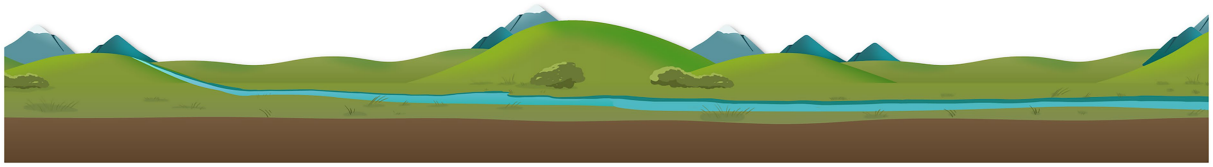


FIGURE 3
Tools Indicator example.

Cost-effectiveness and scalability

Staff time

In the context of measuring and supporting implementation fidelity in both research and professional development contexts, the demands of research and/or coach time are frequently a limiting factor in terms of scale. When assessing fidelity directly *via* in-person observation or in-person recording of video, logistical constraints such as travel time and classroom schedules impact the number of classrooms that can be observed per day. An important question for Tools was how utilizing teacher-recorded video would impact the ability to scale our reach and improve the efficiency of individualized professional development to maximize impact. In order to answer this question, Tools designed the TREE coach portal used by our team to review and respond to videos which allowed us to also internally track the time spent by team members on each video down to the second. Our findings suggest that the use of teacher-recorded video in TREE was both efficient and cost-effective compared to in-person observation. The median review time for submitted videos was 17 min (IQR = 20 min), which included reviewing the video, coding indicators present or absent, and communicating with teachers the time-stamped strengths they observed, and strategies they could apply to strengthen practice and support children's development of target skills.

Security and storage costs

Protecting children's privacy is fundamental to Tools of the Mind's work with teachers, schools, and school districts. To this end, we ensure that data storage and transmission are end-to-end encrypted and utilize industry-leading cloud storage providers who conform to relevant domestic and international security standards for data security such as SOC 1/SSAE 16/ISAE 3402, SOC 2, PCI DSS Level 1, ISO 27001, and FISMA. Additionally, we use industry best practices and align with national and international standards for data security and privacy governance including FERPA, COPPA, state privacy laws, and global regulations.

Recent advances in cloud computing have dramatically reduced the infrastructure and cost associated with uploading and storing large amounts of video data such as those collected in our pilot. During the 2021–2022 school year, Tools collected more than 190 h of video associated with our pilot for a total of 872GB of video file data. At standard industry cloud storage rates, the cost to store these data in secure encrypted online storage works out to approximately $20.07 USD per month, or less than $0.22 per video per year.

Travel cost savings

In the last school year in which in-person support was provided to school districts, Tools spent an average of $329.27 per classroom-served on travel costs. Extrapolating based on these figures, Tools saved nearly $100,000 in travel related expenses by using TREE to provide 2 captures of fidelity of implementation development and monthly sessions of individualized professional development support for teachers in each classroom of the 65 programs participating in the pilot. These savings allow programs like Tools to serve more teachers, and underscore how the use of teacher-recorded video can significantly improve scalability and keep interventions and curriculum affordable to districts wanting to implement them.

Discussion

Our TREE pilot demonstrates that teacher-recorded classroom video offers a feasible, cost-effective, and scalable method for assessing fidelity while enabling targeted, responsive individualized professional development to support teacher learning and implementation. While the specific design and implementation of TREE is uniquely adapted to the Tools of the Mind curriculum, we believe that our experience and learning offers several important takeaways for the conceptualization and implementation of fidelity measurement in SEL interventions more generally.

The approach laid out in this paper adds to the field by providing a new way of thinking about and operationalizing fidelity measurement while providing a cost-effective and scalable solution that delivers on tracking curriculum implementation fidelity. Below we discuss each of the five overarching ideas covered in this paper in more detail.

Conceptualizing fidelity of implementation as a developmental journey

Rather than relying on a set of indicators that are agnostic to where teachers are in the process of learning to teach Tools with fidelity, and where children's development is, we identified a set of "core" indicators (Century et al., 2010) across the year aligned with teacher and children's development. We believe other interventions could benefit from reconceptualizing fidelity of implementation in this more nuanced way.

At the same time, fidelity to a curriculum that embeds an approach to teaching and learning can only be assessed by seeing the impact of teachers' application of the approach on children at multiple developmental levels that change over time. Our experience has convinced us that integrating continuous, short, targeted fidelity snapshots with individualized professional development coaching support is a model that can both better capture fidelity of implementation than single measures and provide a way for intervention and curricula developers to support learning and implementation.

Measurement of fidelity and responsive support for teacher learning can and should be integrated

Our experience with TREE has strengthened our belief that fidelity of implementation and responsive support of teacher learning can and should be integrated. Tools, drawing on Vygotskian theory, works with teachers to empower them to observe children's current levels of development, adjust the challenge levels of activities in response, and uses scaffolding to meet each child where they are and support them in reaching their full potential. We believe teachers deserve and will benefit from the same responsive support. Traditional professional development approaches, such as workshops, are akin to 'whole group instruction' and cannot individualize to meet each teacher where they are and incorporate an understanding of the children in their classrooms and their current levels of development. In addition to not being able to individualize and provide on-going support that fits teachers' unique needs, these large-scale approaches to professional development often come with a high financial cost (Haymore-Sandholtz, 2002; Odden et al., 2002; Pianta et al., 2008). By integrating fidelity measurement with responsive and ongoing coaching support for teachers, TREE helps to minimize professional development and coaching costs while simultaneously measuring implementation fidelity for a larger group of teachers than in-person coaching allows.

Teachers benefit when fidelity of implementation measurement and tracking is an equitable partnership

With equity as a core value, through engaging teachers as partners in collecting and reflecting on data sampled multiple times across the year, we have prioritized transparency, collaboration, and inclusivity as key components of our approach.

In-person, observation-based fidelity measurement approaches limit the opportunity to engage teachers in self-reflection on implementation of a curriculum and its impact on children. Our experience has been that engaging teachers in capturing video and reflecting on the indicators observed in their video has real benefits for teachers, providing the opportunity to focus on children's interactions and development, and reflect on practice and the impact of their teaching on children over time (e.g., Borko et al., 2008; Marsh and Mitchell, 2014). Although Tools' commitment is to continuously collect fidelity of implementation data to measure our real-time impact and improve our support of teachers, we believe that researchers studying new interventions and curriculum would benefit from

this approach. Beyond improving fidelity of implementation, including teachers in this way adds value to them as participants in research, and may make teachers more open to participating in future research studies.

Finally, research has shown the impact of teachers' own SEL skills and wellbeing on students' SEL (Schonert-Reichl, 2017); for this reason alone, as we develop and research the impact of interventions and curricula designed to impact children's SEL, we should at the same time go about this in a way that promotes teacher learning and wellbeing. We believe that the process of having teachers partner in reflecting on video of their classroom to look for children's SEL and other areas of children's development empowers teachers, propels their learning, and contributes to their wellbeing.

The collection of teacher-recorded video in classrooms is feasible, scalable and offers significant long-term benefits for learning and research

Learning and research benefits

Video provides the opportunity for collaborative review and consensus coding, training and building inter-rater reliability as well as coaching skills to support each teacher and the children in their classroom (van der Linden et al., 2022). Video enables fast-cycle trials to evaluate the impact of professional development and individualized coaching, supporting innovation within the coaching staff, and leverages the potential to get rapid feedback from a video within a short period of time to observe if there is an improvement in teachers' application of teaching practices and fidelity of implementation. Moreover, video data can be re-analyzed in the future as fidelity measures are revised, providing an on-going source of valuable data (e.g., Clarke and Hollingsworth, 2002; van der Linden et al., 2022).

Feasibility and scalability

The role of technology both inside and outside of the classroom has changed dramatically over the past decade. As of April 2021, 85% of US adults now own a smartphone (Horowitz and Graf, 2021). Our experience suggests that many early childhood teachers (including those new to teaching and those with decades of experience) are comfortable filming in their classrooms with smartphones. Smartphone video and audio quality is now more than sufficient for assessment of fidelity in a classroom setting, and our team encountered few barriers in being able to see and hear in the over 1,000 classroom videos recorded by teachers.

We also found that teachers were quick to learn how to use the TREE mobile app and platform. A brief instructional video for teachers coupled with a few slides presented in professional development workshops was sufficient to enable the vast majority of teachers in our pilot to download our app and successfully upload video with no additional support. Teachers' only guidance on what to record each cycle consisted of the video focus (activity or time block) and the 5 indicators which were integrated into the TREE portal and phone app. We found these were sufficient for teachers to capture and upload video that enabled shared reflection on and assessment of fidelity of Tools implementation in classrooms over time.

Likewise, the vast majority of school districts and administrators expressed no concerns about having teachers film in classrooms for purposes of receiving individualized on-going professional development support. Administrators in TREE were also provided with continuous high-level data capturing implementation fidelity across classrooms, providing easy access to see how implementation was progressing to provide individual teachers with support as needed.

Finally, our experience suggests that switching from in-person observation to teacher-recorded video can offer significant gains in efficiency, allowing for greater reach to more teachers with the same resources with reduced overhead. Our data suggests it is feasible for a single team member to review, code, and respond to 20–25 submitted videos per workday, a four to five-fold increase over the number of teachers Tools team members could typically observe and provide feedback to during in-person visits to schools.

Acknowledgment of constraints

By design, this was a small pilot study focused on feasibility, and we cannot yet draw broad conclusions without further research. Future work will seek to assess the fidelity of implementation at a larger scale, identify the indicators correlated with school measures of child outcomes and quantify the impact of this new approach on teacher practice or child outcomes. Likewise, demonstrating the psychometric properties of our fidelity measure was outside the scope of this initial pilot, although it will be a focus of future work. Finally, because participating districts chose to participate and engage with Tools of the Mind, our results may not be representative of all teachers and school districts around the country. This was feasible in the population we described but may not be feasible in all settings.

While TREE was developed in the context of a specific curriculum that includes ongoing individualized professional development, we believe that this approach could be applicable to interventions and curricula that operate in other ways. A variety of interventions or curricula could engage teachers in capturing and uploading video and self-reflection as part of research efforts to look at impact and feasibility. Teacher-recorded video can be shared and uploaded on multiple commercially-available platforms and *via* multiple systems, fidelity of implementation indicators can be made transparent to teachers participating in research, and our experience suggests that such an approach can offer significant value.

Conclusion

We believe that our experience with TREE offers an important new approach to conceptualizing and operationalizing fidelity measurement, interweaving individualized support for teachers with real-time measures of program implementation fidelity *via* teacher-recorded video. The design of TREE as an equitable feedback system between teachers and curricular coaches empowers teachers to reflect on their own practices and children's development and identify shared focuses in a transformative and respectful way. Capturing fidelity over time, as opposed to once or twice a year, while also providing support and uplifting teachers in their practice is an approach, we believe could be transformative for other programs and researchers, especially in SEL interventions and curricula looking to support teachers and to examine implementation fidelity and its impact.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: private. Requests to access these datasets should be directed to mspacciapoli@toolsofthemind.org.

Ethics statement

The studies involving human participants were reviewed and approved by Advarra. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

MS co-conceived, wrote, and edited manuscript, oversaw collection of survey data, and contributed to fidelity of implementation indicators. MV wrote and edited manuscript, contributed figures, and contributed to the design of TREE. OS wrote and edited manuscript, co-designed and developed TREE, conducted data analysis, co-conceived of manuscript concept, and contributed figures. JS wrote and edited manuscript and conducted data analysis. TM wrote and edited manuscript. BW-S co-conceived, wrote and edited manuscript, co-designed and developed TREE, and supervised manuscript preparation. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Barnett, W. S., Jung, K., Yarosz, D. J., Thomas, J., Hornbeck, A., Stechuk, R., et al. (2008). Educational effects of the tools of the mind curriculum: a randomized trial. *Early Child. Res. Q.* 23, 299–313. doi: 10.1016/j.ecresq.2008.03.001

Barton, E. A., Whittaker, J. V., Kinzie, M. B., DeCoster, J., and Funari, E. (2017). Understanding the relationship between teachers' use of online demonstration videos and fidelity of implementation in MyTeachingPartner-math/science. *Teach. Teach. Educ.* 67, 189–201. doi: 10.1016/j.tate.2017.06.011

Bickman, L., Brown, J. L., Flay, B. R., Li, K., Pelham, W. Jr., Massetti, G., et al. (2009). Approaches to measuring implementation fidelity in school-based program evaluations. *J. Res. Character Educ.* 7, 75–101.

Blair, C., and Raver, C. C. (2014). Closing the achievement gap through modification of neurocognitive and neuroendocrine function: results from a cluster randomized controlled trial of an innovative approach to the education of children in kindergarten. *PLoS One* 9:e112393. doi: 10.1371/journal.pone.0112393

Borko, H., Jacobs, J., Eiteljorb, E., and Pittman, M. E. (2008). Video as a tool for fostering productive discussions in mathematics professional development. *Teach. Teach. Educ.* 24, 417–436. doi: 10.1016/j.tate.2006.11.012

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., and Balain, S. (2007). A conceptual framework for implementation fidelity. *Implement Sci.* 2, 1–7. doi: 10.1186/1748-5908-2-40

Century, J., Rudnick, M., and Freeman, C. (2010). A framework for measuring fidelity of implementation: a foundation for shared language and accumulation of knowledge. *Am. J. Eval.* 31, 199–218. doi: 10.1177/1098214010366173

Clarke, D., and Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teach. Teach. Educ.* 18, 947–967. doi: 10.1016/S0742-051X(02)00053-7

Collaborative for Academic, Social, and Emotional Learning (CASEL) (2022). Review programs. https://pg.casel.org/tools-of-the-mind/ (Accessed August 5, 2022).

Dede, C., Jass Ketelhut, D., Whitehouse, P., Breit, L., and McCloskey, E. M. (2009). A research agenda for online teacher professional development. *J. Teach. Educ.* 60, 8–19. doi: 10.1177/0022487108327554

de Leeuw, R. R., Boer, A. A., and Minnaert, A. E. M. G. (2020). The proof of the intervention is in the implementation; a systematic review about implementation fidelity of classroom-based interventions facilitating social participation of students with social-emotional problems or behavioural difficulties. *Int. J. Educational Res. Open* 1:100002. doi: 10.1016/j.ijedro.2020.100002

Diamond, A., Barnett, S. W., Thomas, J., and Munro, S. (2007). Preschool program improves cognitive control. *Science* 318, 1387–1388. doi: 10.1126/science.1151148

Diamond, A., Lee, C., Senften, P., Lam, A., and Abbott, D. (2019). Randomized control trial of tools of the mind: marked benefits to kindergarten children and their teachers. *PLoS One* 14:e0222447. doi: 10.1371/journal.pone.0222447

Durlak, J. A., and DuPre, E. P. (2008). Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am. J. Community Psychol.* 41, 327–350. doi: 10.1007/s10464-008-9165-0

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

Dusenbury, L., Brannigan , R., Falco, M., and Hansen, W. B. (2003). A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Educ. Res.* 18, 237–256. doi: 10.1093/her/18.2.237

Hamre, B. K., Justice, L. M., Pianta, R. C., Kilday, C., Sweeney, B., Downer, J. T., et al. (2010). Implementation fidelity of MyTeachingPartner literacy and language activities: association with preschoolers' language and literacy growth. *Early Child. Res. Q.* 25, 329–347. doi: 10.1016/j.ecresq.2009.07.002

Haymore-Sandholtz, J. (2002). Inservice training or professional development: contrasting opportunities in a school/university partnership. *Teach. Teach. Educ.* 18, 815–830. doi: 10.1016/S0742-051X(02)00045-8

Horowitz, J. M., and Graf, N. (2021). Mobile Fact Sheet. Pew Research Center. Available at: https://www.pewresearch.org/internet/fact-sheet/mobile/

Huntley, M. A. (2009). Measuring curriculum implementation. *J. Res. Math. Educ.* 40, 355–362. doi: 10.5951/jresematheduc.40.4.0355

Jones, D. E., Greenberg, M., and Crowley, M. (2015). Early social emotional functioning and public health: the relationship between kindergarten social competence and future wellness. *J. Public Health* 105, 2283–2290. doi: 10.2105/AJPH.2015.302630

Landry, S. H., Swank, P. R., Smith, K. E., Assel, M. A., and Gunnewig, S. B. (2006). Enhancing early literacy skills for preschool children: bringing a professional development model to scale. *J. Learn. Disabil.* 39, 306–324. doi: 10.1177/00222194060390040501

Marsh, B., and Mitchell, N. (2014). The role of video in teacher professional development. *Teach. Dev.* 18, 403–417. doi: 10.1080/13664530.2014.938106

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., et al. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proc. Natl. Acad. Sci. U. S. A.* 108, 2693–2698. doi: 10.1073/pnas.1010076108

Mowbray, C. T., Holter, M. C., Teague, G. B., and Bybee, D. (2003). Fidelity criteria: development, measurement, and validation. *Am. J. Eval.* 24, 315–340. doi: 10.1177/109821400302400303

Odden, A., Archibald, S., Fermanich, M., and Gallagher, H. A. (2002). A cost framework for professional development. *J. Educ. Finance* 28, 51–74.

Pianta, R. C. (2005). "Standardized observation and professional development: a focus on individualized implementation and practices," in *Critical issues in early childhood professional development*. eds. M. Zaslow and I. Martinez-Beck (Baltimore, MD: Paul H. Brookes Publishing), 231–254.

Pianta, R. C., Mashburn, A. J., Downer, J. T., Hamre, B. K., and Justice, L. M. (2008). Effects of web-mediated professional development resources on teacher-child interactions in pre-kindergarten classrooms. *Early Child. Res. Q.* 23, 431–451. doi: 10.1016/j.ecresq.2008.02.001

Quinn, D. M., and Kim, J. S. (2017). Scaffolding fidelity and adaptation in educational program implementation: experimental evidence from a literacy intervention. *Am. Educ. Res. J.* 54, 1187–1220. doi: 10.3102/0002831217717692

Schonert-Reichl, K. A. (2017). Social and emotional learning and teachers. *Futur. Child.* 27, 137–155. doi: 10.1353/foc.2017.0007

Shernoff, E. S., Lekwa, A. J., Reddy, L. A., and Coccaro, C. (2017). Examining teachers' attitudes and experiences with coaching to inform research-based practice: an iterative developmental design study. *J. Educ. Psychol. Consult.* 27, 459–485. doi: 10.1080/10474412.2016.1255850

Sklad, M., Diekstra, R., De Ritter, M., Ben, J., and 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. doi: 10.1002/pits.21641

Stains, M., and Vickrey, T. (2017). Fidelity of implementation: an overlooked yet critical construct to establish effectiveness of evidence-based instructional practices. CBE–life sciences. *Education* 16, 1–11. doi: 10.1187/cbe.16-03-0113

Taylor, R., Oberle, E., Durlak, J. A., and Weissberg, R. P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: a meta-analysis of follow-up effects. *Child Dev.* 88, 1156–1171. doi: 10.1111/cdev.12864

van der Linden, S., van der Meij, J., and McKenney, S. (2022). Teacher video coaching, from design features to student impacts: a systematic literature review. *Rev. Educ. Res.* 92, 114–165. doi: 10.3102/00346543211046984

Zhang, M., Lundenberg, M., Koehler, M., and Eberhardt, J. (2011). Understanding affordances and challenges of three types of video for professional development. *Teacher Teacher Educ.* 27, 454–462. doi: 10.1016/j.tate.2010.09.015

frontiers | Frontiers in Education

TYPE Hypothesis and Theory
PUBLISHED 20 December 2022
DOI 10.3389/feduc.2022.1035730

OPEN ACCESS

EDITED BY
Stephanie M. Jones,
Harvard University,
United States

REVIEWED BY
Ann Partee,
University of Virginia,
United States
Baiba Martinsone,
University of Latvia,
Latvia
Elena Bodrova,
Tools of the Mind Inc., United States

*CORRESPONDENCE
Jaiya R. Choles
jcholes@pdx.edu

SPECIALTY SECTION
This article was submitted to Educational Psychology, a section of the journal Frontiers in Education

RECEIVED 03 September 2022
ACCEPTED 28 November 2022
PUBLISHED 20 December 2022

CITATION
Choles JR, Roeser RW and Mashburn AJ (2022) Extensions beyond program impacts: Conceptual and methodological considerations in studying the implementation of a preschool social emotional learning program. *Front. Educ.* 7:1035730. doi: 10.3389/feduc.2022.1035730

Extensions beyond program impacts: Conceptual and methodological considerations in studying the implementation of a preschool social emotional learning program

Jaiya R. Choles[1]*, Robert W. Roeser[2] and Andrew J. Mashburn[1]

[1]Department of Psychology, Portland State University, Portland, OR, United States, [2]Department of Human Development and Family Studies, Pennsylvania State University, State College, PA, United States

Social–emotional learning (SEL) programs are frequently evaluated using randomized controlled trial (RCT) methodology as a means to assess program impacts. What is often missing in RCT studies is a robust parallel investigation of the multi-level implementation of the program. The field of implementation science bridges the gap between the RCT framework and understanding program impacts through the systematic data collection of program implementation components (e.g., adherence, quality, responsiveness). Data collected for these purposes can be used to answer questions regarding program impacts that matter to policy makers and practitioners in the field (e.g., Will the program work in practice? Under what conditions? For whom and why?). As such, the primary goal of this paper is to highlight the importance of studying implementation in the context of education RCTs, by sharing one example of a conceptualization and related set of implementation measures we created for a current study of ours testing the impacts of a SEL program for preschool children. Specifically, we describe the process we used to develop an implementation conceptual framework that highlights the importance of studying implementation at two levels: (1) the program implementation supports for teachers, and (2) teacher implementation of the curriculum in the classroom with students. We then discuss how we can use such multi-level implementation data to extend our understanding of program impacts to answer questions such as: "Why did the program work (or not work) to produce impacts?"; "What are the core components of the program?"; and "How can we improve the program in future implementations?"

KEYWORDS

fidelity of implementation, social emotional learning, preschool, education program, intervention

Introduction

The recent wide-scale expansion of social emotional learning (SEL) programs in schools and classrooms has been informed by results of research studies that demonstrate, across age groups, SEL programs have positive impacts on students' academic success and well-being. More specifically, SEL programs have been found to produce demonstrably positive impacts on students' social and emotional skills (e.g., perspective taking, identifying emotions, interpersonal problem solving); attitudes toward self and others (e.g., self-esteem, self-efficacy); positive social behaviors (e.g., collaboration, cooperation); reduced conduct problems (e.g., class disruption, aggression); reduced emotional distress (e.g., depression and anxiety); and academic performance (e.g., standardized math and reading; Durlak et al., 2011). Positive impacts of SEL programs are also evident among preschool aged children; results from a meta-analysis of 39 SEL programs in early childhood education settings found small to medium effects (Hedge's g effect size estimates between 0.31 and 0.42) for improvements in children's social and emotional competencies, and reductions in their challenging behaviors (Luo et al., 2022). Results of studies of SEL program impacts on children's social and academic outcomes have gone on to be included in economic studies, and a recent estimate of a $11 return for every $1 invested in school-based SEL programs have compelled policy makers and program administrators nationwide to implement these programs at a wide scale (Belfield et al., 2015).

To date, studies of the impacts of SEL programs have mostly used field-based experimental designs (Boruch et al., 2002; Bickman and Reich, 2015). When applied to education contexts, these experimentally designed studies are commonly referred to as randomized controlled trials, or cluster randomized trials in the cases when the design considers the multiple levels of analyses that are familiar in school-based settings (e.g., classrooms clustered within schools, and children clustered within classrooms). The primary research question that is addressed by the experimental design is "What are the impacts of offering access to an SEL program on students' development of SEL competencies and well-being?." The experimental study then proceeds by randomly assigning schools, classrooms, or children (depending on whether the program is intended to be delivered school-wide, classroom-wide or to individual children) to either the "treatment" group that is offered the SEL program or to a comparison group that either may receive a different program (active control) or carry on with business as usual (control) and may receive the treatment at a later time (waitlist control).

The methodological premise of random assignment is that each school, classroom, or child has an equal chance of being assigned to the SEL program or the comparison group, and as such, there are no expected differences between these groups at the start of the study on any measurable or unmeasurable characteristics. Thus, any differences in outcomes between these two groups after the program concludes can be attributed to the one key difference between these groups—the SEL program itself (e.g., Shadish et al., 2002). Given these methodological strengths, experimentally designed studies, when well implemented, may be more likely to achieve a label of "evidence-based," which is weighted heavily by policy makers and program administrators when considering which programs should be funded for expansion and wide scale use (Boruch, 2005; Mark and Lenz-Watson, 2011).

Despite these important strengths of the experimental design with regard to internal validity of the conclusions about program impacts, there are some notable weaknesses of the experimental design. Such weaknesses are related to the difficulty implementing field-based experimental studies in real world settings. For instance, the external validity of the results of an impact study—to whom the results of the study generalize—often is not clear. In addition, in the experimental study, the program is implemented under "ideal conditions" which may not reflect the actual conditions in which the program is implemented in the real world (e.g., more resources and/or implementation supports). And, despite using random assignment, there may be differences between the treatment and comparison groups at the start of the study, which would weaken the inference that the intervention caused the impacts.

Perhaps most notably, results of experimental studies leave a lot to be desired for researchers and program implementers who ask different questions about SEL programs that extend beyond the question about program impacts that the experimental study is primarily designed to address. This limitation of the experimental design in field-based settings (such as schools) becomes evident in studies where group assignment (treatment or comparison) does not reflect the experiences of those assigned to the condition. That is, random assignment determines that a school, classroom, or child was offered *access to the SEL program*. However, not all schools, classrooms or children that are offered access to the program will actually *participate in the program* at all, or in the ways that the program developers intended. In addition, those schools, classrooms, or children assigned to the comparison group may actually implement components of the SEL program that are intended to be accessed only by the treatment group. As such, the results of an experimental study provide estimated impacts of random assignment to the SEL program without regard to the actual experiences of those who participated in either study condition (Hollis and Campbell, 1999). Thus, this methodology ignores the often rich and meaningful variation in how the program was implemented and what the experiences were of those who participated in the treatment and comparison conditions, and how closely aligned these experiences were to how the program was intended to be implemented by those who developed it.

In sum, impact studies of SEL programs answer important policy questions, but they are limited in their capacity to answer questions that are relevant for education practitioners and non-policy-oriented research. We argue here, as others have (Fixsen and Blase, 2009; Moir, 2018) that studying the implementation of a program in the context of an impact study

creates an ideal context to understand program impacts and subsequently address many other important questions of interest to program developers and practitioners who directly implement these programs.

The field of implementation science offers rich frameworks for researchers to draw on that examine how variation in how the program was implemented is associated with the program's impacts. Carefully designed implementation studies can provide critical contextual information that helps researchers feel confident in answering the questions about impacts thoroughly, and they offer the opportunity to extend research questions to explore active ingredients in the program, explain why an intervention was not effective, and guide efforts to modify interventions that maximize their effectiveness in the future (Durlak and DuPre, 2008). For example, Low et al. (2016) used latent class analysis to study teachers' implementation within a RCT study of the Second Step® SEL curriculum in elementary schools. The authors incorporated multiple implementation measures into their latent structure specification, including adherence, dosage, generalization (i.e., application/integration in the classroom), and student engagement. The authors then used the determined latent structure to predict both teacher report and observed scores of student behavioral and academic outcomes and found a negative relationship between the low engagement latent class and student outcomes, when compared to the low adherence and high quality implementation latent classes.

Other studies have examined moderation of program impacts with factors that may influence implementation. For example, McCormick et al. (2016) examined the role of parent participation in moderating program impacts of the INSIGHTS into Children's Temperament program in the subset of kindergarten and first grade participants of a larger experimental study. Results indicated program impacts for students (reading, math, and adaptive behaviors) were stronger for children of parents categorized in the low participation group. Similarly, Sandilos et al. (2022) explored whether implementing the Social Skills Improvement System SEL Classwide Intervention Program buffered the negative effect of low teacher well-being on the quality of teacher-child interactions in the classroom.

Despite this small but growing body of evidence in SEL programs, implementation is still often not studied in the context of testing impacts of SEL programs (Domitrovich et al., 2012) and mindfulness-based SEL programs (Roeser et al., 2020). Furthermore, these studies of implementation tend not to address the "complex" structure of the intervention design in education settings. That is, the focus of program implementation of SEL programs has been on the teachers (as intervention agents), rather than also considering whether the implementation of the training and supports offered to teachers (by implementation agents) have been effective (Fixsen et al., 2005); and, little consideration is given to the multitude of coordinating pieces required to implement this multi-level structure across diverse settings (Bryk, 2016). Additionally, with the increase in popularity of SEL curricula, monitoring control group practices (i.e., describing business-as-usual practices) has become increasingly important.

In the following sections, we provide a conceptual framework for conducting an implementation study within the context of an experimental study of program impacts. We then apply this framework to a mindfulness-based SEL program for preschoolers and describe our process of creating a robust set of measures of program implementation that we include in our impact study. We then demonstrate how this framework and measures can be used to address interesting and important questions about the SEL program that extend beyond the question of program impacts, such as, "Why did the program work (or not work) to produce impacts?"; "What are the core components of the program?"; and "How can we improve the program in future implementations?." This paper will explore these questions using examples of our measures and offer suggestions to others implementing studies like ours.

Implementation study design and measurement framework

In this section, we highlight the importance of attending to the interconnections between an impact study and a study of program implementation. To do this, we provide an example of our own work assessing program implementation of a preschool mindfulness-based SEL program. We begin by briefly describing the SEL program and the design of the study testing the impacts of the program on children's outcomes. Then, we provide an overview of specific aspects of fidelity of implementation (FOI) that are commonly considered in implementation studies. Lastly, we present the key components of program implementation that are the focus of our implementation study and describe how they are incorporated into our multi-level FOI conceptual framework. In this last section, we also discuss the process of mapping measures onto our conceptual framework, piloting those measures, refining them, and integrating them into the design of our impact study.

Social emotional learning program description

The mindfulness-based SEL program for preschoolers under investigation in this study (MindUP™ PreK—The Goldie Hawn Foundation) consists of four main elements: the mindfulness-based SEL preschool classroom curriculum for students, the curriculum training for teachers, and two additional implementation supports for teachers—monthly community of practice meetings and coaching sessions. A cluster randomized trial testing the impacts of the MindUP™ program on children's social, emotional, and academic outcomes is the context for this implementation study. Specifically, the theory of change of this study hypothesizes that through implementation of the MindUP™

program, preschool children will develop key social and emotional skills (i.e., attentional, social, and emotional) and academic skills (i.e., early literacy and math). We also hypothesize that the impact of MindUP™ may be stronger for children who begin preschool with fewer SEL skills and/or who are in classrooms where students experience lower quality interactions with their teachers. Exploratory follow-up to these impacts will also examine which aspects of implementation of the program are positively associated with children's development of SEL and academic outcomes and whether aspects of FOI vary based on teacher and classroom characteristics measured at baseline.

The impact study consists of three sequential and independent cohorts of preschool classrooms that will be randomly assigned to either participate in the MindUP™ program or in a waitlist control group for one year. We began the MindUP™ trial in fall 2019 and successfully recruited our first cohort of 38 teachers from a range of preschool programs (e.g., private for profit, community-based organizations, Head Start) serving four-year-old children. These teachers were randomized and half were offered access to the MindUP™ training program and supports. However, the trial was interrupted in Spring of 2020 due to the COVID-19 pandemic and was then paused for two subsequent years. In fall of 2022, the trial was re-started with a new first cohort of teachers and minor adaptations to the implementation of the program (i.e., remote training rather than face-to-face training) and the research protocol (i.e., reduced in-person assessments for children).

Mindfulness-based social emotional learning curriculum

The MindUP™ curriculum for preschool students consists of 15 mindfulness and SEL-based lessons, and a daily core practice called the "brain break." Each weekly lesson comprises several related activities that are estimated to take 15 minutes to complete. For this study, teachers were instructed to implement two activities within each lesson per week in their classroom. Additionally, the curriculum included supplementary activities and instructions on ways to integrate the lessons into other classroom experiences. The 15 curriculum lessons are organized into three units. The first unit is *Mindful Me*, in which children learn about the structure and function of the brain and are introduced to the brain break. An example of an activity from this unit is *My Feelings,* in which children build emotion literacy skills by learning to name different emotions and identify how they feel when they experience them. The second unit is *Mindful Senses*, in which children focus on the relationship between their senses, their bodies, and how they think. An example activity from the second unit is *Mindful Touch*, in which children are instructed to be open and curious about touching mystery objects hidden in closed containers. Removing sight from the activity allows children to investigate the objects using only their tactile sense. The third unit is *Mindful Me in the World*, where children learn about mindsets, such as gratitude and perspective-taking, and how to apply mindful behaviors through interactions with the community and world. An example activity from unit three is *The Gratitude Tree*, in which students are instructed to draw something or someone they are grateful for on a paper shaped like a leaf and then the children's work is displayed on a tree visual in the classroom.

The core practice of the MindUP™ curriculum is called the "brain break"—a brief focused attention activity that teachers implement in their classroom with their students four times daily—usually at the start of the day, after recess, after lunch, and at the end of the day. The brain break is initiated with a chime sound that children focus on and listen to in order to settle into their bodies. Children are then instructed to focus on the natural rhythm of their breath. For preschoolers, the brain break initially requires various scaffolds to help focus attention, which can include, for example, a "breathing ball" (e.g., Hoberman Sphere) that expands and contracts to simulate the inhale and exhale of breathing; or the placement of stuffed animals on the diaphragm to help children focus on the rising and falling of their breath. After repeating these mindful breaths two to three more times, the chime is rung a second time to conclude the brain break. Children are instructed to listen for as long as they can hear the chime and then provided time to bring their awareness back to the classroom. Conceptually, the brain break can be considered a focused attention practice (Maloney et al., 2016).

MindUP™ curriculum training

Before implementing the MindUP™ curriculum in the classroom, teachers attend a single day six-hour curriculum training led by a certified trainer (in this study, so named the implementation director). During the training, teachers learn foundational scientific research underpinning the curriculum, review the curriculum book, discuss implementation with other attendees, and practice lesson planning. During this training, teachers are also given a comprehensive materials kit to fully implement the curriculum in their classroom.

Community of practice

Based on our previous research on MindUP™ in the early years (Braun et al., 2018), we developed a new set of implementation supports for the purpose of this study. This included a community of practice for all teachers implementing the MindUP™ curriculum in their classrooms (see Mac Donald and Shirley, 2009). The community of practice component consists of monthly hour-long, face-to-face, small group meetings between participating teachers and the implementation director. At the start of each meeting, teachers are led through a mindfulness activity as a way to center the group and to provide teachers with their own opportunities to focus their attention. These meetings are facilitated by the implementation director and are used as time for teachers to discuss their progress in implementing the curriculum. More specifically, teachers are asked to reflect and share with the group regarding what went well or what was challenging implementing the curriculum activities since the last community of practice meeting. At this time, the implementation director, and other teachers, can offer feedback to support implementation improvement in the future. During these

meetings teachers are also provided with the opportunity to discuss their plan for upcoming lessons and to discuss as a group the purpose of the upcoming curriculum activities. During this segment of the meeting, teachers can also ask the implementation director for support in how to successfully implement the upcoming activities, or to address any aspects of the upcoming implementation that remain unclear.

Coaching

In addition to the community of practice, we also included a coaching program as an additional implementation support. Developed by our implementation director, the coaching program comprises monthly 30-min one-on-one check-in calls between the implementation director and each teacher participating in the MindUP™ program. The coaching model used was adapted from the National Center on Quality Teaching and Learning's *Practice Based Coaching* framework (for example see Snyder et al., 2015) to specifically align with the MindUP™ curriculum. Prior to each check-in call, the implementation director would review the teacher's most recent implementation log information to understand the teacher's implementation progress. During the coaching sessions, the implementation director and each teacher would discuss specific challenges or questions related to their curriculum implementation and brainstorm strategies for improving implementation in the classroom. The implementation director's agenda for each of these coaching sessions drew on the *Practice Based Coaching* framework and included time for discussing shared goals and action planning, teacher self-monitoring, and reflection and feedback related to the program.

Implementation of the MindUP™ program

In this section, we define components of FOI and apply them to the MindUP™ SEL program in particular, to create a conceptual framework for assessing fidelity in our study of the implementation of MindUP™. In our effort to develop a conceptual and assessment strategy for studying FOI in the context of this program, we drew on extant work on FOI. Broadly, implementation can be defined as the study of a program and its components, and how it is delivered in a specific context to optimize program outcomes (Durlak and DuPre, 2008). Variation exists among aspects of FOI terminology, however, common components often include the following: (1) *Dosage*, which describes the strength or quantity (in hours, sessions, etc.) of the program; (2) *Responsiveness*, which consists of the extent to which the program is engaging, interesting, and relevant to participants; (3) *Adherence*, which measures the extent to which the program is implemented as designed or planned; (4) *Quality* describes how well program components are delivered (e.g., clarity, organization); and, (5) *Differentiation* examines the degree to which the program under investigation is similar or different to others like it (Dane and Schneider, 1998; Durlak and DuPre, 2008). More recently, the field of implementation science has identified additional FOI components that should be measured, including (6) Program *adaptations*, which capture modifications (changes, omissions, additions) made to program components; (7) *Program reach*, which describes the generalizability or representativeness of program participants to the broader population of interest for the program; and, more specifically to experimental studies, (8) *Monitoring control group practices*, which seeks to measure the extent to which intervention activities or "intervention-like" activities are conducted by the comparison group (Durlak and DuPre, 2008).

Together, measures of FOI components help to capture the multiple elements of implementation that factor into program success; for instance, *dosage* data that measures the frequency with which participants attended program sessions can be used to calculate the percentage of the total program each participant actually received (i.e., their individual "dose"). However, it is often insufficient to simply measure attendance—research shows participants learn when content is engaging, relevant, and interesting—so measuring *responsiveness* becomes imperative to contextualize dosage. From the implementation support side, *adherence* and *quality* FOI components are similarly interrelated: *adherence* measures can assess the degree to which the program was implemented as planned, but in order to evaluate whether the program was engaging and relevant (i.e., participant responsive), assessing if the program components were delivered with a high level of *quality* is needed as well. Finally, *adaptations* are important to measure due to their potential influence on each of the other components described above. Adaptations may impact implementation in a complex way, such that they could increase the quality of delivery and therefore participant responsiveness, while simultaneously reducing adherence to the program as prescribed. Moving outside of the group receiving the intervention, *monitoring control group practices* is also important in order to fully understand the impact of an intervention in impact studies. If this FOI component is not considered, the true impact of the program on outcomes cannot be determined if the experiences of the control group are not known (Fixsen et al., 2005). In sum, each of these FOI components affect the conclusions that can be drawn about the impacts of the SEL program (Durlak and DuPre, 2008).

Other FOI components—differentiation and program reach—concern the external validity of results of a program and are particularly important to evaluate when comparing across different "evidence-based" programs (Durlak and DuPre, 2008). For instance, *differentiation* data can help answer questions that seek to determine why a specific program should be chosen over another (what makes a program unique?). Clearly outlined information regarding dosage, program components, and resources needed, as well as their associated costs, can help define a program's uniqueness and help determine which may be best used for a particular purpose or community. Similarly, paying attention to *program reach* can help program adopters estimate the extent previous data regarding a program's effectiveness will

generalize when implemented in communities the program was developed for (Durlak and DuPre, 2008). Without this broader lens, researchers remain in the dark on the extent to which a program will be successful when implemented outside of a highly controlled "ideal conditions" scenario of an impacts study.

In education interventions, measuring each of these aspects of implementation becomes increasingly complex due to the fact that program components are diverse and span across multiple levels (e.g., classrooms, teacher professional development, whole schools). In addition, the coordination of these components within and across levels, in a manner that is responsive to a wide range of local contextual and organizational conditions, is key to successful program implementation. Due to this multi-level, ecological complexity surrounding program implementation, Bryk (2016) describes FOI in education contexts as "adaptive integration" to emphasize that as much as program implementation involves adherence and compliance, it also fundamentally involves responsivity and adaptation across levels. Thus, although we adopt the term FOI here, we use this terminology as a broader term that acknowledges and incorporates this systems perspective of "adaptive integration" into our work.

Below we present our implementation study conceptual model in Figure 1, which includes the FOI components we believe are most important to measure during the implementation study that is part of an impacts study (see Bywater, 2012). However, it is important to note that there are numerous other factors that must be considered and coordinated prior to this stage that also influence implementation success. For example, researchers must obtain buy-in from interested parties and ensure schools, or in this case preschool centers, are ready and receptive to implementation and change through program adoption (Fixsen et al., 2005; Moir, 2018). It is also important to ensure motivations are shared between leadership (often the gatekeepers for what programs are considered) and teachers (who often implement the program with children in classrooms; Fixsen et al., 2005). Intervention programs should also be piloted prior to testing their efficacy to determine if the program is accepted by participants and feasible, and to identify any areas that need to be improved for future use (Bywater, 2012; Bryk, 2016). Similarly, aspects of FOI focused on the external validity and generalizability of programs are connected to this work, but remain outside the focus of this paper.

In our implementation study, we believe it is very important to attend to the multi-level nature of the program under study (Fixsen et al., 2005). Specifically, the first level of the SEL program focuses on measuring the transfer of program knowledge from the project implementation team members (in our study, the implementation director) to participating teachers who will be implementing the curriculum in their classrooms with their students. As seen in our conceptual framework displayed in Figure 1, we call this level of measurement "teacher implementation supports." Specifically, this level comprises measuring FOI of the MindUP™ curriculum training, community of practice meetings and coaching sessions.

The second and more frequently considered level of implementation is what we refer to as "curriculum implementation"—teachers' implementation (as intervention agents) of the curriculum with students in the classroom. In our study, this level includes measurement of the program curriculum activities as well as the daily brain break practice. We also posit that it is important to monitor these two levels in the control

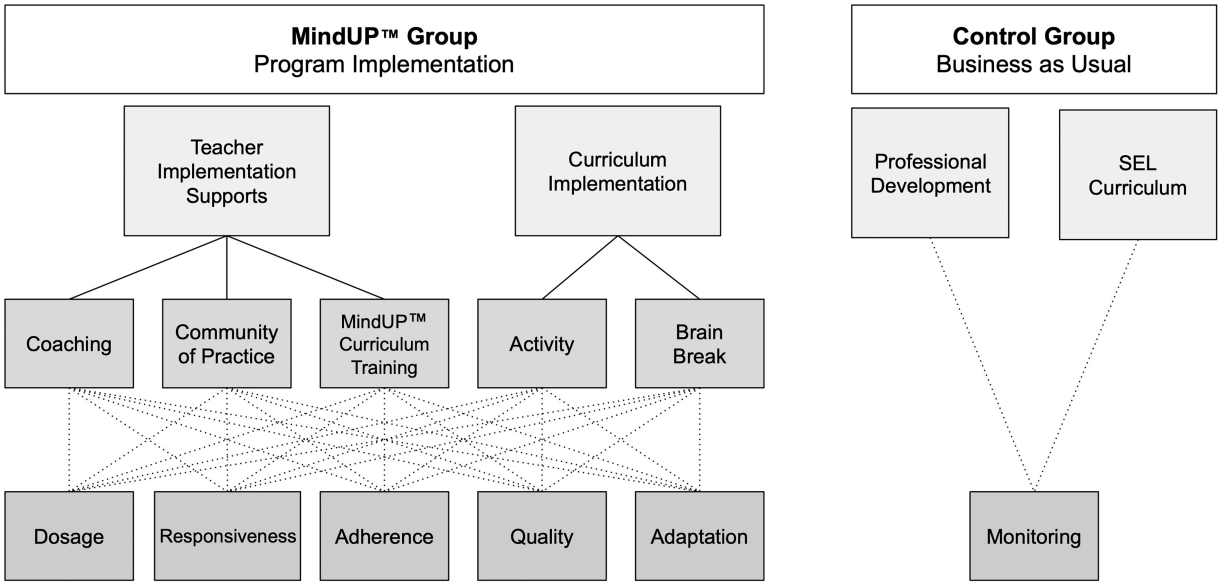


FIGURE 1
Conceptual framework for fidelity of implementation of MindUP™.

group as well. To operationalize similar activities that may take place as part of business-as-usual practices for the control group, we define these two levels more broadly as professional development for teachers and SEL activities for students for this group. The right side of Figure 1 visualizes this component of our framework and highlights the importance of attending to this aspect of FOI in the context of an impacts study.

The lower half of our conceptual model in Figure 1 identifies the comprehensive set of FOI components of the program we sought to measure in our implementation study. Specifically, we chose to measure dosage, responsiveness, adherence, quality, and adaptations for each of the teacher implementation supports (i.e., curriculum training, community of practice, and coaching) and each curriculum implementation component (curriculum activities and brain break). Similarly, we defined the FOI measurement at this level for the control group as "monitoring" of both control group teachers' professional development activities and SEL curricular activities with their students. In Table 1, we provide a glossary that defines all of these terms. Next, we describe in detail how we compiled and developed measures, and when we determined it essential to have measures from multiple informants.

TABLE 1 Definitions of terms.

Construct	Definition
Program	Synonymous with intervention
Fidelity of implementation (FOI)	General term used to define the systematic study of implementation, includes complex interventions that have been defined as "adaptive integrations"[1]
FOI component	Aspects of implementation that can be measured, such as dosage, responsiveness, adherence, quality, or adaptations
Implementation level	Label to differentiate implementation supports to teachers by an implementation agent (first level) and curriculum implementation by the teacher to students in the classroom (second level)
Program element	Factors that make up the program as a whole, including both implementation supports (curriculum training, community of practice, and coaching) and curriculum implementation (curriculum activities and brain break practice)
Implementation agent	A person who provides professional development to the teacher to support their implementation of the program with students in the classroom
Intervention agent	The teacher or other person who delivers the program in the classroom to students

[1]Bryk (2016).

Mapping measures onto the implementation conceptual framework

With a clear conceptualization of FOI in general, and of the MindUP™ program in particular, we then began to develop specific measures of FOI that map onto our conceptualization. We began this mapping process by stating two explicit goals for our set of FOI measures: first, the set of measures should be comprehensive, such that there is at least one measure for each FOI component outlined in the framework for each program element within each program level. Secondly, that when feasible, we believe there should be more than one measure for an FOI element, and that these multiple measures should be from different reporters or sources (i.e., multi-informant). Table 2 presents the results of this process by summarizing the informant or informants for each FOI component measured. Specifically, it highlights whether the informant was the implementation agent (i.e., implementation director), the classroom teacher, a researcher who is a third-party observer, or a combination therein. In our study, it did not seem developmentally appropriate to have preschoolers report on their experiences participating in the curriculum implementation, however, older students can be included as an additional informant source when applicable.

As mentioned above, in order to develop a robust FOI measure set, when feasible and appropriate, we sought triangulation of perspectives and experiences through multi-informant measures (see Table 2). Thus, our study design included third-person observers, with a member of our research group attending and observing the MindUP™ curriculum training and community of practice sessions. For these program elements among those in the "treatment" group, the observer comprehensively rated all FOI components defined in our conceptual framework (i.e., dosage, responsiveness, adherence, quality, and adaptation). We supplemented this complete set of observer ratings with teacher self-report ratings of those FOI components that require the first-person experience of the participant (i.e., responsiveness, quality). We did not feel it was appropriate for a third-party observer to rate the one-on-one coaching sessions, so instead, for this program element, we had the implementation director provide ratings that measure dosage, responsiveness, adherence, and adaptation ratings. Again, we supplemented these measures with responsiveness and quality ratings from the teacher.

For the curriculum implementation level, teachers were the primary source who reported on the complete set of FOI measures (except for ratings of quality) for their implementation in the classroom. This required teachers to fill out an implementation log on a weekly basis as they implemented each program activity. In

TABLE 2 Example of multi-informant measures of fidelity of implementation components across levels and groups.

	Implementation supports (implementor to teacher)				Curriculum implementation (teacher to student)		
FOI component	**Curriculum training**	**Community of practice**	**Coaching**	**Control monitoring (PD)**	**Curriculum activity**	**Curriculum brain break**	**Control monitoring (SEL)**
Dosage	OR	OR	IR	TR	OR; TR	OR; TR	TR
Responsiveness	OR; TR	OR; TR	IR; TR	TR	OR; TR	OR; TR	TR
Adherence	OR	OR	IR		OR; TR	OR; TR	
Quality	OR; TR	OR; TR	TR		OR	OR	
Adaptation	OR	OR	IR		OR; TR	OR; TR	

OR, Observer Report; TR, Teacher Report; IR, Implementor Report; PD, Professional development; SEL, Social emotional learning curriculum.

addition, teachers provided weekly data on the daily brain break practice as well. For triangulation of information at this level, researchers also comprehensively rated all FOI components measured in the "treatment" group, for both a curriculum activity and brain break practice, by visiting teachers in their classrooms on two separate observation occasions during the implementation phase. During these observation sessions, the researchers recorded the number of students who were present for the observation session (as a collective measure of dosage) for both the curriculum activity and the brain break. The observer then rated adherence, adaptations, student engagement, and total minutes spent conducting the activity (a specific measure of dosage) separately for the curriculum activity and brain break (see Figure 2 for an example of how these FOI components were scored by the observer during the Brain Break implementation). After observing both program components, the researcher scored the teacher's quality of delivery of both the curriculum activity and brain break (as a collective measure of quality).

Generally, across both levels of implementation data, the desire for a high degree of adherence to intervention elements needs to be balanced with flexibility and adaptation of the program to meet local school, community, and student needs. Known as the adherence/adaptation trade off, researchers must balance the desire for internal validity with an understanding that achieving high-quality implementation often requires the implementation agent and intervention agents to be afforded flexibility in meeting the needs of participants and contextual demands (Dane and Schneider, 1998). To contend with this issue, Fixsen et al. (2005) recommend requiring adherence to main intervention principles but allowing flexibility in how the principles are implemented (i.e., processes, strategies) in a manner that retains the objective or function of the component.

In our own work, we attempted to produce this balance when developing our measures of adherence and adaptation. First, the way the MindUP™ curriculum was developed helped us with this balance. The curriculum allows for flexibility in the implementation of the brain break practice, for instance, by including multiple scaffolded ways teachers can implement the brain break in the classroom. The manual also discusses the need for flexibility regarding the context in which the practice is implemented, such as the location/time of day (e.g., circle time) and structure (whole group or small group). This flexibility directly translates into the classroom implementation adherence measure we developed, where we capture teachers' use of different scaffolds when offering the brain break (e.g., Hoberman sphere, stuffed animals; see the "extra materials utilized" section of Figure 2) as well as the presence/absence of (what we believe are) core components and the extent to which the component's objective was met (see Jennings Brown et al., 2017; Doyle et al., 2019). For example, Figure 2's section on adherence depicts adherence expectations for teachers during the brain break practice, in which each element is considered essential for full implementation of this practice. Given teachers' diverse expertise and our desire to leave space for developmentally appropriate practice, we believe measuring adherence in multiple ways (e.g., presence/absence and degree objective was met) will capture more nuance and variability across adherence in classroom implementation. Additionally, we included space to qualitatively describe any adaptations to the brain break practice that move outside the scope of flexible options provided within the program.

To monitor control group practices, control group teachers were asked to describe their professional development and SEL curriculum activities as part of their spring (end of program) survey. In terms of FOI components, we focused on asking questions about program dose and responsiveness, and asked teachers to describe their experiences regarding curriculum trainings and their engagement in communities of practice, and/or coaching. Similarly, we asked them whether or not they implemented any SEL curricula and/or attentional practices with their students, and explicitly asked them whether they implemented any portion of the MindUP™ curriculum over the course of the year.

In sum, Table 3 provides a complete list of the comprehensive set of FOI measures developed for and used in this study. Table 4 provides example items drawn from these measures for each of the FOI components measured. Developing this set of measures was one of the main activities during the first year of the project, which was a planning year for the project intended for this purpose. After initial development, we conducted a small-scale pilot study

Brain Break
Indicate below if the brain break was implemented, the level of adherence, any extra materials utilized. Adaptations, and student engagement

Observed: ☐ Yes ☐ No Start Time: _____:_____ End Time: _____:_____

Were each of the following components included?

☐ Children are invited to participate

☐ Teacher tells children to find comfortable position

☐ Beginning Chime ☐ Three breaths ☐ Ending Chime

☐ Time to bring attention back to classroom (30-45 seconds)

Extra materials utilized:

☐ Flower/candle metaphor ☐ Stuffies/duckies ☐ Hoberman sphere ☐ Pinwheels

☐ Other: _____

Notes:

Brain Break Adherence: How closely was the brain break implemented as prescribed? Note: use in conjunction with Activity Adherence Guide

☐ Very high: followed the directions as prescribed (80%-100%)

☐ Moderately high: Followed most of the activity directions (60%-80%)

☐ Average: Followed about half of the activity directions (40%-60%)

☐ Moderately low: Followed less than half of the activity directions (20%-40%)

☐ Very low: did not follow most of the activity directions (<20%)

Was the Brain Break Objective Met? 0 1 2 3 4

Brain Break Adaptations: e.g., were there any practices added from other SEL programs? Any other information added? Anything left out that should have been included? Any disruptions? Any accommodations made for diverse learners?

Student Engagement
For the questions below, please rate student engagement during each implementation component - from 0 (lowest) to 4 (highest)

0 = Very low 1 = Low 2 = Average 3 = High 4 = Very high

Brain Break NA 0 1 2 3 4

Number of students who didn't accept invitation to participate: _____

Alternative activity offered: N/Y _____

FIGURE 2
Example of third-party observer ratings of the brain break practice: adherence, adaptations, and student responsiveness. Dosage was additionally measured by student attendance during the observation session. Quality of the teacher's implementation was rated once for the observation session, such that quality scores comprise both the curriculum activity and brain break. Neither of these measures are represented in this figure.

to test our set of FOI measures, which led to a cycle of revisions of the measures in preparation for their use within the context of the MindUP™ impact study. Also of note, the process of development required close collaboration between the implementation director and researchers, particularly in operationalizing adherence measures of the teacher implementation supports. We also drew on expertise from a FOI consultant and other researchers supporting our project, who had extensive knowledge and many

TABLE 3 Fidelity of implementation measures of program elements by frequency and components measured.

Program element	Measure	Report frequency	FOI components measured
Curriculum training	MindUP™ training teacher survey	Once	Quality, responsiveness
	MindUP™ training observation form	Once	Dosage, adherence, quality, responsiveness, adaptations
Community of practice (CoP)	CoP teacher survey	Monthly	Quality, responsiveness
	CoP observation form	Monthly	Dosage, adherence, quality, responsiveness, adaptations
Coaching	Implementation director coaching log	Monthly	Dosage, adherence, responsiveness, adaptations
	CoP teacher survey	Monthly	Quality, responsiveness
Curriculum activity	Teacher implementation calendar	Weekly	Dosage, adherence, responsiveness, adaptations
	Classroom observation form	Twice	Dosage*, adherence, quality*, responsiveness, adaptations
Brain break	Teacher implementation calendar	Weekly	Dosage, adherence, quality, responsiveness, adaptations
	Classroom observation form	Twice	Dosage*, adherence, quality*, responsiveness, adaptations
Control group practices	Teacher survey	Once	Dosage, responsiveness, descriptions of professional development experiences

*Quality of teacher delivery and dosage (attendance of students) were each rated once for the observation session, such that quality and dosage scores comprise both the MindUP™ Curriculum Activity and Brain Break. Dosage and responsiveness were measured in the control group for both implementation supports and curriculum implementation level program elements.

TABLE 4 Examples of fidelity of implementation items and informants.

FOI component	Measurement type	Sample item
Dosage	OR; IR	Total minutes of each session attended, additive to total minutes of program attended overall
	OR	Attendance (present, arrived late, left early)
Responsiveness	OR	Engagement (e.g., asking questions, active listening, making eye contact, lack of off task behavior)
	TR	Satisfaction with program element (e.g., "How satisfied are you with the training today?")
Quality	TR	Clarity of the [program elements] activities (e.g., "How clear was the content of this unit delivered during the training today?")
	OR	Observer ratings of implementor clarity of [program elements] activities (e.g., delivers instructions for activities, aids participants in completely activities; confusion is noticed and addressed; misinformation is corrected)
Adherence	OR	Presence or absence of each agenda item
	OR; IR	Degree objective was met for each agenda item—scale of 0 (did not meet any participant objective) to 4 (all participant objectives were met at an exemplary level)
Adaptations	OR; IR; TR	Qualitative notes on changes, additions, or omissions from planned program components
Control group monitoring	TR	Description of professional development opportunities (equivalent to implementation supports including curriculum trainings, community of practices, coaching)
	TR	Implementation of social emotional learning curriculum or practices, including MindUP™

Items in table represent a sample of constructs we measured and how they were operationalized. This list is not comprehensive for most fidelity of implementation components presented. OR, Observer Report; TR, Teacher Report; IR, Implementor Report.

example measures from previous studies examining FOI of the same curriculum that we were studying with different age groups (e.g., Schonert-Reichl et al., 2015). We also drew on an established set of FOI measures developed in the context of teacher-focused, mindfulness-based professional development programs (Doyle et al., 2019). Our final set of measures is the result of this joint effort among many researchers and we acknowledge their effort in this work. Those who are interested in additional detail regarding our comprehensive set of measures should contact the first author for more information.

Discussion

Developing a robust, comprehensive, multi-informant set of FOI measures for our multi-level SEL program creates numerous

opportunities to use the data collected to extend beyond program impacts to address other questions about a program of interest to program administrators, researchers, and practitioners alike. In the following section, we discuss several of these additional avenues of inquiry that become answerable with implementation study data. These areas include (a) a richer potential explanation for why program impacts were found (or not found), (b) a refinement of our understanding of the core components or "active ingredients" of the program, and (c) how FOI data can be used to improve SEL programs over time.

Using implementation data to understand program impacts

In our own work, we view implementation data as critical for answering *"Why did the program work (or not work)?."* Our implementation conceptual framework is meant to guide analyses that can assess both implementation and program outcomes, and therefore, how implementation affects intervention effectiveness. In our study, implementation outcomes are those that, when taken together, allow us as researchers to understand whether or not our implementation supports were successful in transferring the knowledge and skills needed to teachers (as intervention agents) to successfully implement the curriculum in their classroom, and whether or not teachers were successful in transferring the knowledge and skills of the curriculum to their students to promote positive outcomes (Fixsen et al., 2005; Dunst and Trivette, 2012).

Examples of how implementation outcomes are hypothesized to impact intervention outcomes in our study are displayed in Figure 3. Specifically, this figure depicts an example chain of hypotheses that link implementation supports for teachers (first level) to teachers' curriculum implementation in the classroom (second level) and then on to intervention outcomes for children, using the FOI components of quality, adherence, and responsiveness to demonstrate. To analyze these relations, first, we intend to use implementation data to understand the extent to which our program implementation was successful. Examples of implementation agent outcomes that would indicate success in this area are: that our implementation director was on time, organized, and prepared for the training, community of practice, and coaching sessions (i.e., prepared to provide a high quality session); that they were knowledgeable and implemented each session with fidelity (i.e., high adherence), and that they were also respectful to participants and responsive to their needs.

Second, we plan to examine relations between implementation agent outcomes and teacher implementation outcomes. As is displayed in the first example of Figure 3 (Implementation Supports Level), we hypothesize that the quality of the implementation director outcomes mentioned above will influence teachers' participation in the program. Examples of teacher outcomes that would indicate success in this area are: that all teachers attend the curriculum training (100% attendance); that they find the training content useful and relevant, and are engaged in the training (high responsiveness); and that they leave the training feeling efficacious in implementing the curriculum with their students. After the training, we also desire to see additional implementation support outcomes that indicate success such as: that teachers attend the majority of ongoing implementation support sessions throughout the study (high dosage of additional supports); that teachers find the sessions engaging and supportive (high responsiveness); and that they will implement the curriculum to a higher level of fidelity through participation in these regular touchpoints and supports. As is displayed in the second example in Figure 3 (Implementation Supports to Curriculum Implementation), we will also be able to determine if greater attendance and engagement in these supports is associated with teachers' fidelity of implementation (adherence) of the curriculum in their classroom, such that teachers who attend support sessions more frequently and/or are more engaged in these sessions will implement the curriculum in their classroom to a higher degree of fidelity compared to teachers who do not attend as frequently, or are not as engaged.

Third, we also plan to examine whether teachers' implementation quality and adherence are associated with student responsiveness to the curriculum (see the third example of Figure 3, Curriculum Implementation Level). Specifically, we hypothesize that teachers who implement curriculum activities on schedule (i.e., high level of adherence), with the majority of students in attendance for each lesson and for each daily Brain Break offered (high dosage), and in a manner that is supportive of where students are developmentally, will lead to higher levels of student engagement and responsiveness overall.

Finally, we plan to test the relations between program implementation and child outcomes. In the final example of Figure 3 (Curriculum Implementation to Outcomes), we hypothesize that implementation support outcomes (indirectly) and curriculum implementation outcomes (directly) will influence child outcomes (impacts). Specifically, we hypothesize that teachers who implement curriculum activities on time, as designed, in a high quality and engaging manner will contribute to larger impacts for their students' social emotional and academic outcomes compared to other teachers who implement activities less frequently, at a lower level of quality, or with a lower degree of adherence. However, we only expect to see larger impacts for students if student attendance and engagement are high.

When displayed in the context of a logic model (such as in Figure 3), the interconnectedness between intervention implementation and outcomes becomes clear. In an ideal scenario, researchers could engage in analyses, such as the ones we have outlined above, to understand if the intervention was implemented as planned. This can help to ensure that the intervention "on paper" matches its effectiveness through its implementation "in practice", as well as any conclusions made about its effectiveness. We believe that it is only when the implementation is understood

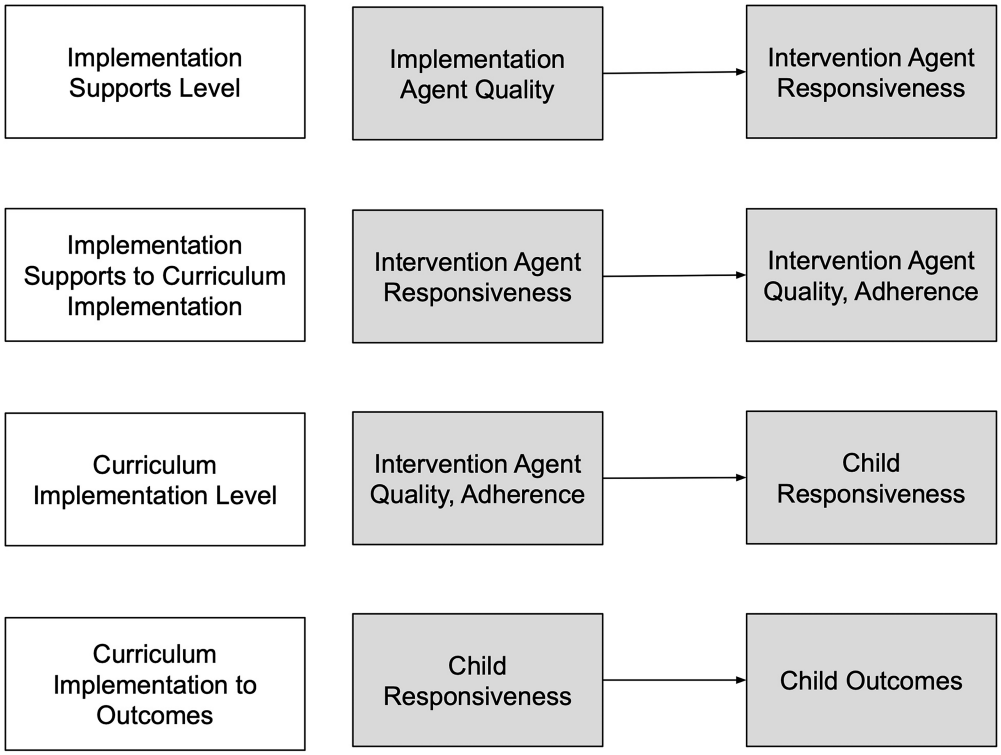


FIGURE 3
Example hypothesized relations between and within intervention implementation levels and outcomes. Outcomes refers to program impacts for children. In our study, implementation agent refers to our implementation director and intervention agent refers to preschool teachers implementing the MindUP™ curriculum in their classroom with students.

in these basic ways that one should move on to testing efficacy for the impact study. Furthermore, publishing implementation and impacts together may reduce the bias present in published findings and support more accurate metanalytic reviews (Torgerson and Torgerson, 2008).

Using implementation data to understand core program components

Once a program has been found to be efficacious, implementation data can also be used to explore why a program was effective and answer "*what are the core components of the program?*" (see Baelen et al., 2022). Core components can be defined as "the most essential and indispensable components of an intervention practice or program" (Fixsen et al., 2005, p. 24). The goal of this framework is to carve away everything nonessential to the program over time to ensure the intervention promotes valuable activities (i.e., "active ingredients") and only demands what is necessary from participants for positive outcomes (Moir, 2018).

Determining the core components of a program requires an iterative process of testing and refining an intervention or program over time (Fixsen et al., 2005). In the initial implementation of a program, implementation data can be used to clarify and solidify the core components of the program being tested (see Baelen et al., 2022). Ideally, researchers will not stop at determining that an intervention is effective, but rather continue investing in a program with additional studies that test the accuracy and impact of specific core components. This can aid in refining the program as needed by adding or subtracting core components and/or delivery methods of program elements (Fixsen et al., 2005). This work between initial testing and future implementation efforts is imperative to engage in a process of continuous improvement, to reduce resources wasted on implementation components that are non-essential, and to explore context-specific considerations of the program (e.g., does it work for everyone? In all setting types? For all ages?). Additionally, the effectiveness of a program would be expected to improve by paring down nonessential elements, as participants may receive a higher dose of the "active ingredients" of the intervention, rather than a combination of inactive and active components (Fixsen et al., 2005).

More specifically, dosage data can be used to determine active ingredients of the intervention, by examining if specific activities or practices were drivers of intervention outcomes. These analyses examine the extent to which dose of the intervention component for each participant relates to their outcomes or benefits gained from the intervention. These explorations provide interesting insights by going beyond the primary question of the impact

study. However, they involve teachers in the treatment group only and rely on quasi-experimental methods (classrooms were not randomly assigned to different levels of implementation). Nonetheless, they can be valuable additional tests that explore questions about what program factors are related to positive intervention impacts (Baelen et al., 2022; Roeser et al., 2022).

In our study of MindUP™, we plan to examine the unique effect of the Brain Break practice, as it is likely a primary core component of this program and also a practice that differentiates this program from many other SEL curricula. We hypothesize that students who receive a "higher dose" of the Brain Break will demonstrate greater gains in cognitive and behavioral measures of self-regulation at the end of the year. Beyond simply examining dosage, we also hypothesize that the impact of the brain break will be greatest for children in classrooms where teachers implemented the Brain Break in a manner that is engaging for students. Thus, we plan to use a combination of the dosage (total number of Brain Breaks offered to the class), student attendance, and child engagement (as a measure of responsiveness) to examine whether the Brain Break is a core component of this program.

Using implementation data to improve the program over time

Analysis of implementation data can also support a process of continuous improvement in the program under evaluation in an impact study. By identifying and understanding the core or active ingredients of an intervention, researchers are in a better position to answer the question *"how do we make the program better?"* in the future. While there are numerous ways to approach program improvements, for this section, we focus on three areas in terms of program optimization, including improvements for teaching, enhancement of active ingredients, and considerations focused on equity.

Implementation data collected as part of an impact study can be used to support teachers' future implementation of a program. In general, FOI data can be synthesized to highlight strengths across the group of implementors, as well as to identify common challenges with implementation. Feedback on adherence can be particularly useful if teachers plan to continue implementing the program after the intervention is over to support high fidelity and counteract program drift over time (see Domitrovich et al., 2012). For example, in our own study, we plan to analyze teachers' use of materials to support and scaffold students in the Brain Break practice and provide teachers with a summary of the most frequently used supplementary materials and tools. This implementation summary resource may be helpful to teachers who found this aspect of implementation challenging.

Similarly, an analysis of FOI components collected regarding coaching may reveal certain aspects of implementation that were particularly challenging for teachers (intervention agents) and that could use careful review for potential revisions in future implementations by the curriculum developers. For example, our coaching log asks teachers whether they would implement each activity in the curriculum again and if they report that they would not, asks them to elaborate on reasons why. We also ask teachers on the coaching log whether they have any feedback for the implementation team regarding each curriculum activity. Having the implementation director read the coaching logs throughout the program allows for a direct line between teachers and the implementation team in a manner that supports this work. Additionally, researchers can acknowledge and honor teachers' expertise by exploring adaptations they implemented, taking careful note of any that may be particularly useful for diverse learners and that could be incorporated into curriculum revisions to improve the program. These are just several examples that highlight how researchers can give back the data to those participating in ways that are useful to them and their profession.

Analyses examining dose–response relations, as described in the previous section, can also inform potential revisions to a program by helping to determine whether the "full-dose" was necessary to produce impacts or if a smaller dose is effective. Given the time demands associated with teaching, if a smaller dose is found to be effective, reducing program demands may actually increase teachers' ability to adhere to the program, and on a broader scale, may increase the total number of teachers who find the program feasible to implement. Furthermore, examining unique effects of each intervention element can establish if specific intervention elements are driving intervention outcomes, in which case these active ingredients could be amplified in future implementations to maximize positive outcomes for participants. We have also learned that incorporating a new FOI measure of generalization (for example see Low et al., 2016) into our future work will likely be important in this vein, to better understand the informal use of the program through integration and reinforcement of program components into daily classroom activities.

Finally, FOI data can also inform future changes to a program if null or negative impacts are found for specific subgroups of participants through moderated impacts analyses (Bywater, 2012). From an equity perspective, this area is essential to ensure an intervention touted as universal does in fact equally benefit individuals from diverse backgrounds. To highlight this issue, Rowe and Trickett (2018) conducted a meta-analytic review examining various diversity characteristics and how they moderated impacts of SEL interventions for students. One prominent finding of this study was the general lack of attention on this issue overall, in which most studies included in the review did not test for differential impacts. For those that did, the authors found mixed program impacts dependent on the diversity characteristic examined. These results speak to the need for examining these issues, particularly in the context of interventions considered appropriate for students of all cultures and backgrounds. Furthermore, we argue these analyses should be examined even if a positive impact of the intervention is found overall for participants, to determine if specific subgroups did not benefit, or alternatively, if certain subgroups benefited significantly

more than others. In the latter case, these analyses can be used to further explore active ingredients of the intervention and potentially inform program changes to increase the impact across all participants in the future.

Regardless of whether we find an overall positive effect on children's outcomes in our study, we plan to examine whether effectiveness of the program varied for different subgroups of children. In particular, we are interested in understanding whether the effectiveness of the program is the same or different based on the classroom setting and/or type of program. As described in our theory of change, we hypothesize that the effects of this program will be stronger in programs in which the quality of interactions between teachers and students is lower at baseline. Furthermore, it will be important to examine whether the program is effective in a variety of program types (e.g., community-based program, Head Start, for-profit) to better understand whether the program context impacts the effectiveness of the program. Lastly, if we find that the impact of the program varies for subgroups of children, it will be important to examine teachers' adherence and quality data, as well as students' dosage and responsiveness data, to understand whether implementation differed across subgroups. If differences in implementation by teachers are not found, but student engagement differs, it may indicate that the program was less relevant for some children. Each of these examinations of differential impacts can inform program changes in the future to enhance the program's effectiveness.

Conclusion

The main purpose of this paper was to demonstrate the interconnected nature between implementation and impacts in the context of studying complex, multi-level education interventions generally, and SEL programs more specifically. We believe, as others do, that robust FOI data provide essential contextual information about an intervention's inner workings. In this vein, these data inform both the internal validity of the conclusions from our impacts study and provide valuable information that can be used to support effective application in real-world contexts. We aimed to illustrate this key point by drawing on our own learnings developing the conceptual framework and associated FOI measures for our ongoing evaluation study testing a mindfulness-based SEL program for preschoolers, and by exploring numerous ways these data can extend beyond simple tests of intervention effectiveness and be used to describe, explain, and optimize education programs.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

JC, RR, and AM contributed to formulating the study design and measures and contributed to the development of the conceptual framework. JC led the development of the fidelity of implementation measures described in depth in this paper, with support from the other authors and those mentioned in the acknowledgments and wrote the first draft of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through grant R305A180376 awarded to Portland State University. RR was supported in this work by the Edna Bennett-Pierce Chair in Caring and Compassion.

Acknowledgments

The authors offer sincere thanks to Sebrina Doyle and Caryn Ward for their contributions to the development of the fidelity of implementation measures used in this study. We also thank the Mindful Pre-Kindergarten research and implementation teams, including Corina McEntire, Tessa Stadeli, Eli Labinger, Brielle Petit, Cristin McDonough, and the outstanding research assistants that have supported this work.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Author disclaimer

The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education.

References

Baelen, R. N., Gould, L. F., Felver, J. C., Schussler, D. L., and Greenberg, M. T. (2022). Implementation reporting recommendations for school based mindfulness programs. *Mindfulness*. doi: 10.1007/s12671-022-01997-2

Belfield, C., Bowden, A. B., Klapp, A., Levin, H., Shand, R., and Zander, S. (2015). The economic value of social and emotional learning. *J. Benefit Cost Analys.* 6, 508–544. doi: 10.1017/bca.2015.55

Bickman, L., and Reich, S. M. (2015). *"Randomized controlled trials: A gold standard or gold plated?" in credible and actionable evidence: The Foundation for Rigorous and Influential Evaluations*. Thousand Oaks, CA: SAGE Publications, Inc., 83–113.

Boruch, R. (2005). Better evaluation for evidence-based policy: place randomized trials in education, criminology, welfare, and health. *Ann. Am. Acad. Pol. Soc. Sci.* 599, 6–18. doi: 10.1177/0002716205275610

Boruch, R., de Moya, D., and Snyder, B. (2002). "The importance of randomized field trials in education and related areas" in *Evidence matters: Randomized trials in education research*. eds. F. Mosteller and R. Boruch (Washington, DC: Brookings Institution Press), 50–79.

Braun, S. S., Roeser, R. W., Mashburn, A. J., and Skinner, E. (2018). Middle school teachers' mindfulness, occupational health and well-being, and the quality of teacher-student interactions. *Mindfulness* 10, 245–255. doi: 10.1007/S12671-018-0968-2

Bryk, A. S. (2016). Fidelity of implementation: is it the right concept? Carnegie commons blog. Available at: https://www.carnegiefoundation.org/blog/fidelity-of-implementation-is-it-the-right-concept/ (Accessed 14 July 2022).

Bywater, T. (2012). "Developing rigorous programme evaluation" in *Handbook of implementation science for psychology in education*. eds. B. Kelly and D. F. Perkins (New York, NY: Cambridge University Press), 37–53.

Dane, A. V., and Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clin. Psychol. Rev.* 18, 23–45. doi: 10.1016/s0272-7358(97)00043-3

Domitrovich, C. E., Moore, J. E., and Greenberg, M. T. (2012). "Maximizing the effectiveness of social-emotional interventions for young children through high quality implementation of evidence-based interventions" in *Handbook of implementation science for psychology in education*. eds. B. Kelly and D. F. Perkins (New York, NY: Cambridge University Press), 207–229.

Doyle, S. L., Jennings, P. A., Brown, J. L., Rasheed, D., DeWeese, A., Frank, J. L., et al. (2019). Exploring relationships between CARE program fidelity, quality, participant responsiveness, and uptake of mindful practices. *Mindfulness* 10, 841–853. doi: 10.1007/s12671-018-1034-9

Dunst, C. J., and Trivette, C. M. (2012). "Meta-analysis of implementation practice research" in *Handbook of implementation science for psychology in education*. eds. B. Kelly and D. F. Perkins (New York, NY: Cambridge University Press), 68–91.

Durlak, J. A., and DuPre, E. P. (2008). Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am. J. Community Psychol.* 41, 327–350. doi: 10.1007/s10464-008-9165-0

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

Fixsen, D. L., and Blase, L. A. (2009). *Implementation: The missing link between research and practice* No. 1; NIRN Implementation Brief; The National Implementation Research Network. Chapel Hill, NC: University of North Carolina at Chapel Hill.

Fixsen, D. L., Naoom, S. F., Blasé, K. A., Friedman, R. M., and Wallace, F. (2005). *Implementation research: A synthesis of the literature* FMHI Publication #231; The National Implementation Research Network. Tampa, FL: University of South Florida.

Hollis, S., and Campbell, F. (1999). What is meant by intention to treat analysis? Survey of published randomised, controlled trials. *Br. Med. J.* 319, 670–674. doi: 10.1136/bmj.319.7211.670

Jennings Brown, J. L., Frank, J. L., Doyle, S., Oh, Y., Davis, R., Rasheed, D., et al. (2017). Impacts of the CARE for teachers program on teachers' social and emotional competence and classroom interactions. *J. Educ. Psychol.* 109:1010. doi: 10.1037/edu0000187, 1028

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

Luo, L., Reichow, B., Snyder, P., Harrington, J., and Polignano, J. (2022). Systematic review and meta-analysis of classroom-wide social–emotional interventions for preschool children. *Top. Early Child. Spec. Educ.* 42, 4–19. doi: 10.1177/0271121420935579

Mac Donald, E., and Shirley, D. (2009). *The mindful teacher*. New York: Teachers College Press.

Maloney, J. E., Lawlor, M. S., Schonert-Reichl, K. A., and Whitehead, J. (2016). *A mindfulness-based social and emotional learning curriculum for school-aged children*. New York, NY: Handbook of mindfulness in education, Springer, 313–334.

Mark, M. M., and Lenz-Watson, A. L. (2011). "Ethics and the conduct of randomized experiments and quasi-experiments in field settings" in *Handbook of ethics in quantitative methodology*. eds. A. T. Panter and S. K. Sterba (New York, NY: Routledge), 185–209.

McCormick, M. P., Cappella, E., O'Connor, E., Hill, J. L., and McClowry, S. (2016). Do effects of social-emotional learning programs vary by level of parent participation? Evidence from the randomized trial of INSIGHTS. *J. Res. Educ. Effect.* 9, 364–394. doi: 10.1080/19345747.2015.1105892

Moir, T. (2018). Why is implementation science important for intervention design and evaluation within educational settings? *Front. Educ.* 3, 1–9. doi: 10.3389/feduc.2018.00061

Roeser, R. W., Galla, B. M., and Baelen, R. N. (2020). Mindfulness in schools: Evidence on the impacts of school-based mindfulness programs on student outcomes in P-12 educational settings. Available at: www.prevention.psu.edu/sel

Roeser, R. W., Greenberg, M., Frazier, T., Galla, B. M., Semenov, A., and Warren, M. T. (2022). Envisioning the next generation of science on mindfulness and compassion in schools. *Mindfulness*. 114, 408–425. doi: 10.1007/s12671-022-02017-z

Rowe, H. L., and Trickett, E. J. (2018). Student diversity representation and reporting in universal school-based social and emotional learning programs: implications for generalizability. *Educ. Psychol. Rev.* 30, 559–583. doi: 10.1007/s10648-017-9425-3

Sandilos, L. E., Neugebauer, S. R., DiPerna, J. C., Hart, S. C., and Lei, P. (2022). Social–emotional learning for whom? Implications of a universal SEL program and teacher well-being for teachers' interactions with students. *Sch. Ment. Heal.* doi: 10.1007/s12310-022-09543-0, 1, 12

Schonert-Reichl, K. A., Oberle, E., Lawlor, M. S., Abbott, D., Thomson, K., Oberlander, T. F., et al. (2015). Enhancing cognitive and social–emotional development through a simple-to-administer mindfulness-based school program for elementary school children: a randomized controlled trial. *Dev. Psychol.* 51, 52–66. doi: 10.1037/a0038454

Shadish, W. R., Cook, T. D., and Campbell, D. T. (2002). *Experimental and quasi-experimental designs for general causal inference*. Boston: Houghton Mifflin.

Snyder, P. A., Hemmeter, M. L., and Fox, L. (2015). Supporting implementation of evidence-based practices through practice-based coaching. *Top. Early Child. Spec. Educ.* 35, 133–143. doi: 10.1177/0271121415594925

Torgerson, D. J., and Torgerson, C. J. (2008). *Designing randomised trials in health, education and the social sciences*. London, UK: Palgrave Macmillan.

frontiers | Frontiers in Psychology

TYPE Methods
PUBLISHED 09 January 2023
DOI 10.3389/fpsyg.2022.1014713

OPEN ACCESS

EDITED BY
Stephanie M. Jones,
Harvard University, United States

REVIEWED BY
Kiat Hui Khng,
Nanyang Technological University, Singapore
Sophie Barnes,
Harvard University, United States
Patricia Snyder,
University of Florida, United States

*CORRESPONDENCE
Brianna L. Devlin
✉ bldevlin@uoregon.edu

†PRESENT ADDRESS
Brianna Devlin, Elyssa Geer, Tracy Zehner and Sara Schmitt, Prevention Science Institute, University of Oregon, Eugene, OR, United States

SPECIALTY SECTION
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

RECEIVED 08 August 2022
ACCEPTED 20 December 2022
PUBLISHED 09 January 2023

CITATION
Devlin BL, Paes TM, Geer EA, Bryant LM, Zehner TM, Korucu I, Morse K, Duncan RJ, Purpura DJ and Schmitt SA (2023) Moving beyond dosage and adherence: A protocol for capturing dimensions of active child engagement as a measure of fidelity for social-emotional learning interventions. *Front. Psychol.* 13:1014713. doi: 10.3389/fpsyg.2022.1014713

Moving beyond dosage and adherence: A protocol for capturing dimensions of active child engagement as a measure of fidelity for social-emotional learning interventions

Brianna L. Devlin[1*†], Tanya M. Paes[1], Elyssa A. Geer[1†], Lindsey M. Bryant[1], Tracy M. Zehner[1†], Irem Korucu[2], Kathleen Morse[1], Robert J. Duncan[1], David J. Purpura[1] and Sara A. Schmitt[1†]

[1]Human Development and Family Studies, Purdue University, West Lafayette, IN, United States, [2]Yale Center for Emotional Intelligence, School of Medicine, Yale University, New Haven, CT, United States

Social-emotional competencies are important for school-readiness and can be supported through social-emotional learning (SEL) interventions in the preschool years. However, past research has demonstrated mixed efficacy of early SEL interventions across varied samples, highlighting a need to unpack the black box of which early interventions work, under what conditions, and for whom. In the present article we discuss the critical implementation component of active child engagement in an intervention as a potential point of disconnect between the intervention as designed and as implemented. Children who are physically present but unengaged during an intervention may lead to decreased average impacts of an intervention. Furthermore, measuring young children's active engagement with an intervention may help to guide iterative intervention development. We propose a four-step protocol for capturing the multi-dimensional and varied construct of active child engagement in a SEL intervention. To illustrate the utility of the protocol, we apply it to data from a pilot study of a researcher-implemented, semi-structured block play intervention focused on supporting the development of SEL and math skills in preschoolers. We then present future directions for the integration of active participant engagement into the measurement of implementation of SEL interventions for young children.

KEYWORDS

social-emotional learning (SEL), social-emotional learning interventions, implementation, child engagement, participant responsiveness, preschool

1. Introduction

Social-emotional competencies, including getting along with others, paying attention to and following directions, and regulating behaviors and emotions are crucial skills for children to develop prior to school entry as they predict later academic achievement and well-being (McClelland et al., 2007, 2013; Schmitt et al., 2017; Mackintosh and Rowe, 2021). However, interventions targeting social-emotional competencies in preschool have had mixed success, and scholars have recently suggested this may be due to issues with implementation fidelity, or perhaps, the way fidelity is typically measured (McClelland et al., 2017). For example, in the large majority of studies, implementation fidelity is typically assessed using simple measures of dosage (i.e., "Did they do it?") and adherence (i.e., "Did it align with the guidelines?"). Active participant engagement (i.e., participation in the intervention) is seldom measured despite its inclusion in theoretical models of implementation (Century et al., 2010; Berkel et al., 2011). When active participant engagement in SEL interventions is included as a measure of fidelity, it almost exclusively concerns the active engagement of adults (e.g., teachers) with the intervention, and not child participant engagement.

Discounting how the autonomy of young children could lead to individual differences in children's active participation during an intervention may lead to biased estimates of the efficacy of early SEL interventions. That is, a child may be physically present at an intervention session, but not actively engaging with the target material. Moreover, they may actively engage with only specific aspects or activities of the intervention but not others. These differences would not be captured by traditional measures of child-level dosage and could result in diminished effectiveness of an intervention as measured by estimates of average impacts. By assessing active child engagement as a multi-dimensional construct of implementation fidelity, we may be better able to capture individual differences in children's experiences that moderate the effect of an intervention. Furthermore, considering the nuanced ways in which young children engage with an intervention may help researchers to iteratively develop interventions that support children from diverse backgrounds (e.g., considering differences by socio-cultural groups, site specific needs, etc.). Consequently, in the present article, we introduce a four-step protocol developed to capture active child engagement with SEL interventions. We use data from a pilot study of a semi-structured block play intervention (Schmitt et al., 2018a) as an example of applying the protocol to capture dimensions of active child engagement.

1.1. Supporting children's early social-emotional learning through targeted interventions

Our broad definition of SEL interventions include intervention programs aimed at supporting the development of social skills, emotion regulation, and cognitive regulation (i.e., executive functions; McClelland et al., 2017). This definition is in line with the Collaborative for Academic Social and Emotional Learning (CASEL)'s framework that emphasizes the importance of competencies spanning social, emotional, and cognitive regulatory processes (CASEL, 2017). Several SEL interventions have been created to bolster these competencies in early childhood (e.g., McClelland et al., 2017; Nesbitt and Farran, 2021), a critical period in which rapid changes occur in cortical brain structures that are vital for SEL and cognitive development (Garon et al., 2008). This developmental period also coincides with children engaging with adults and peers in multiple settings (e.g., home, preschool; Schmitt et al., 2018a), in which SEL skills are necessary for successful relationships.

A common theme across the majority of early SEL interventions is that they are child-centered, which is thought to promote active child engagement with intervention content (Massouleh et al., 2012). For example, the Red Light, Purple Light intervention (Tominey and McClelland, 2011; Schmitt et al., 2015; McClelland et al., 2019) centers children's experiences and active engagement by using fun and age-appropriate music and movement activities designed to promote behavioral self-regulation. During this intervention, children are given agency in interacting with the games as they choose and are also offered opportunities for autonomy in leading games. As another example, the Preschool Alternative THinking Strategies (PATHS) program offers explicit instruction in SEL through teacher-led lessons and extension activities like group games and art projects (Domitrovich et al., 2007). These activities were designed to be fun and engaging for young children while also promoting their SEL skills broadly. Brain Games (Barnes et al., 2021) is another classroom-based intervention for young children, focused on building attention, working memory, and inhibitory control through games.

Despite the fact that many SEL interventions take a child-centered approach, evidence of the efficacy of these interventions is mixed (McClelland et al., 2017; Nesbitt and Farran, 2021). Although some interventions like Red Light, Purple Light (Tominey and McClelland, 2011; Schmitt et al., 2015) and PATHS (Domitrovich et al., 2007) have shown positive effects on their target outcomes in preschoolers, other programs have shown mixed or null effects. For example, in a recent study Nesbitt and Farran (2021) found no positive impacts of the Tools of the Mind curriculum for supporting SEL in early childhood. Of note, Nesbitt and Farran were not the original developers of the Tools of the Mind curriculum. When researchers analyzed the implementation

fidelity factors of adherence and dosage they found that intervention classrooms had varied dosage but implemented about half of what the developers expected, on average. However, neither the adherence to nor amount of time spent on the Tools curriculum was statistically related to children's outcomes.

1.2. Integrated models of intervention implementation

To unpack the nuances of which early SEL interventions work under which conditions and for whom, it is essential to consider multiple factors of implementation fidelity. Theoretical frameworks of program implementation include factors controlled by the intervention designers (e.g., differentiation of intervention practices from currently enacted practices), factors controlled by the interventionist (e.g., script adherence and quality of delivery), and factors controlled by the participant (e.g., attendance and active engagement; Dane and Schneider, 1998; Carroll et al., 2007; Berkel et al., 2011). For example, Berkel et al.'s (2011) integrated model of program implementation differentiated behaviors that occur at the time of implementation that are directed by the interventionist from those directed by the participant (i.e., participant responsiveness). Furthermore, interventions can sometimes include multiple levels of implementation, such as when researchers train providers, who go on to train teachers or parents, who then teach children. Responsiveness to the intervention as intended includes attendance and dosage, retention, satisfaction, and active participant engagement, or participation in an intervention (Berkel et al., 2011). However, reviews of implemented educational interventions have demonstrated that interventionist-controlled factors (especially adherence) are more commonly reported than measures of participant responsiveness (O'Donnell, 2008). When responsiveness is considered, dosage is the most often-reported measure (Bos et al., 2022). Measurement of active participant engagement in early childhood interventions has primarily been concentrated on the adult participants who implement the intervention with children (e.g., teachers and parents' active participation with the intervention). Some work has started to unpack the complex relations among fidelity indicators, including active participant engagement, at differing levels of an intervention. For example, Berkel et al. (2018) tested a theoretical cascade model in which facilitator delivery (e.g., adherence of the provider) predicted participant responsiveness (e.g., parents' home practice), which in turn lead to improvements in the targeted outcomes (e.g., children's mental health). However, very few studies have considered participant responsiveness fidelity indicators, such as active participant engagement, at the child-level of interventions designed to improve outcomes for young children.

2. Capturing dimensions of active child engagement

Given the lack of studies that report active child engagement—especially for young children—we sought to develop a general protocol for capturing young children's active engagement with interventions as a measure of fidelity. The development of this protocol was guided by a conceptual framework comprised of the following assumptions:

- Individual differences in active child engagement may influence intervention efficacy.
- Contextual and individual factors may shape differences in active child engagement with an intervention.
- Active child engagement with an intervention can be measured through observing behavior.
- There are multiple dimensions to active child engagement with an intervention.
- There is variability in active child engagement with an intervention.

2.1. Individual differences in active child engagement may influence intervention efficacy

Active child engagement occurs within each intervention session, and thus "serves as a potential source of disconnect between the program as intended and the program as implemented" (Berkel et al., 2011, p. 23), making it a worthwhile focus of effort for measuring implementation fidelity. Dosage is a commonly reported measure of children's participation in an intervention, but this fidelity indicator does not provide information about whether children experienced the intervention as it was intended. That is, children who are physically present but unengaged during intervention sessions may lead to decreased efficacy of the intervention as measured by estimates of average impacts. Therefore, it is important to consider how active child engagement may moderate the effect of an intervention.

2.2. Contextual and individual factors may shape differences in active child engagement with an intervention

Beyond considering how individual differences in engagement may influence intervention efficacy, it is also

important to consider what factors predict differences in engagement with an intervention. School level factors such as discipline policies or classroom-level factors such as classroom culture and norms may affect active child engagement with a SEL intervention. For example, when schools and centers enact disciplinary measures that involve pulling children from an activity or classroom, they increase the chances of disrupting active child engagement. Classroom factors like the number of children in play groups have been associated with children's positive engagement with peers, as smaller groupings promote more cooperative play (Howes et al., 2011). Intervention factors such as who is implementing the intervention (e.g., teachers or researchers) may also influence active child engagement. Finally, child level factors such as demographic background and baseline skills may influence children's active engagement. By conducting in depth analyses of how children engage with an intervention as a measure of responsiveness, we can gather information that may challenge researchers' assumptions about how children respond to materials and prompts. This is an especially important undertaking when working with historically under-represented populations in past research. This information may be used to inform iterative development of the intervention in an effort to create interventions that promote active child engagement for children from diverse backgrounds.

2.3. Active child engagement with an intervention can be measured through observing behavior

Although most research focused on active participant engagement has utilized participant report with adult participants (Carroll et al., 2007), engagement can be assessed through observing behavior, opening the door for assessment of engagement in young children. For example, Ling and Barnett (2013) used behavior observations time-sampled in 15-s intervals to assess groups of preschoolers' engagement in circle time activities. Another researcher-developed measure is the inCLASS (Downer et al., 2010), which is an observational tool focused on capturing preschoolers' interactions with teachers, peers, and classroom activities. In the context of early SEL interventions, Schmitt et al. (2018b) used a similar measure to assess preschoolers' on-task behavior during the Positive Action intervention, aimed at improving social-emotional competence and health behaviors. One efficacy study of the Tools of the Mind intervention also included a measure of child engagement with the intervention using a Likert scale based on observations of child behavior ranging from completely off-task to intense focus across time-sampled intervals (Child Observation in Preschool, COP; Nesbitt and Farran, 2021).

2.4. There are multiple dimensions to active child engagement with an intervention

The aforementioned work that measured and reported active child engagement in interventions approached it as a unidimensional construct and did not consider the complex nature of the intervention environment in which young children often have varied points of opportunity for engagement with an intervention as intended. For example, a child may appear to be actively engaged in an intervention activity (e.g., by being socially engaged with their peers) without engaging with the targeted SEL content. As another example, in a randomized trial of the Red Light, Purple Light intervention, the authors describe that although "the majority of children actively participated in all of the playgroup games... a few children chose to watch on occasion" (Tominey and McClelland, 2011 p. 513). Although these children were actively engaged with the content of the games by watching, they were not actively engaged with the physical movements of the games, a separate dimension of engagement. By considering active child engagement as a unidimensional construct, we miss capturing varied levels of engagement with the intervention, which may provide biased estimates of program outcomes. Furthermore, considering multiple dimensions of active child engagement allow for the ability to test which child participant-involved components are most important for growth in the target outcomes. Thus, a key extension of existing research is our consideration of active child engagement with an intervention as a multi-dimensional construct.

2.5. There is variability in active child engagement with an intervention

A final assumption behind our approach to measuring active child engagement is that it is varied, both between and within children. Decades of research in cognitive developmental psychology have converged to support the theory that development is defined by variability (Siegler, 2007). That is, differences in children's thinking and behavior exist between individual children, within the same child in different contexts, and even within the same child in the same context at a different point in time (Siegler, 1994). We believe that it is important to take this variability into account when approaching the measurement of active child engagement with an intervention. For example, a child's engagement may vary across intervention sessions in response to the content, or across the course of a single session. By operationalizing engagement in a way that takes variability into account, researchers can capture nuances such as the "implementation dip," or a decrease in performance or adherence in response to change (Fullan, 2001) in the context of active child engagement with SEL interventions.

3. Four-step protocol for capturing dimensions of active child engagement

In line with past theoretical frameworks that use a multi-step approach as a solution for measuring the complex construct of fidelity (e.g., Hulleman et al., 2013), we propose a four-step protocol for capturing active child engagement:

1. Identify points of opportunity for active child engagement with the intervention to specify multiple dimensions of active child engagement.
2. Operationalize and measure the dimensions of active child engagement.
3. Analyze the dimensions of active child engagement.
4. Link the dimensions of active child engagement to other variables.

3.1. Applying the four-step protocol to measure preschoolers' active engagement with the Block Play Intervention

In this section, we use data from a pilot study of the Block Play Intervention to illustrate application of the four-step protocol. The Block Play Intervention is a brief, semi-structured, play-based intervention aimed at supporting preschoolers' cognitive regulation, a critical component of SEL (McClelland et al., 2017), and mathematics through small group interactions. The intervention includes twice-weekly sessions of small group play (two to three children) with wooden unit blocks. Children are given specific building goals at the start of the session by a researcher interventionist (e.g., *"Today your job is to build a tower together!"*) but are then allowed to build freely. Over the course of the intervention sessions, the prompts gradually become more complex so that children's cognitive regulation is challenged, not just used (e.g., Week five: *"Today your job is to build a house together. . . It needs to have four walls, a roof, a way to get inside like a door, and at least two rooms"*; Week seven: *"Today I am going to show you a picture of a structure. Your job is to work together to build the structure you see in this picture."*). These prompts were targeted at priming children to work together and engage in social interaction to build structures in collaboration with their groupmates while also building their self-regulation skills through avoiding distractions and engaging in goal-oriented behavior.

A pilot study of this intervention included a sample of 59 children (M_{age} = 55.20 months), randomly assigned to the intervention group (n = 24) or a business as usual (BAU; n = 35) control group. On average, children in the intervention condition participated in 13 of the 14 sessions, with a range of 11 to 14 sessions attended. Researcher interventionists adhered to the building prompt scripts 94% of the time. The application of the four-step protocol for capturing active child engagement to the pilot data of the Block Play Intervention is presented visually in **Figure 1**. We use the context of this researcher-implemented pilot study as a straightforward example of the protocol's application but believe the four-step protocol can be scaled and utilized with much larger and more complex SEL interventions with young children.

3.1.1. Step 1: Identify points of opportunity for engagement to specify dimensions of active child engagement

The first step is to identify opportunities for potential active child engagement with the intervention. As with all aspects of fidelity, the identification of active child engagement opportunities should be closely tied to the intervention's overarching theory of change (Darrow, 2013). It may be helpful to think of these points of opportunity as participant-involved core intervention components, or the aspects of the intervention in which child participants are directly involved that are theorized to lead to change in the child-level outcome variables. The theory of change of the Block Play Intervention is presented in **Figure 2**.

In the Block Play Intervention, we hypothesized that participation in semi-structured block play sessions would lead to gains in the outcomes of interest through three points of potential active child engagement: the block play itself, positive interactions with peers, and working toward a provided, increasingly complex goal (see Row 1 of **Figure 1**). Block play provides children with opportunities to practice working with abstract concepts and representations, which may help to develop cognitive regulation (Wolfgang et al., 2001; Hadani and Rood, 2018). Children also get the opportunity to practice fine motor skills through block play, which relate to cognitive regulation and mathematical cognition development in the early years (Gashaj et al., 2019; McClelland and Cameron, 2019). Working cooperatively with peers by negotiating and engaging in prosocial behaviors may help children to develop their language and interpersonal skills (Sluss and Stremmel, 2004), as well as emotional and behavioral self-regulation skills like inhibiting prepotent responses (e.g., knocking down a peers' tower when angry). Finally, the goal-directed aspect of the activity may also help to strengthen children's cognitive regulation and academic skills as children are required to utilize metacognitive resources to remember and plan their building behavior in accordance with the goal (Schmitt et al., 2018a).

After identifying these points of opportunity for potential engagement, we used them to generalize and name the multiple dimensions of participant engagement that we wished to capture. For the Block Play Intervention, this resulted in three general dimensions of child engagement: Engagement with

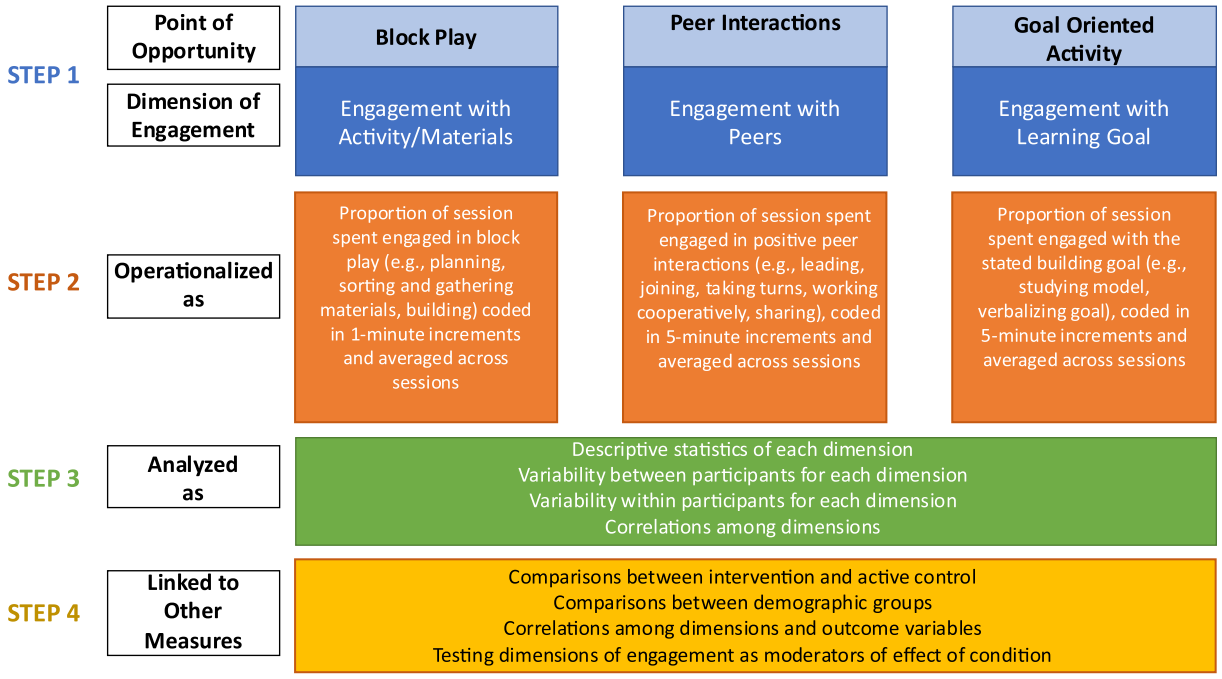


FIGURE 1
Applying the four-step protocol to measure preschoolers' active engagement with the Block Play Intervention.

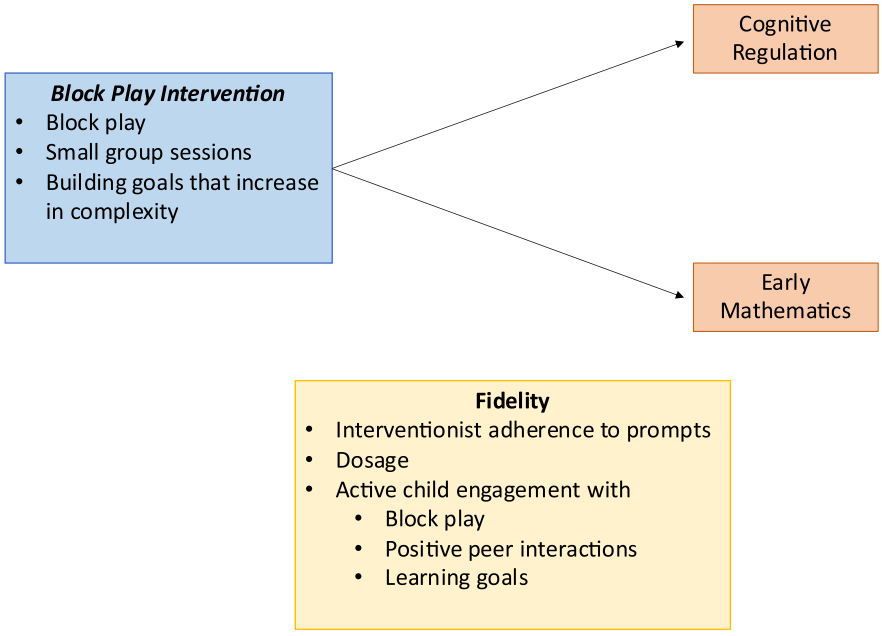


FIGURE 2
Theory of change for the Block Play Intervention.

the activity/materials, engagement with peers, and engagement with the learning goal (see Row 2 of Figure 1). Although our specified dimensions may apply to many play-based or semi-structured early SEL interventions, researchers wishing to use this protocol can alter or substitute these dimensions as needed. For example, SEL interventions that include teacher or parent interactions may need to add a dimension for children's engagement with a caregiver. Furthermore, researchers testing

more comprehensive interventions may wish to draw the points of opportunity for potential engagement from the specific activities in the intervention or create multiple levels (e.g., engagement with materials and peers within specific intervention activities like shared book reading, role playing, or group games).

3.1.2. Step 2: Operationalize and measure dimensions of active child engagement

The second step is to operationalize each dimension of active child engagement (see Row 3 in **Figure 1**) so that it can be quantified. Measurement decisions should be guided by pilot work, past literature, and the intervention theory of change. Researchers may choose to use previously developed measures of child engagement (e.g., inCLASS, Downer et al., 2010; COP, Farran, 2011) or to develop their own observational measures of specific behaviors.

3.1.2.1. Engagement with the activity/materials in the Block Play Intervention

Aligned with our conceptual framework, we chose to operationalize each dimension of child engagement with the Block Play Intervention through observation of behavior captured on video recordings of the 15-min block play sessions. For the engagement with activity/materials dimension, we used time-sampled observations of "on-task" behavior (Ling and Barnett, 2013). Coders watched videos and documented codes for each child individually, replaying sections of the video as needed. As past pilot work suggested that children would engage in on-task block play behavior for much of the session, we chose to use momentary time-sampling of whether or not children were actively engaged in on-task block play behavior after every minute of play. That is, it was important to use a higher frequency of coding instances (after every minute) to adequately capture variability and examine individual differences in children's engagement with the activity and materials. Children received a score of 0–15 for each session.

3.1.2.2 Engagement with peers in the Block Play Intervention

When operationalizing the engagement with peers dimension, we chose to focus on positive peer interaction behaviors as they were hypothesized to be related to gains in the outcome variables. We drew the positive peer interaction behaviors of interest from two sub-scales of the Minnesota Preschool Affect Checklist Revised/Shortened (M-PAC-R/S; Denham et al., 2012)–the Leading and Joining and the Empathy and Prosocial Behavior sub-scales. We coded in 5-min whole-interval time-sampled increments (a coarser level of analysis than the engagement with activity/materials dimension) in line with the protocol for using the M-PAC-R/S. Children therefore received a score of 0–3 for each session corresponding to positively engaging with peers during none of the three 5-min intervals (0) to positively engaging with peers during all of the three 5-min intervals (3). To code, we determined if a child engaged in any of these five behaviors at any time in the prior 5 min: successfully leading an activity, successfully joining an activity, facilitating turn-taking, cooperating with a peer or group to achieve a common goal, or exhibiting sharing behavior. For example, consider this dyad of children who were given the goal of building a castle for a king and queen. In the process, they engaged in leading and joining behaviors, taking turns in adding blocks to the structure, and working together cooperatively:

> Child A: [Pointing to blocks] "*So this is the king, that's the queen. King, queen. And we need a bottom slip. That will help. . . So that will go right here.*"
> Child B: "*Put it right on this.*"
> Child A: "*We need another one of these pieces.*"
> Child B: "*. . . I found one!*"

A different dyad of children took another cooperative building approach, where one child led the building activity, and another gathered and shared materials:

> Child C: "*Wait, wait. . . I need baby triangles. . . I need another triangle. . . A baby one. Will you find it for me?*"
> Child D: *"I will find it for you."*
> Child C: [taking triangle from Child D] "*This goes here.*"

3.1.2.3 Engagement with learning goals in the Block Play Intervention

Finally, we operationalized engagement with the learning goal as the proportion of the session a child spent engaged with the explicit building goal given to them by the interventionist, as evidenced by verbal and non-verbal behaviors. This dimension was also coded in whole-interval 5-min increments by determining whether a child spent any of the time in the previous 5 min building the assigned structure for the session. Children therefore received a score of 0–3 for each session corresponding to engaging with the building goal during none of the three 5-min intervals (0) to engaging with the building goal during all of the three 5-min intervals (3). Examples of children who did not engage with the building goal varied. For example, when asked to build a tower, one child responded, *"No thank you. We are making a playground."* Other children chose not to engage with the given building goal, and instead requested to be able to build freely. For example, when children were asked to model their structure after a picture, one child responded, *"I'm not gonna build that. Can I just build?"*

For each dimension of engagement, we calculated the proportion of the session spent engaged in the specified behaviors of that dimension. That is, for the engagement

10.3389/fpsyg.2022.1014713

TABLE 1 Descriptive statistics by session.

Dimension of engagement		Session number													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Activity/Materials	*M*	12.3	12.4	12.3	11.9	12.4	10.5	12.6	11.6	12.3	13.0	12.5	12.6	11.1	12.7
0–15	*SD*	3.1	2.3	3.3	2.6	2.1	4.7	2.4	3.4	2.3	2.0	3.2	3.1	3.8	2.2
1-min intervals	Min	2.0	8.0	3.0	6.0	7.0	1.0	7.0	3.0	9.0	7.0	1.0	4.0	1.0	8.0
	Max	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0
Peers	*M*	1.9	2.0	2.2	1.7	2.0	1.4	1.8	1.2	1.5	0.9	2.3	1.8	1.5	2.1
0–3	*SD*	1.1	1.2	0.9	1.2	1.1	1.2	1.3	1.2	1.2	1.0	0.9	1.3	1.1	0.9
5-min intervals	Min	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Max	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
Learning goals	*M*	2.4	2.5	1.9	2.3	2.6	2.2	2.4	2.4	2.5	2.5	2.2	2.1	2.0	2.0
0–3	*SD*	0.9	0.9	1.4	1.1	0.7	0.6	0.6	0.9	0.8	0.5	1.1	1.0	1.1	1.0
5-min intervals	Min	0.0	0.0	0.0	0.0	1.0	1.0	1.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0
	Max	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0

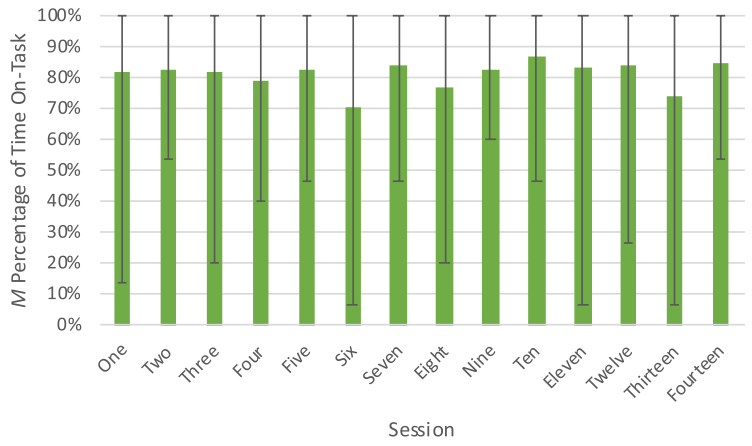


FIGURE 3
Engagement with the activity/materials across sessions of the Block Play Intervention. Figure shows percentage of engagement out of total coded time (15 min) with the activity/materials at each session. The range from minimum to maximum engagement is represented by error bars.

with activity/materials dimension, we calculated the proportion of minutes children were engaged in on-task block play behavior out of the total 15 min. For the engagement with peers and learning goals dimensions, we calculated the proportion of intervals children were engaged in either leading, joining, taking turns, working cooperatively and sharing behaviors or goal-oriented behaviors out of the total three coded intervals. Proportions were averaged across the 14 sessions to create an overall engagement score for each dimension throughout the course of the intervention. As we were also interested in considering within-session variability, we also averaged across sessions to create a variable of average engagement in each dimension for the first 5 min, middle 5 min, and final 5 min of the 15-min sessions.

3.1.3. Step 3: Analyze dimensions of active child engagement

The third step of the protocol is to intensively analyze the dimensions of engagement variables. The overarching goal of this step is to answer the questions of whether children actively engaged in the intervention as intended across dimensions, and how their engagement in each dimension varied within and across sessions. See Row 4 of **Figure 1** (Step 3) for suggestions on how to analyze these variables.

3.1.3.1. Engagement with the activity/materials in the Block Play Intervention

Data was converted to percentages of the calculated proportions for interpretation, but descriptive statistics of the raw data by session are presented in **Table 1**. Children in

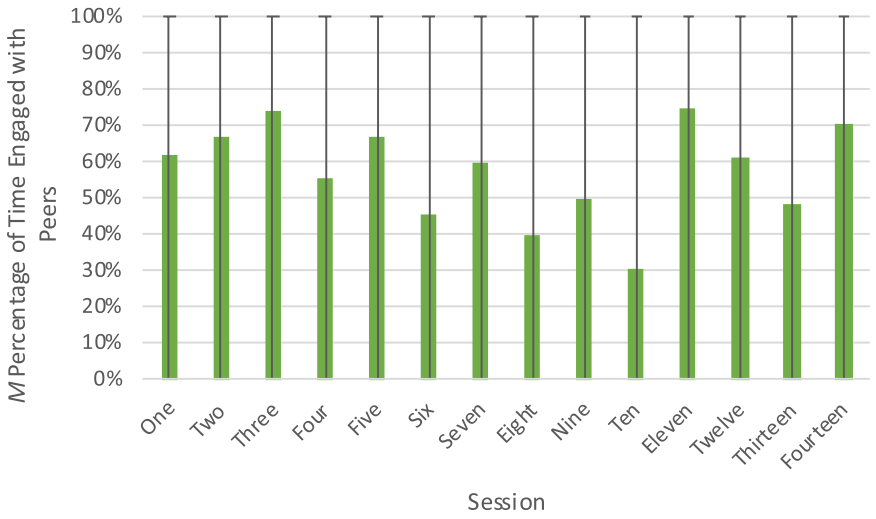


FIGURE 4
Engagement with peers across sessions of the Block Play Intervention. Figure shows percentage of engagement out of total coded time (three time intervals) with peers at each session. The range from minimum to maximum engagement is represented by error bars.

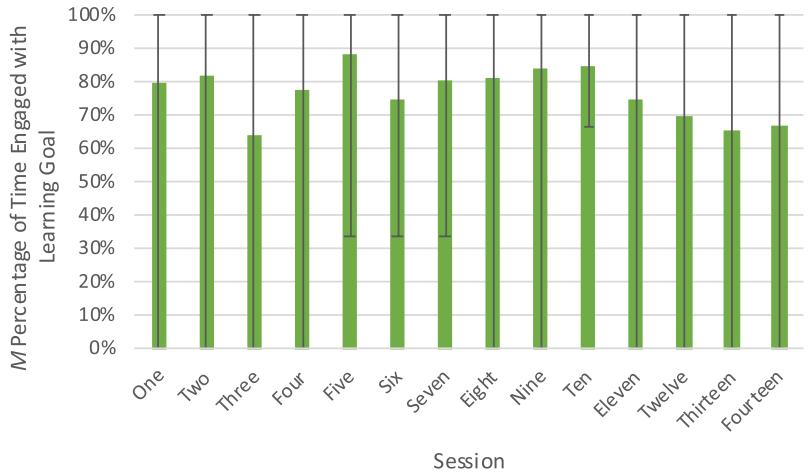


FIGURE 5
Engagement with learning goals across sessions of the Block Play Intervention. Figure shows percentage of engagement out of total coded time (three time intervals) with the learning goal at each session. The range from minimum to maximum engagement is represented by error bars.

the intervention condition were engaged with on-task block building activity 81.0% of the time on average. There was variability between individuals, with the least engaged child engaging in on-task behavior an average of 58.5% of the time and the most engaged child engaging in on-task behavior an average of 95.2% of the time. Examining within-individual variability also provided valuable information about whether the design of the intervention was able to sustain children's engagement in on-task block play over time. Children were fairly consistent with their engagement over the course of the 15-min sessions, although there was some decline; they engaged in on-task behavior 84.4% of the time, on average, in the first 5 min of sessions, 81.4% of the time in the next 5 min, and 73.6% of the time in the last 5 min. There was also some variability in engagement with the activity/materials across sessions (see Figure 3), but there was not an overall decline in engagement with block play as the intervention went on. Overall, our analysis of this dimension suggests that although on-task behavior was high, there was between-individual variability which could potentially influence the strength of efficacy of the intervention. Furthermore, our design of the block play sessions seems to have been successful at keeping children actively engaged in on-task behavior across the intervention sessions.

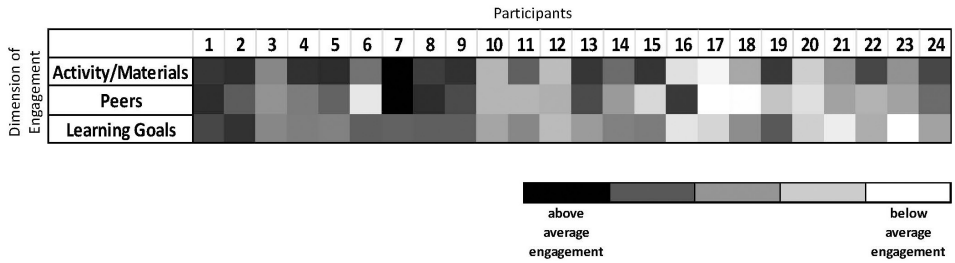


FIGURE 6
Heat map of individual children's multi-dimensional active engagement with the Block Play Intervention. Figure shows individual children's engagement with activity/materials, peers, and learning goals across the Block Play Intervention. Each column represents an individual child, and the rows illustrate that child's engagement with a specific dimension, relative to other children in the sample. Darker shades demonstrate that a child was increasingly above average, compared to the rest of the sample, on a dimension of engagement. Lighter shades demonstrate that a child was increasingly below average on a dimension of engagement.

3.1.3.2. Engagement with peers in the Block Play Intervention

Descriptive statistics of the raw data by session are presented in Table 1. Children were engaged in positive peer interactions an average of 57.0% of the intervals sampled. There was substantial variability between individual children, with the least engaged child only engaged in positive interactions with peers an average of 12.0% of the time, and the most engaged child engaged an average of 97.7% of the time. Children were most engaged in positive interactions with peers near the start of the session (61.5% in the first 5 min, on average), which declined over the course of the session (54.2% in the next 5 min and 50.0% in the last 5 min, on average). There was also considerable variability in children's engagement with peers across sessions (see Figure 4). However, there was not an overall pattern of children positively engaging with peers more or less as the intervention went on.

Overall, analyses of the Engagement with Peers dimension highlighted that despite the prompts for children to work together on their building, children only positively engaged with peers an average of 57% of the sampled intervals. As cooperative play typically emerges between 4 and 5 years of age (Parten, 1932), this level of peer interaction seems developmentally appropriate. Children were more likely to engage positively with peers just after the building prompt was given (in the first 5 min), suggesting that adding scaffolding of peer interactions or reminders to work together throughout the course of the sessions in future iterations of the intervention may increase positive peer engagement.

3.1.3.3 Engagement with learning goals in the Block Play Intervention

Descriptive statistics of the raw data by session are presented in Table 1. Children were engaged with the provided building goal an average of 77.0% of the intervals. There was variability between the least engaged child (39%, on average) and the most engaged child (100%). Again, children's engagement with the goal generally declined over the course of the session, with a drop-off near the end of the sessions (M = 84.3% in first 5 min; 80.0% in second 5 min; 63.9% in the last 5 min). There was some variability in goal engagement across the course of the intervention (see Figure 5). With the exception of one session early on, engagement with the goal was lowest for the most complex building goals, in which children were required to copy a model of a sophisticated structure, given in the final four sessions.

3.1.3.4. Engagement across dimensions in the Block Play Intervention

To demonstrate multi-dimensionality of engagement with the Block Play Intervention for the 24 children in the pilot intervention condition, we standardized the dimensions of engagement variables and have presented them as a heat map in Figure 6. Each column represents an individual child, with increasingly above average engagement in a specific dimension represented as a darkening shade. Below average engagement is indicated with lighter shades. Some children were uniformly engaged or disengaged with the intervention across the dimensions. For example, Participant 1 was above the average of the sample in their engagement with the activity/materials, with peers, and with the learning goal. Conversely, Participant 17 was below average in their engagement across all dimensions. However, other children showed more nuanced engagement with different aspects of the intervention. Participant 16 was highly positively engaged with their peers compared to the rest of the sample but was below average in their engagement with the activity and the goal. Specifically, this child was frequently distracted from building (e.g., walking around the room) and when they did engage in building, they preferred to set their own goals. However, they were highly social with peers, offering positive encouragement and engaging in sharing behaviors during play. Participant 19 was engaged with block play activity in alignment with the provided building goal but was not as positively engaged with peers as the rest of the sample, as this child preferred to build alone. When a peer did attempt to engage with them, the interactions were often negative (e.g.,

refusing to let another child help with the structure). Finally, Participant 13 was highly engaged with block play and with peers compared to the rest of the sample but was slightly below average in their engagement with the provided building goal. These results demonstrate that, even in this small sample, children were differentially engaged with the multiple points of opportunity for engagement. By separating the multiple dimensions of engagement, we can test the most important dimensions of engagement for intervention efficacy in the fourth step of the protocol.

3.1.4. Step 4: Link active child engagement to other measures

In the final step of the protocol, the dimensions of active child engagement should be linked to other variables. First, researchers can explore the assumption that differences in active child engagement can be shaped by contextual and individual factors, like school-level, demographic or baseline skill variables. Next, it is important to link the dimensions of active child engagement to outcome measures to assess how active child engagement may influence intervention effectiveness.

3.1.4.1 Individual predictors of active child engagement in the Block Play Intervention

To consider individual-level predictors of active child engagement with the Block Play Intervention, we estimated correlations among the dimensions of engagement and the demographic variables of age and parent education level, as well as children's baseline social skills and problem behaviors as rated by their classroom teacher using the Social Skills Improvement System rating scale (Gresham and Elliott, 2008; see **Table 2**). Teachers were asked to use the scale to rate children's frequency of behaviors in social skills including communication and cooperation and problem behaviors including hyperactivity/inattention. We also calculated an overall engagement score by summing the standardized scores of engagement across each dimension. Results of these correlations should be interpreted cautiously, given the small sample size. Age was not significantly correlated with any dimension or overall engagement (all p values > 0.05). However, parent education was significantly correlated with overall engagement ($r = 0.54$, $p = 0.01$), engagement with activity/materials ($r = 0.43$, $p = 0.04$), and engagement with peers ($r = 0.48$, $p = 0.02$). Parent education and engagement with the learning goals were not significantly correlated at $r = 0.31$, $p = 0.16$.

Teacher-ratings of children's social skills were also significantly correlated with overall engagement ($r = 0.40$, $p = 0.04$), engagement with the activity/materials ($r = 0.46$, $p = 0.03$) and with the learning goals ($r = 0.60$, $p = 0.002$). Although the correlation between baseline social skills and engagement with peers was not statistically significant, it was positive ($r = 0.11$, $p = 0.57$). Conversely, teacher-rated child problem behaviors were significantly negatively correlated with overall engagement ($r = -0.55$, $p = 0.01$), engagement with the activity/materials ($r = -0.56$, $p = 0.01$), and with goals ($r = -0.76$, $p < 0.001$). Again, the correlation with the engagement with peers dimension was not statistically significant, but was negative ($r = -0.26$, $p = 0.32$). Overall, these analyses revealed interesting insights important for future iterative development of the intervention. First, although previous work found that the Block Play Intervention was particularly efficacious for children whose parents had less educational attainment (Schmitt et al., 2018a) we found that children whose parents have higher educational attainment were more actively engaged with the intervention than their peers. This finding highlights a need to unpack potential reasons and explore additional supports for increasing active child engagement in future iterations as we strive to create equitable interventions appropriate for children from diverse backgrounds. We also found that children who were rated as engaging in more problem behaviors in the classroom by their teacher were less engaged with the intervention, and particularly with the learning goal. Again, in future iterations of the Block Play Intervention, we hope to explore how to scaffold these children's goal-orientation.

3.1.4.2 Linking active child engagement to outcome measures in the Block Play Intervention

Next, we moved to considering how children's overall active engagement and engagement with the activity/materials, peers, and learning goals in the Block Play Intervention related to gains in self-regulation and math outcomes. We calculated pre-to post-gain scores in three outcomes of interest: behavioral self-regulation, measured by the Head-Toes-Knees-Shoulders task (HTKS; McClelland et al., 2014); cognitive flexibility, measured by the Three-Dimensional Change Card Sort Task (DCCS; Zelazo, 2006); and mathematics-specific language, measured by the researcher-developed Math Language Assessment (Purpura and Logan, 2015). See Schmitt et al. (2018a) for assessment details. We estimated correlations among these gain scores and children's active engagement with the activity/materials, peers, and learning goals, as well as overall engagement (see **Table 2**). Four children were missing post-test data on these measures. None of the dimensions of engagement or overall engagement were significantly correlated with the gain scores (all p values > 0.05). Despite this, the pattern of correlation coefficients revealed interesting associations among active child engagement and growth in the outcome variables. For example, overall engagement with the intervention was similarly related to gains in behavioral self-regulation ($r = 0.20$, $p = 0.39$) and cognitive flexibility ($r = 0.19$, $p = 0.44$). However, when the dimensions of engagement variables were considered separately, gains in behavioral self-regulation were similarly related to engagement with the activity/materials ($r = 0.18$, $p = 0.44$), peers ($r = 0.19$, $p = 0.43$), and goals ($r = 0.17$, $p = 0.49$),

TABLE 2 Descriptive statistics and correlations among dimensions of engagement and variables of interest for Step 4 of the protocol from the *Block Play Intervention*.

		Correlations			
	Mean (*SD*)	Engagement with Activity/Materials	Engagement with Peers	Engagement with Learning Goals	Overall engagement
Engagement with Activity/Materials	81% (11.9%)	–			
Engagement with Peers	57% (25.0%)	0.60**	–		
Engagement with Learning Goals	77% (17.3%)	0.77**	0.51**	–	
Overall engagement (*z-composite*)	0 (2.35)	0.89**	0.85**	0.86**	–
Age in months	57.6 (6.32)	0.11	−0.12	0.20	0.05
Parent education	5.61 (2.41)	0.43**	0.48**	0.31	0.54**
Social skills	2.07 (0.36)	0.46**	0.11	0.60**	0.40**
Problem behaviours	0.47 (0.35)	−0.56**	−0.26	−0.76**	−0.55**
Gain in behavioral self-regulation	18.8 (22.3)	0.18	0.19	0.17	0.20
Gain in cognitive flexibility	3.47 (5.18)	0.06	0.39*	−0.03	0.19
Gain in mathematical language	1.6 (1.8)	0.22	0.08	−0.04	0.07

**Indicates $p < 0.05$, *indicates $p < 0.10$; Parent education ranged from 8th grade or less (1) to doctoral degree (9). On average, parents had some college experience. Social skills and problem behaviors were measured by teacher rating scale (SSIS). Behavioral self-regulation was assessed using HTKS task, cognitive flexibility was assessed using DCCS task, and mathematical language was assessed using the math language assessment.

but gains in cognitive flexibility were more related to active child engagement with peers ($r = 0.39$, $p = 0.09$) than with the activity/materials ($r = 0.06$, $p = 0.71$) or goals dimensions ($r = -0.03$, $p = 0.91$). Gains in mathematics-specific language may also be more related to one dimension of engagement with the intervention (activity/materials, $r = 0.22$, $p = 0.35$) than the others (peers, $r = 0.08$, $p = 0.75$; goals, $r = -0.04$, $p = 0.74$). Although these correlation analyses are underpowered, the pattern of results supports our theoretical assumptions that active child engagement with an intervention is multi-dimensional, and engagement with specific dimensions may be particularly important for supporting different outcomes.

In larger samples, along with traditional measures of fidelity (e.g., dosage and adherence) the separate dimensions of engagement should be considered as moderators of treatment effects on outcome measures. For example, in future iterations of the Block Play Intervention, we plan to test dimensions of engagement as moderators of the effect of condition. We will also test interactions among these dimensions and hypothesize that children who are highly engaged across dimensions will benefit the most from the intervention.

4. Conclusion and future directions

Despite theoretical models that include multiple factors of participant responsiveness, the crucial aspect of active participant engagement is often overlooked, especially for young child participants. Some early SEL interventions (e.g., Nesbitt and Farran, 2021) have mixed evidence of efficacy but measuring active engagement with an intervention at the child level may help the field to unpack these mixed results. Furthermore, the multi-dimensionality and variability of children's active engagement with the intervention must be considered to capture children's experiences in the complex intervention environment. In this article, we have presented a four-step protocol for identifying, operationalizing, and analyzing children's active engagement with multiple dimensions of an intervention. We encourage fellow researchers to prioritize incorporating participant engagement at the child level into the measurement

of implementation fidelity in SEL interventions with young children and conclude by presenting suggestions to further this work. Although this protocol has been initially applied to a pilot study of a researcher-implemented intervention with a small sample size, we believe it can be scaled for use with larger and more comprehensive intervention studies.

In the fourth step of the protocol, we suggested that future work using this protocol may wish to explore contextual factors that predict the dimensions of engagement. For example, research specifically focused on supporting SEL through play-based learning would benefit from the examination of play as it is defined within different cultural contexts. Children's play-based learning experiences are scaffolded by their culture and given the impact of globalization, successful interventions would require consideration of the different cultural belief systems of the children and families included in the study (Harkness and Super, 1993). Children have different exposure to play activities within the home environment which affects the variety of play experiences that children are exposed to and possibly their level of active engagement in various dimensions of an intervention activity. Future work is also needed to consider how individual children's engagement may affect peers' engagement, especially during group work.

Our example, the Block Play Intervention, did not include an active control condition, but researchers that include a closely aligned active control condition may wish to code active child engagement with the active control condition. For example, in another iteration of the Block Play Intervention, we plan to test the efficacy of the semi-structured play sessions against a free block play condition, in which children are not given specific building goals. We will code and compare engagement with the activity and with peers to compare to the semi-structured condition to test whether there are differences in engagement by condition and if individual differences in engagement relate to gains in the target outcomes.

We focused this article on child participants' engagement in interventions, but future work is needed to consider the bi-directional relations between interventionist and child-directed factors of implementation. For example, theoretical models posit that interventionist-directed behaviors influence participants' responsiveness to an intervention (e.g., a cascade model; Berkel et al., 2018). However, child engagement across the dimensions likely also influences interventionist behaviors (Berkel et al., 2011). That is, interventionist adherence to a script and quality of delivery is likely influenced by children's engagement in the activity at hand, as the interventionist responds to the children's behavior. The Block Play Intervention was a researcher-implemented intervention, but researchers testing teacher-implemented SEL interventions may also include measures of active teacher participant engagement to consider relations among teacher and child engagement variables. Untangling and determining the best ways to capture the interaction between these factors is an important next step for applying integrated theoretical models of program implementation to practice.

Finally, protocols are also needed to guide the nuanced measurement and analysis of other understudied aspects of fidelity. For example, quality of the delivery of an intervention also includes multiple dimensions and is varied between and within interventionists. Intervention work with school-aged children has explored some of the dimensions of this domain (e.g., practice opportunities, modeling, feedback, scaffolding; Doabler et al., 2021). Additional work is needed to unpack quality in interventions designed for young children and to create general protocols for guiding the identification, operationalization, and analyses of these variables.

In summary, we hope to encourage SEL intervention researchers to consider active child engagement as a worthwhile area of focus in the measurement of implementation fidelity. We have introduced the four-step protocol as a general guide for capturing dimensions of active child engagement. In doing so, the field may be better able to discover which early SEL interventions work, under which conditions, and for whom.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Purdue University Institutional Review Board. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

BD conceptualized the protocol, conducted analyses, interpretation of data, and led writing of manuscript. BD, TP, EG, and LB wrote the manuscript. TZ, IK, and KM conducted data coding and analysis. SS conceived of original study, directed data collection, provided input on conceptualization of protocol, and interpretation of data. All authors provided critical feedback and approved the submitted version.

Funding

This work was supported by the National Science Foundation, Award #2000641.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Barnes, S. P., Bailey, R., and Jones, S. M. (2021). Evaluating the impact of a targeted approach designed to build executive function skills: A randomized trial of brain games. *Front. Psychol.* 12:655246. doi: 10.3389/fpsyg.2021.655246

Berkel, C., Mauricio, A. M., Sandler, I. N., Wolchik, S. A., Gallo, C. G., and Brown, C. H. (2018). The cascading effects of multiple dimensions of implementation on program outcomes: A test of a theoretical model. *Prev. Sci.* 19, 782–794. doi: 10.1007/s11121-017-0855-4

Berkel, C., Mauricio, A. M., Schoenfelder, E., and Sandler, I. N. (2011). Putting the pieces together: An integrated model of program implementation. *Prev. Sci.* 12, 23–33. doi: 10.1007/s11121-010-0186-1

Bos, S. E., Powell, S. R., Maddox, S. A., and Doabler, C. T. (2022). A synthesis of the conceptualization and measurement of implementation fidelity in mathematics intervention research. *J. Learn. Disabil.* 222194211065498. doi: 10.1177/00222194211065498

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., and Balain, S. (2007). A conceptual framework for implementation fidelity. *Implement. Sci.* 2:40. doi: 10.1186/1748-5908-2-40

CASEL (2017). *What is the CASEL framework?* Chicago, IL: CASEL.

Century, J., Rudnick, M., and Freeman, C. (2010). A framework for measuring fidelity of implementation: A foundation for shared language and accumulation of knowledge. *Am. J. Eval.* 31, 199–218. doi: 10.1177/1098214010366173

Dane, A. V., and Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clin. Psychol. Rev.* 18, 23–45. doi: 10.1016/S0272-7358(97)00043-3

Darrow, C. L. (2013). The effectiveness and precision of intervention fidelity measures in preschool intervention research. *Early Educ. Dev.* 24, 1137–1160. doi: 10.1080/10409289.2013.765786

Denham, S. A., Bassett, H. H., Thayer, S. K., Mincic, M. S., Sirotkin, Y. S., and Zinsser, K. (2012). Observing preschoolers' social-emotional behavior: Structure, foundations, and prediction of early school success. *J. Genet. Psychol.* 173, 246–278. doi: 10.1080/00221325.2011.597457

Doabler, C. T., Clarke, B., Kosty, D., Fien, H., Smolkowski, K., Liu, M., et al. (2021). Measuring the quantity and quality of explicit instructional interactions in an empirically validated tier 2 kindergarten mathematics intervention. *Learn. Disabil. Q.* 44, 50–62. doi: 10.1177/0731948719884921

Domitrovich, C. E., Cortes, R. C., and Greenberg, M. T. (2007). Improving young children's social and emotional competence: A randomized trial of the preschool "PATHS" curriculum. *J. Prim. Prev.* 28, 67–91. doi: 10.1007/s10935-007-0081-0

Downer, J. T., Booren, L. M., Lima, O. K., Luckner, A. E., and Pianta, R. C. (2010). The individualized classroom assessment scoring system (inCLASS): Preliminary reliability and validity of a system for observing preschoolers' competence in classroom interactions. *Early Child. Res. Q.* 25, 1–16. doi: 10.1016/j.ecresq.2009.08.004

Farran, D. C. (2011). *Child observation in preschool (With adaptations for use in the Tools of the Mind curriculum evaluation)*. Nashville, TN: Peabody Research Institute, Vanderbilt University.

Fullan, M. (2001). *The new meaning of educational change*. London: Routledge.

Garon, N., Bryson, S. E., and Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychol. Bull.* 134, 31–60. doi: 10.1037/0033-2909.134.1.31

Gashaj, V., Oberer, N., Mast, F. W., and Roebers, C. M. (2019). The relation between executive functions, fine motor skills, and basic numerical skills and their relevance for later mathematics achievement. *Early Educ. Dev.* 30, 913–926. doi: 10.1080/10409289.2018.1539556

Gresham, F. M., and Elliott, S. N. (2008). *Social skills improvement system: Rating scales*. London: Pearson Assessments.

Hadani, H. S., and Rood, E. (2018). *The roots of STEM success: Changing early learning experiences to build lifelong thinking skills*. San Francisco, CA: Center for Childhood Creativity.

Harkness, S., and Super, C. M. (1993). "The developmental niche: Implications for children's literacy development," in *Early intervention and culture: Preparation for literacy*, eds L. Eldering and P. Lesemen (Paris: UNESCO), 115–132.

Howes, C., Guerra, A. W., Fuligni, A., Zucker, E., Lee, L., Obregon, N. B., et al. (2011). Classroom dimensions predict early peer interaction when children are diverse in ethnicity, race, and home language. *Early Child. Res. Q.* 26, 399–408.

Hulleman, C. S., Rimm-Kaufman, S. E., and Abry, T. (2013). "Innovative methodologies to explore implementation: Whole-part-whole—construct validity, measurement, and analytical issues for intervention fidelity assessment in education research," in *Applying implementation science in early childhood programs and systems*, eds T. Halle, A. Metz, and I. Martinez-Beck (Baltimore, MD: Paul H Brookes Publishing), 65–93.

Ling, S. M., and Barnett, D. W. (2013). Increasing preschool student engagement during group learning activities using a group contingency. *Topics Early Child. Spec. Educ.* 33, 186–196. doi: 10.1177/0271121413484595

Mackintosh, B. B., and Rowe, M. (2021). Baseline inequalities: Social skills at preschool entry moderate math development. *J. Res. Child. Educ.* 35, 1–21. doi: 10.1080/02568543.2020.1728446

Massouleh, N. S., Branch, R., and Jooneghani, R. B. (2012). Learner-centered instruction: A critical perspective. *J. Educ. Pract.* 3, 50–59.

McClelland, M. M., Acock, A. C., Piccinin, A., Rhea, S. A., and Stallings, M. C. (2013). Relations between preschool attention span-persistence and age 25 educational outcomes. *Early Child. Res. Q.* 28, 314–324. doi: 10.1016/j.ecresq.2012.07.008

McClelland, M. M., and Cameron, C. E. (2019). Developing together: The role of executive function and motor skills in children's early academic lives. *Early Child. Res. Q.* 46, 142–151.

McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., and Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Dev. Psychol.* 43, 947–959. doi: 10.1037/0012-1649.43.4.947

McClelland, M. M., Cameron, C. E., Duncan, R., Bowles, R. P., Acock, A. C., Miao, A., et al. (2014). Predictors of early growth in academic achievement: The head-toes-knees-shoulders task. *Front. Psychol.* 5:599. doi: 10.3389/fpsyg.2014.00599

McClelland, M. M., Tominey, S. L., Schmitt, S. A., and Duncan, R. (2017). SEL interventions in early childhood. *Future Child.* 27, 33–47.

McClelland, M. M., Tominey, S. L., Schmitt, S. A., Hatfield, B. E., Purpura, D. J., Gonzales, C. R., et al. (2019). Red light, purple light! Results of an intervention to promote school readiness for children from low-income backgrounds. *Front. Psychol.* 10:2365. doi: 10.3389/fpsyg.2019.02365

Nesbitt, K. T., and Farran, D. C. (2021). Effects of prekindergarten curricula: Tools of the mind as a case study. *Monogr. Soc. Res. Child Dev.* 86, 7–119.

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Rev. Educ. Res.* 78, 33–84.

Parten, M. B. (1932). Social participation among pre-school children. *J. Abnorm. Soc. Psychol.* 27, 243–269. doi: 10.1037/h0074524

Purpura, D. J., and Logan, J. A. R. (2015). The nonlinear relations of the approximate number system and mathematical language to early mathematics development. *Dev. Psychol.* 51, 1717–1724. doi: 10.1037/dev0000055

Schmitt, S. A., Geldhof, G. J., Purpura, D. J., Duncan, R., and McClelland, M. M. (2017). Examining the relations between executive function, math, and literacy

during the transition to kindergarten: A multi-analytic approach. *J. Educ. Psychol.* 109, 1120–1140. doi: 10.1037/edu0000193

Schmitt, S. A., Korucu, I., Napoli, A. R., Bryant, L. M., and Purpura, D. J. (2018a). Using block play to enhance preschool children's mathematics and executive functioning: A randomized controlled trial. *Early Child. Res. Q.* 44, 181–191. doi: 10.1016/j.ecresq.2018.04.006

Schmitt, S. A., Lewis, K. M., Duncan, R. J., Korucu, I., and Napoli, A. R. (2018b). The effects of positive action on preschoolers' social–emotional competence and health behaviors. *Early Child. Educ. J.* 46, 141–151. doi: 10.1080/10409289.2013.825565

Schmitt, S. A., McClelland, M. M., Tominey, S. L., and Acock, A. C. (2015). Strengthening school readiness for head start children: Evaluation of a self-regulation intervention. *Early Child. Res. Q.* 30, 20–31. doi: 10.1016/j.ecresq.2014.08.001

Siegler, R. S. (1994). Cognitive variability: A key to understanding cognitive development. *Curr. Dir. Psychol. Sci.* 3, 1–5. doi: 10.1111/1467-8721.ep10769817

Siegler, R. S. (2007). Cognitive variability. *Dev. Sci.* 10, 104–109. doi: 10.1111/j.1467-7687.2007.00571.x

Sluss, D. J., and Stremmel, A. J. (2004). A sociocultural investigation of the effects of peer interaction on play. *J. Res. Child. Educ.* 18, 293–305. doi: 10.1080/02568540409595042

Tominey, S. L., and McClelland, M. M. (2011). Red light, purple light: Findings from a randomized trial using circle time games to improve behavioral self-regulation in preschool. *Early Educ. Dev.* 22, 489–519. doi: 10.1080/10409289.2011.574258

Wolfgang, C. H., Stannard, L. L., and Jones, I. (2001). Block play performance among preschoolers as a predictor of later school achievement in mathematics. *J. Res. Child. Educ.* 15, 173–180. doi: 10.1080/02568540109594958

Zelazo, P. D. (2006). The dimensional change card sort (DCCS): A method of assessing executive function in children. *Nat. Protoc.* 1, 297–301.

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 09 January 2023
DOI 10.3389/feduc.2022.1044730

OPEN ACCESS

EDITED BY
Stephanie M. Jones,
Harvard University,
United States

REVIEWED BY
Thandi Sule,
Oakland University,
United States
Gretchen Brion-Meisels,
Graduate School of Education and
Harvard University,
United States
Robert Jagers,
Collaborative for Academic Social and
Emotional Learning (CASEL), United States

*CORRESPONDENCE
Meiko Lin
✉ research@liberatedsel.com

†These authors have contributed equally to this work and share first authorship
‡These authors have contributed equally to this work and share last authorship

SPECIALTY SECTION
This article was submitted to Educational Psychology, a section of the journal Frontiers in Education

RECEIVED 15 September 2022
ACCEPTED 15 December 2022
PUBLISHED 09 January 2023

CITATION
Lin M, Olsen S, Simmons DN, Miller M and Tominey SL (2023) "Not try to save them or ask them to breathe through their oppression": Educator perceptions and the need for a human-centered, liberatory approach to social and emotional learning. *Front. Educ.* 7:1044730. doi: 10.3389/feduc.2022.1044730

"Not try to save them or ask them to breathe through their oppression": Educator perceptions and the need for a human-centered, liberatory approach to social and emotional learning

Meiko Lin[1,2*†], Svea Olsen[1,3†], Dena N. Simmons[1,2†], Miriam Miller[1,4‡] and Shauna L. Tominey[1,3‡]

[1]LiberatED, Oakland, CA, United States, [2]Institute for Racial Justice, Loyola University Chicago, Chicago, IL, United States, [3]College of Public Health and Human Sciences, Oregon State University, Corvallis, OR, United States, [4]Department of Educational and Counselling Psychology, and Special Education, University of British Columbia, Vancouver, BC, Canada

Introduction: Social and emotional learning (SEL) has been identified as one approach to promote positive mental health outcomes while alleviating the stressors of systemic racism and a global pandemic. As the United States turns to SEL as a remedy for mental health challenges and the current civil unrest, it becomes increasingly relevant to understand what SEL means to those who use it the most to strengthen the implementation of current programs as well as to inform the development of new programs to fill existing gaps.

Methods: This abductive qualitative study expands prior research by exploring how in-service educators define SEL (N = 427).

Results: Our findings highlight that educators perceive SEL as more expansive than current competency-based models. Educators describe SEL as a praxis that can be responsive to student and community needs, facilitate healing, and center humanity along with racial and social justice.

Discussion: We discuss implications that highlight the potential risks and harm that can be perpetuated by the current practice of SEL and, like the educators in our study, advocate for dismantling white supremacy structures in education through the co-creation of a humanizing SEL approach.

KEYWORDS

social and emotional learning, SEL, education, qualitative research, liberatory pedagogy, human-centered education, racial justice, community-based research

Introduction

The COVID-19 pandemic sparked a sense of urgency for the nation to address existing and new mental health needs. The lives of many young people were upended with school closures and reopenings (Youth Truth, 2021), inconsistent messages about pandemic protocols, and, for some, illness or loss of loved ones. As a result, many youth experienced heightened stress, anxiety, and depression (Fair Health, 2021; Mott, 2021; Varma et al., 2021). Alongside the pandemic, there was increased focus on racist violence (Curtis et al., 2021), which contributed to additional anxiety, fear, and angst. The coronavirus pandemic and civil unrest contributed to more pronounced racial tension (Barret, 2022). Youth who already experienced systemic marginalization [e.g., Black, Indigenous, and People of Color (BIPOC), youth in poverty, students with disabilities, or lesbian, gay, bisexual, transgender, and queer (LGBTQ) youth] were particularly vulnerable to mental health challenges and pandemic-related stressors (OECD, 2020; Reliefweb, 2020; Vasquez Reyes, 2020). It is worth noting that some young people did better in the face of these challenges. For instance, youth who were removed from oppressive school contexts no longer experienced daily exposure to hostility, discrimination, and microaggressions to the same degree (Miller, 2021). Youth in supportive homes developed a stronger sense of identity and critical consciousness during the periods of remote learning (Miller, 2021). For youth to thrive, during a pandemic or otherwise, they need to learn in environments that support their whole selves, promote their well-being, and are free from harm.

To respond to the challenges of both the pandemic and civil unrest, there were calls for schools to bolster commitments to mental health and diversity, equity, and belonging (Jones et al., 2022; Rutgers Center for Effective School Practices, 2022). Many school districts and policymakers advocated for social and emotional learning (SEL) programming, which aims to foster life skills that support people in experiencing, managing, and expressing emotions meaningfully, making sound decisions, and fostering rewarding interpersonal relationships (Modan, 2020; Sanders, 2020). The Office of Child Care Initiative to Improve Social–Emotional Wellness of Children published a guide recommending SEL as a strategy to help meet the needs of students (Childcare Technical Assistance Network, 2021). Thirty-eight states referenced SEL in their response plans to the pandemic (Yoder et al., 2020). With an uptick of attention to SEL, many prominent SEL organizations and programs including the Collaborative for Academic, Social, and Emotional Learning (CASEL) and the Social and Emotional Learning Alliance for the United States (SEL4US) responded by producing webinars, online courses, and resources for educators. CASEL produced a roadmap to infuse SEL and mental health promotion as schools reopened amidst COVID-19 [Collaborative for Academic, Social, and Emotional Learning (CASEL), 2020a,b].

Although the pandemic heightened awareness of SEL's importance, SEL had already been gaining traction among school communities for decades because of growing evidence of its short- and long-term benefits for students including enhanced relational skills and attitudes, improved academic achievement, and reduced anxiety, stress, and depression (Durlak et al., 2011; Taylor et al., 2017). Generally, SEL programs are designed to be implemented as a complement to other school curricula by teachers trained in their use. Schools can choose from many SEL programs, most of which generally share the goal of enhancing a range of social and emotional skills (e.g., understanding one's own and other's emotions). Many SEL programs follow a widely-adopted SEL framework proposed by CASEL, which categorizes SEL into five competencies: self-awareness, self-management, social-awareness, relationship skills, and responsible decision-making (Lawson et al., 2019; Frye et al., 2022).

Founded in 1994, CASEL published the first major book on school-based SEL programming in which they identified a research-based framework for implementing effective programs for building SEL (ASCD, 1997). The framework eventually evolved into the CASEL 5 SEL Competencies Framework (hereafter referred to as the CASEL 5) and has been widely adopted across the SEL field and schools across the nation. For example, 20 large school districts, serving 1.7 million students, have used the CASEL 5 to establish preschool to high school SEL standards that articulate what students should know and be able to do [Collaborative for Academic, Social, and Emotional Learning (CASEL), 2021]. The CASEL 5 has also been used to guide the development and evaluation of many school-based SEL approaches and research, thus shaping educational practice [Collaborative for Academic, Social, and Emotional Learning (CASEL), 2013a,b].

Although the evidence supporting the effectiveness of SEL programs continues to grow, there has been some critique that SEL alone is not enough to address the current educational and societal climate without a commitment to confront and eradicate racist practices, policies, and curricula (Simmons, 2021). Some argue that SEL without a racial justice lens could contribute to the continued harm and dehumanization of our nation's BIPOC students, who continue to be systemically and institutionally oppressed (Camangian and Cariaga, 2021; Simmons, 2021). As SEL programs have been identified as falling short in addressing social and racial justice, scholars have criticized SEL for perpetuating and upholding systems of oppression and contributing to harmful narratives about the need to "fix" BIPOC youth (e.g., DeMartino et al., 2022; Miller et al., 2022). Given the education system in the United States (U.S.) is based on tenets of white supremacy (Brooks and Theoharis, 2018) and that our nation's schools serve mostly BIPOC students (Riser-Kositsky, 2019), SEL faces the risk of becoming "white supremacy with a hug" (Madda, 2019; Simmons, 2021) without dedicated efforts to combat racial and social injustice (Simmons, 2019b). There is an urgent need for antiracist, culturally affirming, and responsive SEL that centers educator and student voices (Abolitionist Teaching Network, 2022; DeMartino et al., 2022). Calls for fearless SEL (Simmons, 2019b) and transformative SEL (Jagers et al., 2019) have pushed the SEL field to be more equity-responsive.

While SEL holds promise for improving mental health outcomes among youth (e.g., Weare, 2017; Grové and Laletas, 2020), SEL programs may not always be implemented effectively for all students in the current sociopolitical context of the United States (Forman et al., 2022; Leonard and Woodland, 2022). Many SEL programs have small to moderate effects with effectiveness varying widely across contexts and populations (Durlak et al., 2011; Boncu et al., 2017). Even the use of the same program can lead to variability in outcomes depending on the fidelity of implementation as well as the unique needs of the students and educators (e.g., Hunter et al., 2022). Few programs have collected data on what is working, for whom, and in what contexts. Fewer still have explored in-depth fidelity data to understand differential effects although research in this area is growing (Thayer et al., 2019). As the nation turns to SEL to support young people's mental health in the midst of a pandemic and social unrest, it is imperative to understand what SEL means to those who use it most to strengthen its implementation and to inform the development of new programs and practices that aim to fill existing gaps.

Educators and other school personnel have often been left out of the creation of widely-used SEL approaches. Thus, exploring how educators and other school staff define and practice SEL is an important step toward understanding how best to support students in a way that honors their lived experiences and identities while recognizing their needs. This qualitative study fills this gap and expands prior research by exploring how educators define SEL. We believe that educators are important partners in the creation of curricula, student learning experiences, and the classroom environment, and that their perspectives are vital not only to their students, but also to the education field. This study provides an opportunity to understand how educators' perspectives of SEL align with predominant SEL conceptualizations and program implementation.

Research design

We approached this research based on the insight of Indigenous scholars, who argue for centering and honoring the wisdom of communities of study and for relational accountability between researchers and study participants (Tuck and Yang, 2012; Wilson, 2020). Relational accountability requires that researchers engage in respectful and mutually beneficial relationships with study participants (Wilson, 2020). Thus, we employed a community-based research design both to honor and respect our study participants and to gain a grassroots perspective of learning environments, where SEL facilitates belonging, healing, care, and justice. The goal of community-based research is to educate, improve practices, and bring about social change (Atalay, 2010; Tremblay et al., 2018). In the context of community-based research, "community" is not a geographic location but rather a community of interest or a collective identity with shared goals, interests, or problems (Alinsky, 1971; Israel et al., 2005). Within education, the historical roots of community-based research can be traced to critical pedagogy (Boyd, 2020). Two major proponents of critical pedagogy, John Dewey and Paolo Freire, argued that meaningful learning may lead to social change through repeated critical analysis, reflections, and actions (Freire, 1970; Peterson 2009). The work of Michel Foucault and Thomas Kuhn further informed the development of community-based research by challenging the "hierarchization of knowledge" (Ritzer, 1996, p. 463). Specifically, they raised questions of *how we know what we know* and *what it is that we value as knowledge* (Wicks et al., 2008). Their work shaped the development of community-based research to reflect "a democratization of the research process and a validation of multiple forms of knowledge, expertise, methodologies" (Boyd, 2020, p. 750).

Our research aim was to understand what SEL means to in-service educators by exploring our overarching research question: "how do educators conceptualize SEL?" In the context of this study, we aimed to address the research question through recognizing the collective voices and knowledge of educators. In line with community-based research, we conducted reflexivity practices to acknowledge how our identities, values, experiences, and attitudes influence our research (Reinharz, 1992; Israel et al., 1998). Our research team included individuals with intersectional racial, cultural, and sexual identities. Our lived experiences as students, teachers, and researchers shape our research. We practiced self- and group-reflexivity to recognize the complexities that our identities, values, and lenses posed in relation to this study. As a research team, we routinely challenged one another around the beliefs and interpretations we ascribed to the educators participating in this study.

For this study, we chose an abductive qualitative methodology, ensuring that educators' voices were at the forefront of our research while also acknowledging the prevalence of an existing definition of SEL. Qualitative analysis enables researchers to construct understanding entirely from the data without any preconceived ideas of what they may find (Creswell and Poth, 2016). An abductive approach involves the systematic combination of both inductive and deductive methodologies (Dubois and Gadde, 2002). In this study, though we used the CASEL 5 as the foundation to build our initial codebook, we did not influence or force the data to fit into it or other preconceived conceptualizations.

Materials and methods

Participants and procedures

We collected data at a free virtual event in the summer of 2021 that was focused on the intersection of SEL, racial justice, and healing in educational settings. We chose participants at this event for our study because they likely shared a dedication

to learning about and/or implementing and centering these topics in their educational practice. To register for the event, participants completed an online survey comprising structured-response and open-ended questions. Participants were 918 registrants who completed the registration survey. The survey took approximately 10–20 min to complete and comprised 12 structured-response questions (e.g., demographics, SEL curriculum implemented at their school) and 7 open-ended questions (e.g., participants' perceptions of SEL). We limited our sample to only include participants who had responded to the one open-ended question related to SEL, "What does SEL mean to you?" We additionally limited our sample to those who were currently working in a pre-K to 12th-grade school setting [e.g., classroom teachers, school psychologists, school administrators, etc. (see Table 1 for a complete list) referred to as 'educators' hereafter] in the United States ($N = 427$). As illustrated in Table 1, participants' racial backgrounds included white (54%), Black/African American (29%), Latinx (14%), and Asian (8%). The majority of respondents were female (79%), followed by male (10%), and transgender, nonbinary, gender non-conforming, or androgynous (3%). Most respondents were 35–44 years old (range: 18–74). Ninety-two percent of participants reported they implemented an SEL curriculum at their school (see Table 2).

TABLE 1 Sociodemographic characteristics of participants ($N=427$).

Demographic characteristics	*n*	%	Examples of self-identification
Age[a]			
18–24	6	1.4	
25–34	75	17.6	
35–44	180	42.2	
45–54	124	29.1	
55–64	39	9.1	
65–74	2	0.5	
Race/ethnicity[b]			
Self-describe*	16	3.8	*Puerto Rican, Haitian by marriage, Lebanese American, West Indian American, Middle Eastern
American Indian or Alaska Native	7	1.6	
Asian	33	7.7	
Black or African American	124	29.1	
Latinx and/or Hispanic (non-white)	33	7.7	
Latinx and/or Hispanic (white)	25	5.9	
Native Hawaiian or Pacific Islander	2	0.5	
White	231	54.2	
Gender[c]			
Self-describe**	3	0.7	**Gender non-conforming, androgynous
Man	43	10.1	
Woman	337	79.1	
Non-binary	6	1.4	
Trans	2	0.5	
Role[d]			
Assistant school leader (assistant principal, assistant director, vice principal, and assistant head)	42	9.8	***Librarian, Trainer, Curriculum Manager, Instructional Coach
Paraprofessional/school aide	1	0.2	
Parent	15	3.5	
School counselor	31	7.3	
School leader (principal, director, head)	1	0.2	
School psychologist	5	1.2	
Social worker	15	3.5	
Subject area or content specialist/coach	59	13.8	

(Continued)

Analyses

Two coders who are authors of this manuscript conducted the analyses for this study. To promote and attain researcher reflexivity, we wrote personal reflexivity statements about how our lived experiences and social positioning might influence how we understood and interpreted survey responses before coding (Birks et al., 2008; Creswell and Poth, 2016). The first coder is a Chinese-Canadian, cisgender, straight woman who previously taught at preschools in Taiwan and Japan. The second coder is a white, cisgender, queer woman with experience working with immigrant and refugee families and living in Spain, the Dominican Republic, and Australia. Both coders previously held research positions in an academic research center focused on social and emotional learning in educational settings.

Our first step to analysis was to create *a priori* codes based on the CASEL 5. We then randomly selected and reviewed 10 responses to familiarize ourselves with the type of data and develop a preliminary codebook (Miles and Huberman, 1994). We separately coded these 10 responses using the *a priori* codes and inductively generated new codes as they arose in the data. We wrote memos noting which *a priori* codes we applied, any new codes we generated, and questions or comments that arose during the review process. We then met to discuss our processes, reconcile discrepancies in our coding, and come to a consensus on an updated codebook. Once we completed this

TABLE 1 (Continued)

Demographic characteristics	*n*	%	Examples of self-identification
Superintendent	3	0.7	
Teacher	166	38.9	
Other***	96	22.5	

[a]Percentages for age add up to less than 100% because one response is missing.
[b]Percentages for race/ethnicity do not add up to 100 because it is a multi-select question.
[c]Percentages for gender do not add up to 100 because it is a multi-select question.
[d]Percentages for role do not add up to 100 because it is a multi-select question.
* Provides some sample self-describe race/ethnicity in the Example of self-identification column.
** Provides some sample self-describe gender in the Example of self-identification column.
*** Provides some sample other [education] roles in the Example of self-identification column.

TABLE 2 Social and emotional learning curriculums used.

Demographic characteristics	*n*	%	Examples of self-identification
4R's	13	3.1	
Character first	2	0.5	
Good behavior game	6	1.4	
Lion's quest	3	0.7	
MindUP	24	5.6	
None	25	5.9	
Open circle	12	2.8	Mosaic, cool tools, mindful schools, conscious discipline, restorative practices, toolbox
Other	129	30.3	
PATHS	4	0.9	
PBIS	195	45.8	
Positive action	7	1.6	
RULER	80	18.8	
Sanford harmony	40	9.4	
School/self-designed	129	30.3	
Second step	101	23.7	
Zones of regulation	97	22.8	

Percentages do not equal 100% because some responses are missing and it is a multi-select question.

stage, we compiled the ten responses that we used for review with all other data so that these responses would also be coded during analysis.

Afterward, we conducted a thematic analysis, a process of identifying, analyzing, organizing, and reporting patterns within data resulting in emergent themes that are then refined and interpreted (Creswell and Poth, 2016). We began our thematic analysis of all data using Atlas.ti (Version 22) following the steps identified by Braun and Clarke (2006). We open-coded all data, conducting line-by-line analysis of all responses to create codes (Braun and Clarke, 2006). We derived codes from key ideas, quotes, and words that reflected participants' perceptions. Codes were broad, multiple, and overlapping. To substantiate the reliability of the coding process, we double-coded 20 % of the data in four rounds (Seale, 1997; Bauer and Gaskell, 2000; Yardley, 2000; O'Brien et al., 2014); the primary coder coded approximately 100 responses per round and the secondary coder coded approximately 20 responses per round.

Following recommendations from Birks et al. (2008), we independently wrote memos after each coding round to note any difficult coding decisions, new codes, potential themes, and any biases that we may have brought into the process. Memos are self-reflections of a researcher's thoughts and insights that explicitly acknowledge the subjective influences of the researcher and promote and attain researcher reflexivity (Birks et al., 2008; Creswell and Poth, 2016). We met after each round of coding. In these meetings, we merged double-coded responses in Atlas.ti, reconciled any coding discrepancies to reach 100% consensus, and identified any updates to the codebook. Before reconciling, our average Krippendorf's alpha was 0.76 with all but the first coding round at or above 0.80, the cutoff threshold for representing good intercoder reliability (O'Connor and Joffe, 2020). We wrote memos during reconciliation to create an audit trail recording how we made decisions and reached conclusions throughout the research process (Birks et al., 2008; Speziale et al., 2011).

After our final reconciliation meeting, we went back through all previously coded data to account for changes in the codebook resulting from previous reconciliation meetings. In the final stage of coding, we clustered codes and identified groups of codes representing similar underlying constructs. These clusters became emerging themes, which we compared within and across responses, reviewing and refining themes in an iterative process to ensure they accurately represented coded material. We then shared our emergent themes with another author on this study whose expertise and lived experience as a Black, queer, cisgender woman, former middle school teacher, and current teacher educator and SEL practitioner and researcher provided important insight into refining our themes. We discussed and further clarified our themes based on this researcher's feedback.

Findings

Three themes emerged related to how educators define SEL. Educators described SEL as "developing skills and competencies" (Theme 1), which aligns with existing SEL definitions. In addition, however, educators conceptualized SEL in ways that are not part of current SEL definitions. Some educators mentioned how SEL, as it is currently taught, can be harmful to BIPOC students and a superficial solution to deeper problems such as racism and other social injustices. Educators expressed that SEL should "not try to save them [students] or ask them to breathe through their oppression" (Theme 2). Educators also highlighted "focusing on a child's well-being" as part of SEL (Theme 3). This theme encompassed several subthemes related to how SEL could be taught and implemented. Subthemes included honoring identity, centering humanity and the whole child,

promoting healing and liberation, and advancing social justice. We present the frequency of codes for each theme in Appendix A.

Theme 1: "Developing skills and competencies"

Many educators in our study defined SEL as competency-based. Specifically, several educators mentioned that, for them, SEL was synonymous with the CASEL 5. In fact, CASEL was the only SEL framework educators referenced by name. One educator wrote that thinking of SEL *"evokes [the] CASEL framework."* Educators found the competency-based definition of SEL helpful and were appreciative of it. They perceived that a framework such as CASEL's outlines "*[the] key skills and competencies that lead all human beings to know themselves,*" and that SEL skills are important in "*developing one's understanding of oneself.*"

Some educators did not refer to the CASEL 5 framework in its entirety but still defined SEL using one or more of the CASEL 5 competencies. For example, educators conceptualized SEL as skills that help them "*identify, recognize, and control their emotions*" (CASEL's self-awareness competency), or described SEL as skills that "*establish and maintain supportive relationships*" (CASEL's relationship skills competency), and "*make responsible and caring decisions*" (CASEL's responsible decision-making competency). Educators also spoke about SEL's pivotal role in helping students "*develop skills and apply strategies for understanding and managing their emotions*" (CASEL's self-management competency). One educator described SEL as "*learning how my awareness, emotional management, relationship skills, decision making, and social awareness impact the people I work with or teach.*"

Theme 2: "Not try to save them or ask them to breathe through their oppression"

The second theme revealed that SEL, as it is currently practiced and taught in schools, can cause harm, especially to BIPOC students. Educators quoted scholar Dena Simmons (Madda, 2019) saying that SEL can be "*white supremacy with a hug*" and noted instances in which SEL programs involved "*white teachers retraumatizing Black, Brown, and Queer youth.*" One warned that a heavily-scripted SEL curriculum "*would further white supremacy and institutional racism.*" Another cautioned that educators may "*unknowingly strip our students of their culture and language and identity in the name of SEL.*" Similarly, one educator described their experience of "*SEL practices which actually were meant to colonize, pacify, and/or make some sort of example of me, my background, my community, and my experiences.*"

Educators also stressed "*what SEL should* ***not*** *be*" (emphasis added) to avoid inflicting further harm on students. One educator, in writing about their students, said that SEL should "*not try to save them or ask them to breathe through their oppression.*" Another expressed that SEL should be about *"providing support and modeling SEL, not policing."* Relatedly, another educator highlighted that "*SEL is not a way to control.*" To avoid inflicting harm to students, one educator mentioned that SEL must be "*culturally and historically responsive, …authentic, and leave no one behind to be managed by practices of white supremacy.*"

Theme 3: "Focusing on a child's well-being"

In contrast to educators who described SEL as a set of competencies and those who described SEL as harmful to BIPOC students, many educators focused on how SEL could be implemented to promote overall well-being. This theme comprises four subthemes, each addressing a different aspect of well-being: (1) honoring identity, (2) centering humanity and the whole child, (3) promoting healing and liberation, and (4) advancing social justice.

Subtheme 3.1: Honoring identity

Educators described how SEL could be used to honor and support students' identities in co-created spaces. Educators discussed how this involved intentional action, both to create safe spaces and to facilitate youth exploration of their identities. In fact, educators defined SEL as "*a dynamic process of learning about and sharing one's identity,*" and "*…the process through which all young people and adults acquire and apply the knowledge, skills, and attitudes to develop healthy identities.*" One respondent wrote that "[SEL] *means understanding who you are and how you show up in the world and understanding who those around you are and how they experience life because of their identities.*"

Some educators described SEL as a means to connect identity with social justice, creating opportunities for students to learn "*…about their identity and emotional lives along with each other for collective social justice*" and to "*show up in our full identities - learning about our identities, understanding our identities in the context of our community and society.*" Additionally, educators stated that SEL skills must be taught *"within a context that is applicable to the cultural, linguistic, racial, and gender identities of my students."* Educators stated that creating this context involved developing spaces and facilitating one's identity exploration. One educator wrote:

> *SEL means ensuring that we make space for children and adults to explore their own identities, honor each other's identities, learn how to listen to their bodies, understand how to regulate when they become dysregulated, learn how to build authentic relationships and partnerships to reach personal and collective goals for the good of all of us.*

Others echoed this sentiment: "*Social Emotional Learning means creating an environment where children are able to have important conversations related to their identity.*" They noted that

SEL includes creating "*authentic opportunities for adults and youth to learn about each other, about oneself/one's identities.*"

In addition to providing spaces and opportunities, educators mentioned that "*SEL means knowing students' identities and who they aspire to be,*" "*helping students understand their identities and themselves,*" "*taking [students'] personal experiences and identities into account,*" and "*honoring students' identities and lives before anything else.*" One wrote that SEL, "*celebrates [students'] identity while affirming their social and emotional selves.*" Educators considered how "*teachers can facilitate this learning when honoring a learner's full identity and lived experience*" and identified themselves as facilitators in ensuring "*students are encouraged to explore all aspects of their identities.*"

Subtheme 3.2: Centering humanity and the whole child

Educators also described getting to know students as whole people and honoring their humanity as key elements of SEL. When done well, educators reflected that SEL "*means understanding students as individual people with complex emotions, past experiences, and lives outside of the classroom.*" One educator wrote that SEL should be about "*educating our children while taking into account the whole child and their life experiences that influence the way they think and learn and how they view the world and their place in it.*" Another expressed that "*SEL is the vehicle through which people learn to embrace their own humanity and the humanity of others.*"

A component of this subtheme was "*centering students' humanity above their productivity.*" Suggestions for this included addressing "*… the needs of students and teachers beyond content and pedagogy to include learning and teaching for the whole person*" and "*giving youth the opportunity to simply exist in the K-12 space and be humans, not test scores.*" One person simply stated that, "*[SEL] is radical humanity wrapped in love and care.*" Another educator summed up this concept by explaining SEL as "*taking care of kids as people, not academics.*"

Descriptions of humanizing practices extended beyond the individual to building community with others and, by doing so, building a better society overall. A respondent shared, "*[SEL] means deeply connecting to students and giving them space to deeply connect with themselves and others. It means acknowledging humanity and creating space to build community.*" Others echoed this sentiment, describing key elements of SEL as "*recognizing and supporting the humanity of individuals and the community you build together in a learning space.*" Some educators described how SEL could "*help individuals learn how they process and interpret emotions in order to learn how they can better engage with their community and society,*" thereby "*contribut[ing] to safe, healthy, and just communities.*"

Subtheme 3.3: Promoting healing and liberation

Educators described promoting healing and liberation as facets of SEL. Some described "*SEL [as] a way of healing*" and how SEL involved "*making the classroom a place of safety and healing for all students.*" One respondent reflected on "*the equal need for addressing trauma and healing through culturally rich spaces that are welcoming and responsive.*" Some educators also regarded SEL as a way to equip students with skills that would promote well-being later in life. Related descriptions of SEL included, "*learning how to heal and strengthen our hearts and minds*" and "*work grounded in healing and justice to support the well-being of our students and families.*" One educator reported that SEL could "*hopefully equip students to disrupt the causes of stress, anxiety, and trauma.*" Another succinctly wrote, "*SEL is proactive mental health care.*"

Moreover, to the educators, not only was healing considered an important outcome of SEL, it was also regarded as a mechanism for liberation. Many educators spoke about healing and liberation as parallel goals or about healing as a step toward liberation. This was present in responses such as, "*…I would define SEL as collective and community wellness, healing, and liberation*" and "[SEL] *means healing, collaboration, collective care, and freedom.*" One person wrote that SEL means "*creating healing, dignifying, liberating, and transformative spaces.*" Another shared that SEL was an opportunity to "*heal together and work towards our shared liberation.*"

Subtheme 3.4: Advancing social justice

Tying into the previous subthemes of humanizing and healing, educators identified dismantling inequitable systems as an aspirational component of SEL that promotes liberation. For instance, one educator said "*[SEL] means to develop self-awareness and emotional intelligence to interrogate systems of oppression and work towards individual and collective liberation.*" Another described SEL as "*educating the whole child in a way that names and confronts oppressive lies while helping students build their own, liberated, proud sense of self in the world.*" To one educator, SEL "*…means being able to feel deeply, then using those feelings to emancipate you and your community.*"

Educators communicated that SEL was a tool to "*dismantle inequitable narratives in schools and society*" and to "*advocate/take part in breaking down systems that dehumanize everyone.*" Educators believed that "*SEL advances educational equity*" through "*help[ing] address various forms of inequity.*" Some stated that SEL could be used "*to recognize and confront injustice and inequity,*" "*address historic imbalances in power,*" and "*reduce suffering and create a more just and equitable world.*" Other educators regarded SEL as "*a relationship-driven approach and a social justice orientation*" including "*things like self-awareness, communication, social justice.*" Specifically, educators stated that in order for SEL to be implemented well, programs need to address social and racial justice. For example, one educator stated, "*SEL, when done effectively, addresses social and racial injustices with the goal of producing healing and solution-based learning experiences,*" while another said, "*if done well, SEL creates an environment where all can feel seen and heard.*" Some mentioned justice more broadly in their definitions, reflecting that "*SEL includes justice…and*

reflection" and that "*SEL are the skills we need to be justice-oriented change makers.*" One educator considered SEL to include "*recognizing injustices and teaching about standing up for justice and many different ways to take action;*" another identified "*learning how to promote and advocate effectively for justice of all peoples*" as a key component of SEL.

For several educators, SEL meant creating spaces with students that emphasize learning and genuine dialogue on social injustices while maintaining students' social and emotional safety. Educators brought up their roles in creating brave spaces through SEL. One person, referring to their students, said, "*it's my role as the adult to help facilitate and support them in a way that supports inclusivity while disrupting harm.*" Others said that their jobs were, "*creating a teaching space in which students are encouraged to explore all aspects of their identities and address all of the 'isms' including, but not limited to, racism and white supremacy*" and "*making sure students feel safe and there is a racial and socio-economic lens to equitably approach emotional support for students.*" Creating these spaces involved "*being transparent and the equal need for addressing trauma and healing through culturally rich spaces that are welcoming and responsive.*" Such spaces enabled students "*to learn without fear of being discriminated against for being your true authentic self*" and to "*be seen, heard, and understood in an educational environment built on connection with a commitment to justice.*"

Educators also spoke about the need for SEL to be culturally affirming and antiracist. Educators' definitions were aspirational. One educator concluded, "*…we need antiracist SEL.*" Another wrote that SEL "*should affirm emotions and responses to emotions that are aligned with students' cultures- or be understood from a cultural lens (not white culture).*" Others described SEL as "*the foundation for racial equity*" and, "*culturally responsive pedagogy, antiracism, equity, inclusion, belonging, and wellness.*" Several educators wrote that SEL "*…includes being honest about racism, bias, and white supremacy and its role in perpetuating hurt.*" Another participant, referring to SEL, said, "*it is antiracist teaching that can be incorporated into every content area of the curriculum.*" Concisely tying up these sentiments, one educator defined SEL as, "*in a nutshell, trauma-informed and antiracist.*"

Discussion

Our study demonstrated the power of listening to a community of educators and honoring their genius in shaping students' educational experiences and fostering learning environments that honor, uplift, and support humanity while also creating educational content that centers students' identities and confronts injustices and white supremacy. Educators are a largely untapped resource in shaping SEL programming and its implementation, research, and policy. As policymakers and education leaders aim to prepare young people for the world that they will inherit, it is critical to include educators in all aspects of the decision-making that happens too often without them.

When asked to define SEL, educators described SEL as a set of competencies using the historical understanding of SEL coined more than two decades ago by youth development professionals as part of the formation of CASEL. This definition of SEL is: "*identifying and labeling feelings, expressing feelings, assessing the intensity of feelings, managing feelings, delaying gratification, controlling impulses, and reducing stress*" (Consortium on the School-Based Promotion of Social Competence, 1994). This is the most used, known, and disseminated definition given CASEL's leadership in the field (Graczyk et al., 2000; Schonert-Reichl et al., 2017). Therefore, we were not surprised that educators in this study defined SEL using CASEL's competencies.

We found, however, that educators also perceive SEL to be more expansive than current competency-only models. Educators in our study identified limitations and potential harm that may be caused by current SEL approaches and emphasized that SEL is a praxis whose benefit to students is dependent on how it is implemented. At face value, educators acknowledged that current SEL definitions and practices are at risk of perpetuating a "one size fits all" approach that emphasizes neutrality and assumes that everyone's emotions are perceived and welcome equally, often ignoring the impact of systemic oppression on those who have been consistently marginalized. Implementing SEL with a goal of neutrality teaches personal strategies for understanding and managing feelings without taking into account *and honoring* the unique and diverse cultures, backgrounds, and lived experiences across all students and especially of BIPOC and other students forced to the margins (Simmons, 2020a).

Our study participants described honoring identity as a prominent aspect of effective SEL to promote justice and student well-being. An SEL praxis that intentionally honors student identity aligns with literature and theory on culturally responsive pedagogy (CRP; Ladson-Billings, 1995). CRP is an approach infused into all aspects of teaching and learning to promote equity and dismantle harmful practices and narratives by meeting students where they are instead of imposing values and practices from the dominant culture on people who cannot relate to them (Simmons, 2019a) and who are not reflected by them. Though rarely discussed in the United States, most educational curricula, including SEL curricula, are based on a dominant white, Western, and individualistic culture (Picower, 2009; Kasun and Saavedra, 2016). Current SEL programs and frameworks, including CASEL, center whiteness in their glorification of productivity and employability (Committee for Children, 2016), which aligns with capitalism (Simmons, 2020b). This emphasis can result in teachers imposing a particular set of narrow values and beliefs on students about behavior, emotion management and expression, and conflict resolution.

Cultural differences in emotion regulation (Matsumoto et al., 2008) and emotion display rules (Safdar et al., 2009) between teachers and students can result in misunderstanding and miscommunication in addition to disproportionate rates of exclusionary discipline practices, academic failure, and school disengagement for youth who are institutionally marginalized

(Gregory et al., 2010; Brown-Jeffy and Cooper, 2011; Skiba et al., 2011). In addition, research has found that when educators misidentify their students' emotions, it leads to student disengagement (Hargreaves, 1998; Hargreaves, 2000). These are only a few examples of why SEL cannot ignore students' cultural, social, and political realities or be race-neutral, as neutrality often translates into centering whiteness and white comfort. Rather, it is crucial for educators to recognize and confront their implicit biases, humanize their students by getting to know them as whole people, and engage in culturally affirming SEL practices. SEL must actively and deliberately be culturally responsive, antiracist, and anti-oppressive (Simmons, 2020b). Educators in this study recognized a current gap in SEL programming which is a lack of attention to the diverse values and beliefs across individuals, groups, and cultures, affirming the need for CRP to guide and inform how we approach, implement, and "live" SEL moving forward. Participants in this study clarified that SEL cannot be practiced as skills alone and must center humanity, healing and liberation, and social justice, thus honoring students' identities, realities, and lives in our ever-complex world.

When students do not experience any reflection of themselves in curricula, this can result in the trauma of erasure (Simmons, 2020a). Traditionally, SEL programming has not been intentionally culturally responsive, contributing to this erasure, which can lead to student disengagement and other adverse social, emotional, and academic outcomes (Bottiani et al., 2020). For example, a white, straight, cisgender, female teacher, who lacks self-awareness about her positionality and who has no exposure to culturally responsive pedagogy, may teach SEL through her worldview without consideration of how her lens does not reflect the reality of her students, especially those who have been systemically marginalized. She might include content relevant to her lived experiences but fail to add content that is relevant to her students' lives, contributing to students' feeling like they do not belong. She might even teach a particular skill that is inconsistent with students' cultures such as demanding eye contact as a sign of respect when, in many Asian, Latinx, and African cultures, eye contact could be regarded as a sign of defiance and disrespect (Wages, 2015). That is, culture plays a role in how SEL competencies are developed and expressed (Hecht and Shin, 2015) and thus needs specific attention. In contrast, that same teacher can build awareness of how to implement SEL in ways that honors students for their identities and experiences rather than restricting them to a narrow view of what is "acceptable" or "appropriate." She would welcome and teach multiple expressions of respect (and other emotions) instead of ostracizing students, who lower and avert their eyes as a sign of respect. Through practicing self-reflection and self-awareness, developing cultural humility, and gaining exposure to culturally responsive pedagogy, educators can expand their knowledge of how to use SEL to humanize and honor all students.

Educators in our study regarded humanizing others as an important piece of SEL implementation. In particular, many described SEL in its ideal form as a praxis that can be responsive to student and community needs, facilitate healing, and center humanity and racial and social justice. In fact, humanization is a process that educators and scholars of education have identified as a key component of education. Philosopher and educator Paulo Freire, whose practices aimed to liberate students, wrote that humanizing students promotes freedom and justice (Freire, 1970). Similarly, scholar and writer bell hooks deemed education as a practice of freedom, where students transgress against racial, sexual, and class boundaries to achieve freedom (Hooks, 1994). Respecting the humanity of others promotes educational, social, and cultural justice and can benefit students' social and emotional health and well-being (Paris and Winn 2013). Humanizing others as part of an SEL practice allows educators to get to know students as complete beings with cultures, languages, and histories, including the intersecting systems of privilege or oppression that may shape their lives. By centering humanity in their instruction, educators can place people first over productivity and academic achievement. When school communities feel valued through humanization, they can engage more respectfully and harmoniously with each other and have the necessary room and safety for collective healing and liberation.

In addition to identifying the need for SEL to be conceptualized as a humanizing praxis, educators in our study identified healing as a crucial outcome of SEL and a way to promote liberation. For us, liberation is defined as living, learning, and thriving in the comfort of one's own skin (LiberatED SEL, 2022). This could happen through fostering belonging by inviting, welcoming, and centering students' lived experiences in instruction and interactions and incorporating healing opportunities in classrooms by ensuring all aspects of students are invited and welcomed, for example. We define healing as a regenerative process with the goal of restoring collective and individual wellbeing at the emotional, spiritual, social, psychic, and physical levels (Chavez-Diaz and Lee, 2015). The fact that nearly all educators who spoke of healing also spoke of liberation is meaningful and indicates the need for schools to be more liberating spaces.

Current SEL frameworks and practices focus more on "remaining neutral" than on building awareness, critical consciousness, and actions that address racialized stress and trauma among BIPOC students (Ginwright, 2015). For healing to happen, SEL practitioners, scholars, and educators must acknowledge and confront the racial harm and social injustice as a result of society's oppressive policies and practices as well as their own individual and institutional practices. Likewise, there needs to be an intentional and explicit practice of healing too. One method of beginning the healing process in schools is through healing-centered engagement (Ginwright, 2018), a holistic approach to trauma that involves a focus on culture, identity, and collective healing. This approach is strengths-based and focuses on human possibility and potential. Healing-centered engagement identifies healing as a collective effort rather than an individual one, and it strives to address the systemic and pervasive root causes of trauma (Ginwright, 2018). This is needed now more than

ever given the coronavirus pandemic and an increase in hate crimes (Federal Bureau of Investigation, 2021). Another approach to healing is through storytelling which challenges the validity of accepted premises or myths held by majority groups, who are often in positions of power (Delgado and Stefancic, 2001). In a classroom, students and teachers can tell personal stories or write personal narratives about what life is like for them and invite others into their worlds as they feel comfortable sharing (Solorzano and Yosso, 2002).

Collectively, healing and working toward shared liberation means taking explicit actions to dismantle oppressive social forces (Ginwright, 2018). Healing work is political, as it involves shifting the blame for harm or well-being from the individual and onto systemic and historical inequities and injustices. Educators in this study noted the importance of creating spaces to identify and actively work toward dismantling oppressive social forces, teaching about social justice, equity, and antiracism, and using culturally-affirming practices as important steps towards liberation. Educators expressed a need for SEL that includes opportunities for promoting social justice and eradicating practices and policies that contribute to inequity. Their comments echoed what SEL scholar and co-author of this manuscript, Dena Simmons, has stated about the need for SEL to address our sociopolitical reality and combat racial and social injustice (Simmons, 2019b). It is our nation's imperative to center social justice in our SEL programming, instruction, and practice so that it can live up to its full promise and facilitate connection across differences, truly reflecting the aspirational definitions of SEL that educators shared in their responses. Simply, SEL alone cannot solve racism or other forms of oppression without deliberately confronting and combating injustice (Simmons, 2020a).

A critical finding from our study was that educators highlighted how SEL, as it is currently conceptualized and thus how it may be implemented, can be weaponized against BIPOC students. This perpetuates racial harm and has been called out in recent literature (e.g., DeMartino et al., 2022). We observed this in educators' responses about how SEL can be presented as a savior for BIPOC students as well as a way to manage and control student behavior. SEL practitioners and scholars have warned about exactly what we found in our study in their descriptions of SEL as a tool that reinforces compliance and control (Simmons, 2019c; Kaler-Jones, 2020) and a way to save BIPOC students (Simmons, 2017). Regarding SEL as the savior of BIPOC students implicitly creates a power imbalance between the "savior" and those in need of saving, disempowering BIPOC youth and perpetuating racial hierarchy with white people atop. This also positions BIPOC students as a problem to solve (Simmons, 2017). When SEL is presented as a behavioral intervention for BIPOC students, it sends the message that BIPOC students behave in unacceptable ways and need to be corrected and taught to conform and comply. Thus, it is imperative that the SEL field be mindful of the social, political, and historical context in which SEL is implemented (Simmons, 2019d). Failing to do so can further marginalize and oppress BIPOC students.

Moreover, it is important to note that the education field continues to be a profession that is predominantly white while our student body is primarily BIPOC (U.S. Department of Education, 2016; Riser-Kositsky, 2019). This racial and cultural mismatch between teachers and students can have deleterious academic, behavioral, social, and emotional outcomes, especially if teachers regard SEL as a behavioral intervention and savior. On top of the disproportionate disciplinary practices that contribute to the school-to-prison nexus for BIPOC students (U.S. Department of Education, 2018), the hyper-surveillance that BIPOC students experience at school contributes to trauma and negative social and emotional health outcomes (Camangian and Cariaga, 2021), thus diminishing the promise of SEL.

Our education system must evolve to meet the unique needs of educators, students, and their families given our current sociopolitical context, shifting demographics, and technological advances. The same is true for SEL. How SEL is defined, implemented, and researched must change to meet the pressing needs of our country. Through the educators in the study, we have the beginning of a redefinition of SEL that centers identity, humanity, healing, and justice. Currently, these areas are not explicit components of most SEL definitions, models, and programs even though studies suggest that attention to them can significantly contribute to improving the quality of SEL programs and policies at schools (e.g., Davis et al., 2022; Forman et al., 2022). This study is a critical step toward understanding, redefining, and transforming SEL through the voices and perspectives of educators with the aim of SEL implementation that ensures all students can live, learn, and thrive at school in the comfort of their own skin.

As currently written, the CASEL 5 is limited, as it suggests that SEL is a set of skills that are the same for everyone and can be taught in the same way to all students. It does not yet address the shifting demographics of our nation's more racially and ethnically diverse student body (National Center for Education Statistics, 2022). Yet, years of research suggest that students learn differently and benefit from both differentiated instruction (Tomlinson and McTighe, 2006) as well as culturally responsive pedagogy (Gay, 2018; Ladson-Billings, 2021), that meets students' individual learning needs. Supporting the optimal social, emotional, and academic development of youth requires that educators and educational leaders effectively apply SEL in ways that are meaningful, relevant, and affirming to the identities and lived experiences of youth. This work requires flexibility, adaptability, representation, attention to racial and social justice, and creativity (e.g., Tan et al., 2021).

CASEL's response to calls for equity-responsive SEL has been to add "equity elaborations" to each competency as part of their enhanced SEL called "transformative SEL" (Jagers et al., 2019). Though a potential step in the right direction, these calibrations are add-ons to existing competencies, which contrast with how educators in our study described honoring identity, humanizing students, promoting healing and liberation, and advancing social justice as foundational components of SEL praxis, not as

afterthoughts. Given its institutional power, CASEL is well-positioned to shift SEL from a competency framework to a model that prioritizes a human-centered, culturally affirming, and historically responsive praxis. Doing so could help ensure that all students–and especially those who have been historically and systematically marginalized–obtain the benefits of SEL, feel a sense of belonging at school, and can flourish in their full personhood, thereby improving their mental health and well-being.

Educators' more expansive definitions of SEL in this study align with models of culturally responsive pedagogy, healing-centered education, and social and racial justice. Our findings demonstrate that SEL is not "value-neutral and independent from practices, histories and the contexts of its production" (Stetsenko, 2014, p. 181). SEL programming and professional development must evolve to ensure instruction, policies, and practices are culturally and contextually relevant and honor student experiences. Educators play a critical role in SEL's necessary evolution. With the rise in attention toward promoting mental health in schools resulting from the pandemic and civil unrest comes an opportunity to promote a holistic, culturally-affirming, and liberatory approach to SEL that supports all students. Schools are turning to SEL as a solution to the numerous challenges students are facing, but our education system must guarantee that SEL is not harmful or oppressive despite good intentions. Now is the time to reassess and redefine SEL and to move away from SEL solely as a set of competencies that may not be reflective or useful for all (and may, in fact, cause harm to some), and toward a future in which SEL honors identity for all, centers humanity and the whole child, promotes healing and liberation, and advances social and racial justice.

Limitations

As with all research, our study has limitations. For one, our question "What does SEL mean to you?" was broad. We were not always able to differentiate whether open-ended responses were reflective of an educator's current SEL practices or their aspirations for SEL unless explicitly stated. In future studies, we would clarify our question and explore both educator aspirations and current practices to measure whether there are discrepancies between the two. Additionally, the majority of educator responses were short; many consisted of only one or a few sentences. Nonetheless, these responses offered an important glimpse into educator conceptualizations of SEL. Future studies would benefit from approaches that promote greater depth of response and allow for follow-up probes such as interviews or focus groups. Future studies would also benefit from exploring how educators' definitions of SEL relate to the program implementation of SEL and student experiences and outcomes.

Another limitation is that our sample comprised educators who self-selected to attend an event focused on the intersection of SEL, racial justice, and healing in educational settings. Our participants most likely possessed a more advanced understanding of or a greater interest in the intersection of SEL, racial justice, and healing than other educators who did not register for such an event. It is likely that respondents were primed to think about SEL, racial justice, and healing based on the focus of the event and given their interest in attending. Though we believe that the educators in this study shared important, illuminating, and relevant information for the exploration of our topics of interest, we acknowledge that they may not be representative of the larger educator population. Future studies could use a random sample of educators to obtain more generalizable results.

Implications and future directions

Our study has practice, policy, and research implications for SEL. For one, our study highlighted identity affirmation, humanity, healing and liberation, and social justice as necessary and central components of SEL. Given our findings, the education system as well as education researchers and curriculum developers must employ practices that create opportunities to honor students' identities, humanize students, and facilitate healing, liberation, and justice. To be effective, these efforts must be paired with access to curricula and other relevant resources, opportunities to engage in professional development and coaching for educators and school leaders, and policies supportive of these practices. After nearly two and a half years in a pandemic with corresponding movements to ban antiracist efforts, culturally relevant and responsive practices in schools, and SEL, there is an urgent need for policymakers to confront these challenges by standing firmly for collective humanity and healing and rejecting calls to ban anything (e.g., books and curricula) that has to do with SEL, race, difference, or identity. Similarly, leaders in the SEL field could more strongly advocate for culturally affirming SEL grounded in racial and social justice, healing, and liberation instead of advocating for neutrality, which further marginalizes youth who are already disenfranchised. Our education system must create opportunities to include the BIPOC student community and other youth who are institutionally marginalized when making instructional decisions and creating academic content related to SEL and beyond so that student learning is grounded meaningfully in their own lives. Our study clarified the importance of doing so.

Future research could expand upon findings from this study by probing between-person differences in conceptualizations of SEL to identify response patterns using qualitative (e.g., open-ended survey items or interviews with educators), quantitative (e.g., latent class analysis), or mixed-method approaches (e.g., a convergent parallel design exploring how educators' responses to open-ended prompts align with responses to close-ended survey items, and if and how these represent categories of people or types of responses). Our team is already beginning to conduct this research. Additionally, because students should have a central role in their own learning and are generally more engaged when their instruction centers their lives (Byrd, 2016), an important future research direction is to gain insight directly from students about how they conceptualize and experience SEL as well as how they

want to engage in SEL at school. In particular, we will continue to adopt a community-based research design for future studies. Rooted in reciprocal partnerships, community-based research design seeks to democratize knowledge, improve practices, and bring about social change by recognizing and valuing the unique strengths and perspectives of all members involved in the research process (Atalay, 2010; Tremblay et al., 2018). This approach provides a framework for using culturally responsive, constructivist, and interpretivist strategies to address injustice, and ensures that we center the voices of educators and uplift their genius so that they have a key role in shaping their profession.

Conclusion

Our current moment is marked by exacerbated mental health challenges and trauma as a result of the COVID-19 pandemic as well as heightened racial tension and racism. While SEL has been one way our nation's schools have selected to address the resulting mental health challenges that many students and communities are experiencing from enduring a global pandemic, it will not facilitate the atonement needed in many schools and communities without centering individual and collective healing, racial and social justice, and a commitment to dismantling white supremacy in education and beyond.

Findings from this study show that SEL has the potential to inflict harm on students if it is not intentionally implemented through a culturally responsive and racially just lens. One of the primary findings of our study was the importance of centering humanity and healing in education. While the backlash against racial and social justice initiatives in education is evidence of the necessity of centering humanity in our educational instruction and practices, research demonstrates racial justice education can be a helpful tool to combat the current divisiveness in the U.S.' education system (Williams, 2021; Scientific American, 2022). SEL has tremendous potential to help us come together, understand one another, build relationships, manage conflict, and elicit social change if infused with an ideology and practice of humanization, healing, social justice, and identity affirmation, and if approached with the goal of collective liberation. When teaching, researching, or creating policies around SEL, we must pay attention to the sociopolitical and racial contexts and work to eradicate the inequities that students experience and navigate daily inside and outside of school (Madda, 2019). In sum, our nation's schools must do the deliberate work to become healing and liberating spaces so that all students have the privilege of experiencing the freedom to be who they are without repercussions, punishment, or fear of harm.

This study provides a hopeful blueprint for educational practices and policies that include educator voices and lean on their resources and experiences as we dream of and cultivate liberatory educational experiences and systems for the future. The learning gained from this study will enable those who support educators and students to adapt and adjust SEL to meet their needs more responsively, effectively, and authentically. Overall, this study lays a foundation for improving SEL implementation so that current and future programs can meet the needs of all students, especially those who have been most disenfranchised. By tapping into educators' experiences and knowledge, we honor them as co-creators of a future vision for a more humanizing and healing approach to SEL—one that centers collective liberation and examines and disrupts oppression, bias, and bigotry—so that educators and students are prepared to engage in the social change required for achieving our nation's values of equity and justice for all.

Data availability statement

The datasets presented in this article are not readily available because the dataset belongs to LiberatED. Requests to access the datasets should be directed to ML, research@liberatedsel.com.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

SO and ML analyzed and interpreted all data. SO, ML, MM, DS, and ST wrote the manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgments

The authors would like to thank Jey Blodgett, Heidi Forbes, and Madeline Ruo for their help with the study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Abolitionist Teaching Network. (2022). Available at: http://abolitionistteachingnetwork.org/ (Accessed September 9, 2022).

Alinsky, S. (1971). *Rules for Radicals: A Pragmatic Primer for Realistic Radicals*. New York, NY: Random House.

ASCD. (1997). *Social and emotional learning*. Available at: https://www.ascd.org/el/social-and-emotional-learning (Accessed September 9, 2022).

Atalay, S. (2010). "We don't talk about Catalhoyuk, we live it": sustainable archeological practice through community-based participatory research. *Archeol. Contemp. Soc.* 42, 418–429. doi: 10.1080/00438243.2010.497394

Barret, P. (2022). *Social unrest is rising, adding to risks for global economy*. Available at: https://www.imf.org/en/Blogs/Articles/2022/05/20/social-unrest-is-rising-adding-to-risks-for-global-economy (Accessed November 23, 2022).

Bauer, M. W., and Gaskell, G. (2000). "Classical content analysis: a review" in *Qualitative Researching With Text, Image and Sound: A Practical Handbook* (Thousand Oaks, CA: Sage), 131–151.

Birks, M., Chapman, Y., and Francis, K. (2008). Memoing in qualitative research. *J. Res. Nurs.* 13, 68–75. doi: 10.1177/1744987107081254

Boncu, A., Costea, I., and Minulescu, M. (2017). A meta-analytic study investigating the efficiency of socio-emotional learning programs on the development of children and adolescents. *Rom. J. Psychol.* 19, 35–41. doi: 10.24913/rjap.19.2.02

Bottiani, J. H., McDaniel, H. L., Henderson, L., Castillo, J. E., and Bradshaw, C. P. (2020). Buffering effects of racial discrimination on school engagement: the role of culturally responsive teachers and caring school police. *J. Sch. Health* 90, 1019–1029. doi: 10.1111/josh.12967

Boyd, M. R. (2020). "Community-based research: a grass-roots and social justice orientation to inquiry" in *The Oxford Handbook of Qualitative Research*. ed. P. Leavy (Oxford: Oxford Academic)

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa

Brooks, J. S., and Theoharis, G. (2018). *Whiteucation: Privilege, Power, and Prejudice in School and Society*. London: Routledge.

Brown-Jeffy, S., and Cooper, J. E. (2011). Toward a conceptual framework of culturally relevant pedagogy: an overview of the conceptual and theoretical literature. *Teach. Educ. Q.* 38, 65–84.

Byrd, C. M. (2016). Does culturally relevant teaching work? An examination from student perspectives. *SAGE Open* 6:215824401666074. doi: 10.1177/2158244016660744

Camangian, P., and Cariaga, S. (2021). Social and emotional learning is hegemonic miseducation: students deserve humanization instead. *Race Ethn. Educ.* 25, 901–921. doi: 10.1080/13613324.2020.1798374

Chavez-Diaz, M., and Lee, N. (2015). *A Conceptual Mapping of Healing Centered Youth Organizing: Building a Case for Healing Justice*. Oakland, CA: Urban Peace Movement.

Childcare Technical Assistance Network (2021). Resource guide for developing integrated strategies to support the social and emotional wellness of children. Available at: https://childcareta.acf.hhs.gov/resource/resource-guide-developing-integrated-strategies-support-social-and-emotional-wellness

Collaborative for Academic, Social, and Emotional Learning (2013a). *CASEL Schoolkit: A Guide for Implementing Schoolwide Academic, Social, and Emotional Learning*. Chicago, IL: CASEL.

Collaborative for Academic, Social, and Emotional Learning (2013b). *2013 CASEL Guide: Effective Social and Emotional Learning Programs—Preschool and Elementary School Edition*. Chicago, IL: CASEL.

Collaborative for Academic, Social, and Emotional Learning (CASEL) (2020a). An initial guide to leveraging the power of social and emotional learning as you prepare to reopen and renew your school community. Available at: https://s3.us-east2.amazonaws.com/ohiomea/covid/1589978073.pdf

Collaborative for Academic, Social, and Emotional Learning (CASEL) (2020b). Reunite, renew, and thrive: Social and emotional learning (SEL) roadmap for reopening school. Available at: https://casel.s3.us-east-2.amazonaws.com/RefocusOnSELRoadmapCASEL1.pdf

Collaborative for Academic, Social, and Emotional Learning (CASEL). (2021). Collaborating district initiative. Available at: https://casel.s3.us-east-2.amazonaws.com/RefocusOnSELRoadmapCASEL1.pdf

Committee for Children. (2016). Why social and emotional learning and employability skills should be prioritized in education. Available at: https://www.cfchildren.org/wp-content/uploads/policy-advocacy/sel-employability-brief.pdf

Consortium on the School-Based Promotion of Social Competence (1994). "The school-based promotion of social competence: theory, research, practice, and policy" in *Stress, Risk, and Resilience in Children and Adolescents: Processes, Mechanisms, and Interventions*. eds. R. J. Haggerty, L. R. Sherrod, N. Garmezy and M. Rutter (New York, NY: Cambridge University Press), 268–316.

Creswell, J. W., and Poth, C. N. (2016). *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. Thousand Oaks, CA: SAGE Publications.

Curtis, D. S., Washburn, T., Lee, H., Smith, K. R., Kim, J., Martz, C. D., et al. (2021). Highly public anti-black violence is associated with poor mental health days for black Americans. *Proc. Natl. Acad. Sci.* 118:e2019624118. doi: 10.1073/pnas.2019624118

Davis, S. J., Lettis, M. B., Mahfouz, J., and Vaughn, M. (2022). Deconstructing racist structures in K-12 education through SEL starts with the principal. *Theory Pract.* 61, 145–155. doi: 10.1080/00405841.2022.2036061

Delgado, R., and Stefancic, J. (2001). *Critical Race Theory: An Introduction*. New York, NY: New York University Press.

DeMartino, L., Fetman, L., Tucker-White, D., and Brown, A. (2022). From freedom dreams to realities: adopting transformative abolitionist social emotional learning (TASEL) in schools. *Theory Pract.* 61, 156–167. doi: 10.1080/00405841.2022.2036062

Dubois, A., and Gadde, L. E. (2002). Systematic combining: an abductive approach to case research. *J. Bus. Res.* 55, 553–560. doi: 10.1016/s0148-2963(00)00195-8

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

Fair Health. (2021). The impact of COVID-19 on pediatric mental health: a study of private healthcare claims. Available at: https://bit.ly/3t3b95F (Accessed September 9, 2022).

Federal Bureau of Investigation. (2021). *FBI Released 2020 Hate Crime Statistics*. Washington, DC: Federal Bureau of Investigation.

Forman, S. R., Foster, J. L., and Rigby, J. G. (2022). School leaders' use of social-emotional learning to disrupt whiteness. *Educ. Adm. Q.* 58, 351–385. doi: 10.1177/0013161X211053609

Freire, P. (1970). *Pedagogy of the Oppressed*. New York, NY: Penguin Classics.

Frye, K. E., Boss, D. L., Anthony, C. J., Du, H., and Xing, W. (2022). Content analysis of the CASEL framework using K–12 state SEL standards. *Sch. Psychol. Rev.*, 1–15. doi: 10.1080/2372966X.2022.2030193

Gay, G. (2018). *Culturally Responsive Teaching: Theory, Research, and Practice*. New York, NY: Teachers College Press.

Ginwright, S. (2015). *Hope and Healing in Urban Education*. London: Routledge.

Ginwright, S. (2018). The future of healing: Shifting from trauma informed care to healing centered engagement. *Medium*. Available at: https://ginwright.medium.com/the-future-of-healing-shifting-from-trauma-informed-care-to-healing-centered-engagement-634f557ce69c

Graczyk, P., Greenberg, M., Elias, M., and Zins, J. (2000). The role of the collaborative to advance social and emotional learning (CASEL) in supporting the implementation of quality school-based prevention programs. *J. Educ. Psychol. Consult.* 11, 3–6. doi: 10.1207/s1532768xjepc1101_2

Gregory, A., Skiba, R. J., and Noguera, P. A. (2010). The achievement gap and the discipline gap. *Educ. Res.* 39, 59–68. doi: 10.3102/0013189x09357621

Grové, C., and Laletas, S. (2020). "Promoting student wellbeing and mental health through social and emotional learning" in *Inclusive Education for the 21st Century: Theory, Policy and Practice*. ed. L. Graham (Oxfordshire: Taylor and Francis), 317–335.

Hargreaves, A. (1998). The emotional practice of teaching. *Teach. Teach. Educ.* 14, 835–854. doi: 10.1016/S0742-051X(98)00025-0

Hargreaves, A. (2000). Mixed emotions: teachers' perceptions of their interactions with students. *Teach. Teach. Educ.* 16, 811–826. doi: 10.1016/S0742-051X(00)00028-7

Hecht, M. L., and Shin, Y. (2015). "Culture and social and emotional competencies" in *Handbook of Social and Emotional Learning: Research and Practice*. eds. J. Durlak, C. Domitrovich, R. Weissberg and T. Gullotta (New York, NY: Guilford), 50–64.

Hooks, B. (1994). *Outlaw Culture: Resisting Representations*. New York, NY: Taylor and Francis US.

Hunter, L. J., DiPerna, J. C., Hart, S. C., Neugebauer, S., and Lei, P. (2022). Examining teacher approaches to implementation of a classwide SEL program. *Sch. Psychol. Forum* 37, 285–297. doi: 10.1037/spq0000502

Israel, B., Eng, E., Schulz, A., and Parker, E. (2005). *Methods in Community-based Participatory Research for Health*. San Francisco, CA: Jossey–Bass.

Israel, B., Schulz, A., Parker, E., and Becker, A. (1998). Review of community-based research: assessing partnership approaches to improve public health. *Am. Rev. Public Health* 19, 173–202. doi: 10.1146/annurev.publhealth.19.1.173

Jagers, R. J., Rivas-Drake, D., and Williams, B. (2019). Transformative social and emotional learning (SEL): toward SEL in service of educational equity and excellence. *Educ. Psychol.* 54, 162–184. doi: 10.1080/00461520.2019.1623032

Jones, S. E., Ethier, K. A., Hertz, M., DeGue, S., Le, V. D., Thornton, J., et al. (2022). Mental health, suicidality, and connectedness among high school students during the COVID-19 pandemic - adolescent behaviors and experiences survey. United States, January-June 2021. *Morbidity Mortality Weekly Rep. Suppl.* 71, 16–21. doi: 10.15585/mmwr.su7103a3

Kaler-Jones, C. (2020). When SEL is used as another form of policing. *Communities for Just Schools Fund*. Available at: https://medium.com/@justschools/when-sel-is-used-as-another-form-of-policing-fa53cf85dce4

Kasun, G. S., and Saavedra, C. M. (2016). Disrupting ELL teacher candidates' identities: indigenizing teacher education in one study abroad program. *TESOL Q.* 50, 684–707. doi: 10.1002/tesq.319

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *Am. Educ. Res. J.* 32, 465–491. doi: 10.3102/00028312032003465

Ladson-Billings, G. (2021). *Culturally Relevant Pedagogy: Asking a Different Question*. New York, NY: Teachers College Press.

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

Leonard, A. M., and Woodland, R. H. (2022). Anti-racism is not an initiative: how professional learning communities may advance equity and social-emotional learning in schools. *Theory Pract.* 61, 212–223. doi: 10.1080/00405841.2022.2036058

LiberatED SEL. (2022). Available at: https://linktr.ee/liberated_sel (Accessed September 9, 2022).

Madda, M. J. (2019). *Dena Simmons: Without Context, Social-emotional Learning Can Backfire*. Burlingame, CA: EdSurge.

Matsumoto, D., Yoo, S. H., and Nakagawa, S. (2008). Culture, emotion regulation, and adjustment. *J. Pers. Soc. Psychol.* 94, 925–937. doi: 10.1037/0022-3514.94.6.925

Miles, M. B., and Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. Thousand Oaks, CA: Sage.

Miller, E. (2021). *For Some Black Students, Remote Learning has Offered a Chance to Thrive*. Washington, DC: NPR.

Miller, D. W., Zyromski, B., and Brown, M. J. (2022). Reimagining SEL as a tool to deconstruct racist educational systems. *Theory Pract.* 61, 141–144. doi: 10.1080/00405841.2022.2043708

Modan, N. (2020). Pandemic-induced trauma, stress leading to "uptick" in SEL need. K-12 dive. Available at: https://www.k12dive.com/news/pandemic-induced-trauma-stress-leading-to-uptick-in-sel-need/576710/

Mott (2021). Mott poll report: How the pandemic has impacted teen mental health. Available at: https://mottpoll.org/sites/default/files/documents/031521_MentalHealth.pdf

National Center for Education Statistics. (2022). Racial/ethnic enrollment in public schools. Available at: https://nces.ed.gov/programs/coe/indicator/cge

O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., and Cook, D. A. (2014). Standards for reporting qualitative research. *Acad. Med.* 89, 1245–1251. doi: 10.1097/acm.0000000000000388

O'Connor, C., and Joffe, H. (2020). Intercoder reliability in qualitative research: debates and practical guidelines. *Int J Qual Methods* 19:160940691989922. doi: 10.1177/1609406919899220

OECD. (2020). *The Impact of COVID-19 on Student Equity and Inclusion: Supporting Vulnerable Students During School Closures and School Re-openings*. Paris: Organisation for Economic Co-operation and Development.

Paris, D., and Winn, M. T. (2013). *Humanizing Research: Decolonizing Qualitative Inquiry With Youth and Communities*. Thousand Oaks, CA: Sage.

Peterson, T. H. (2009). Engaged scholarship: reflections and research on the pedagogy of social change. *Teach. High. Educ.* 14, 541–552. doi: 10.1080/13562510903186741

Picower, B. (2009). The unexamined whiteness of teaching: how White teachers maintain and enact dominant racial ideologies. *Race Ethn. Educ.* 12, 197–215. doi: 10.1080/13613320902995475

Reinharz, S. (1992). *Feminist Methods in Social Research*. Oxford: Oxford University Press.

Reliefweb. (2020). COVID-19: Most marginalized children will bear the brunt of unprecedented school closures around the world. Available at: https://reliefweb.int/report/world/covid-19-most-marginalised-children-will-bear-brunt-unprecedented-school-closures

Riser-Kositsky, M. (2019). Education statistics: facts about American schools. *Education Week*. Available at: https://www.edweek.org/leadership/education-statistics-facts-about-american-schools/2019/01

Ritzer, G. (1996). *Modern Sociological Thought* (*4th edn.*). New York, NY: McGraw–Hill.

Rutgers Center for Effective School Practices (2022). Social-emotional learning in response to COVID-19. Available at: https://cesp.rutgers.edu/sites/default/files/RU.CESP_Research.Brief_Covid.SEL.RTI.pdf

Safdar, S., Friedlmeier, W., Matsumoto, D., Yoo, S. H., Kwantes, C. T., Kakai, H., et al. (2009). Variations of emotional display rules within and across cultures: a comparison between Canada, USA, and Japan. *Can. J. Behav. Sci./Revue Canadienne Des Sciences Du Comportement.* 41, 1–10. doi: 10.1037/a0014387

Sanders, B. (2020). The power of social and emotional learning: why SEL is more important than ever. *Forbes*. Available at: https://www.forbes.com/sites/forbesnonprofitcouncil/2020/12/07/the-power-of-social-and-emotional-learning-why-sel-is-more-important-than-ever/?sh=6636e96a7a29

Schonert-Reichl, K. A., Kitil, M. J., and Hanson-Peterson, J. (2017). *To Reach the Students, Teach the Teachers: A National Scan of Teacher Preparation and Social and Emotional Learning. A Report Prepared for CASEL. Collaborative for Academic, Social, and Emotional Learning*. Chicago, IL: CASEL.

Scientific American. (2022). Teaching about racism is essential for education. *Scientific American*. Available at: https://www.scientificamerican.com/article/teaching-about-racism-is-essential-for-education

Seale, C. (1997). Ensuring rigor in qualitative research. *Eur. J. Public Health.* 7, 379–384. doi: 10.1093/eurpub/7.4.379

Simmons, D. N. (2017). How to change the story about students of color. *Greater Good Magazine*. Available at: https://greatergood.berkeley.edu/article/item/how_to_change_story_of_students_of_color (Accessed September 9, 2022).

Simmons, D. N. (2019a). You can't be emotionally intelligent without being culturally responsive: why FCS must employ both to meet the needs of our nation. *J. Fam. Consum. Sci.* 111, 7–16. doi: 10.14307/jfcs111.2.7

Simmons, D. N. (2019b). How to be an antiracist educator. ASCD. Educ. Update, 61(10). Available at: https://www.ascd.org/el/articles/how-to-be-an-antiracist-educator

Simmons, D. N. (2019c). A call to action for SEL's next 25 years. 2019 SEL exchange, 2 October, Chicago, IL.

Simmons, D. N. (2019d). Why we can't afford whitewashed social-emotional learning. ASCD. Educ. Update, 61(4). Available at: https://www.ascd.org/el/articles/why-we-cant-afford-whitewashed-social-emotional-learning

Simmons, D. N. (2020a). Healing black students' pain. Education. *Leadership* 78, 80–81.

Simmons, D. N. (2020b). Abolitionist teaching and the future of our schools [video]. *Haymarket Books*. Available at: https://www.haymarketbooks.org/blogs/179-abolitionist-teaching-and-the-future-of-our-schools (Accessed September 9, 2022).

Simmons, D. N. (2021). Why SEL alone isn't enough. *Educ. Leadersh.* 78, 30–34.

Skiba, R. J., Horner, R. H., Chung, C. G., Rausch, M. K., May, S. L., and Tobin, T. (2011). Race is not neutral: a national investigation of African American and Latino disproportionality in school discipline. *Sch. Psychol. Rev.* 40, 85–107. doi: 10.1080/02796015.2011.12087730

Solorzano, D. G., and Yosso, T. J. (2002). Critical race methodology: counter-storytelling as an analytical framework for education research. *Qual. Inq.* 8, 23–44. doi: 10.1177/107780040200800103

Speziale, H. S., Streubert, H. J., and Carpenter, D. R. (2011). *Qualitative Research in Nursing: Advancing the Humanistic Imperative*. Philadelphia, PA: Lippincott Williams and Wilkins.

Stetsenko, A. (2014). "Transformative activist stance for education: Inventing the future in moving beyond the status quo," in *Psychology in Education: Critical Theory~Practice*. ed. T. Corcoran (Rotterdam, The Netherlands: Sense), 181–198.

Tan, K., Wegmann, K., Patino, R., Hand, B., Mitchell, J., and Moser, K. (2021). Social and emotional learning group work during the COVID-19 pandemic, the reopening, and the mobilization for racial justice. *Child. Sch.* 43, 118–122. doi: 10.1093/cs/cdab002

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

Thayer, A. J., Campa, D. M., Weeks, M. R., Buntain-Ricklefs, J., Low, S., Larson, M., et al. (2019). Examining the differential effects of a universal SEL curriculum on student functioning through the dual continua model of mental health. *J. Prim. Prev.* 40, 405–427. doi: 10.1007/s10935-019-00557-0

Tomlinson, C. A., and McTighe, J. (2006). *Integrating Differentiated Instruction and Understanding by Design: Connecting Content and Kids*. Minnesota: ASCD.

Tremblay, M. C., Martin, D. H., McComber, A. M., McGregor, A., and Macaulay, A. C. (2018). Understanding community-based participatory research through a social movement framework: a case study of the Kahnawake schools diabetes prevention project. *BMC Public Health* 18:487. doi: 10.1186/s12889-018-5412-y

Tuck, E., and Yang, K. W. (2012). Decolonization is not a metaphor. Decolonization: Indigeneity. *Educ. Soc.* 1, 1–40.

U.S. Department of Education. (2016). 2013-2014 civil rights data collection: data highlights on school climate and safety in our nation's public schools. Available at: https://www2.ed.gov/about/offices/list/ocr/docs/school-climate-and-safety.pdf

U.S. Department of Education. (2018). 2015-2016 civil rights data collection: key data highlights on equity and opportunity gaps in our nation's public schools. Available at: https://www2.ed.gov/about/offices/list/ocr/docs/2013-14-first-look.pdf

Varma, P., Junge, M., Meaklim, H., and Jackson, M. L. (2021). Younger people are more vulnerable to stress, anxiety and depression during COVID-19 pandemic: a global cross-sectional survey. *Prog. Neuro-Psychopharmacol. Biol. Psychiatry* 109:110236. doi: 10.1016/j.pnpbp.2020.110236

Vasquez Reyes, M. (2020). The disproportional impact of COVID-19 on African Americans. *Health Hum. Rights* 22, 299–307. PMID: 33390715

Wages, M. (2015). *Creating Culturally Responsive Schools: One Classroom at a Time*. Lanham, MD: Rowman & Littlefield.

Weare, K. (2017). "Promoting social and emotional wellbeing and responding to mental health problems in schools," in *Global Mental Health* eds. S. Bährer-Kohler, F., Carod-Artal (Springer, Cham), 113–125.

Wicks, P. G., Reason, P., and Bradbury, H. (2008). "Living inquiry: personal, political and philosophical groundings for action research practice," in *The Sage Handbook of Action Research: Participative Inquiry and Practice* eds. P. Reason and H. Bradbury (SAGE Publications Ltd), 15–30.

Williams, M. B. (2021). A reckoning of systemic racism and U.S. history is long overdue. Available at: https://www.teachforamerica.org/one-day/opinion/a-reckoning-of-systemic-racism-and-us-history-is-long-overdue

Wilson, S. (2020). *Research is Ceremony: Indigenous Research Methods*. Nova Scotia: Fernwood Publishing.

Yardley, L. (2000). Dilemmas in qualitative health research. *Psychol. Health* 15, 215–228. doi: 10.1080/08870440008400302

Yoder, N., Posamentier, J., Godek, D., Seibel, K., and Dusenbury, L. (2020). From response to reopening: state efforts to elevate social and emotional learning during the pandemic. Committee for Children. Available at: https://casel.org/state-efforts-elevate-social-emotional-learning-during-pandemic/

Youth Truth. (2021). Students weigh in, part II: Learning and well-being during COVID-19. Available at: http://youthtruthsurvey.org/wp-content/uploads/2021/02/YouthTruth-Students-Weigh-In-Part-II-Learning-and-Well-Being-During-COVID-19.pdf

Appendix A

Frequency of codes by theme.

Theme/subtheme	Frequency
Theme 1: Developing skills and competencies	290
Theme 2: "Not try to save them or ask them to breathe through their oppression"	28
Theme 3: "Focusing on a child's well-being"	
Subtheme 1: Honoring identity	70
Subtheme 2: Centering humanity and the whole child	152
Subtheme 3: Promoting healing and liberation	68
Subtheme 4: Advancing social justice	140

frontiers | Frontiers in Education

TYPE Systematic Review
PUBLISHED 10 January 2023
DOI 10.3389/feduc.2022.965538

OPEN ACCESS

EDITED BY
Andres Molano,
University of the Andes, Colombia

REVIEWED BY
Lori-Ann Rosalind Sacrey,
University of Alberta, Canada
Laura Sokal,
The University of Winnipeg, Canada
Celene Domitrovich,
Georgetown University, United States

*CORRESPONDENCE
Triin Ulla
✉ triin.ulla@tlu.ee

SPECIALTY SECTION
This article was submitted to Educational Psychology, a section of the journal Frontiers in Education

RECEIVED 09 June 2022
ACCEPTED 12 December 2022
PUBLISHED 10 January 2023

CITATION
Ulla T and Poom-Valickis K (2023) Program support matters: A systematic review on teacher- and school related contextual factors facilitating the implementation of social-emotional learning programs.
Front. Educ. 7:965538.
doi: 10.3389/feduc.2022.965538

Program support matters: A systematic review on teacher- and school related contextual factors facilitating the implementation of social-emotional learning programs

Triin Ulla* and Katrin Poom-Valickis

School of Educational Sciences, Tallinn University, Tallinn, Estonia

School- and teacher-related contextual factors are those that often influence the quality of social-emotional learning (SEL) program implementation, which in turn has an impact on student outcomes. The current paper was interested in (1) Which teacher- and school-related contextual factors have been operationalized in articles that focus on the relationship between implementation quality indicators 200 and contextual factors in SEL program implementation in schools? (2) Which contextual factors would demonstrate the highest frequency of statistically significant relationships with SEL program implementation quality indicators and could therefore be more essential for ensuring the program outcomes? Determining the more significant contextual factors would allow for more focused and better-informed teacher professional development for supporting students' social and emotional skills, it can also be useful for hypothesis development for quasi- experimental research designs of SEL program implementation on the school level. A systematic literature search was conducted in seven electronic databases and resulted in an initial sample of 1,281 records and additional journal and citation sampling of 19 additional records. 20 articles met the final inclusion criteria for the study (19 quantitative and one mixed methods). Inductive content analysis and quantitative analysis were employed to map the variables and estimate the relative frequency of statistically significant relationships across studies. Four categories of contextual factors were revealed: program support, school, teacher, and student categories. The results of the study reveal the diversity in contextual factors studied across SEL program implantation quality and bolster the relevance of program support factors (modeling activities during coaching and teacher–coach working relationship) for ensuring implementation quality. A link between teacher burnout and program dosage was revealed. Student factors emerged as a

separate contextual level in school, with special attention to student baseline self-regulation that may influence SEL program implementation quality.

KEYWORDS

social-emotional learning, teachers, school, program support, students, implementation quality, contextual factors, systematic literature review

1. Introduction

The purpose of general education, in addition to cultivating learners' cognitive and academic skills, is to facilitate social change, and therefore, children's social, emotional, and character development has become increasingly more emphasized and intertwined with compulsory education (Jones and Bouffard, 2012; Kochenderfer-Ladd and Ladd, 2016; Elias, 2019). Social and emotional learning (SEL) refers to the process through which children and adults acquire and effectively apply the knowledge, attitudes, and skills necessary to understand and manage emotions, set and achieve positive goals, feel and show empathy for others, establish and maintain positive relationships, and make responsible decisions (Weissberg et al., 2015). SEL serves as an interdisciplinary field that aligns areas for educators, researchers, and policymakers that address students' capacities to coordinate cognition, affect, and behavior, as well as navigate daily challenges and succeed in life, career, and college (Osher et al., 2016). International policymakers have emphasized the importance of social and emotional skill development to assure students' readiness as future citizens in a world characterized by more turbulence and uncertainty (OECD, 2021a,b).

Studies, where peer and teacher report measures have been used, indicate that children, when starting school, may not demonstrate the basic skills needed for effective collaboration, emotional, and behavioral control, and could be at risk of educational failure (Rabiner et al., 2016; Suntheimer and Wolf, 2020). The risk factors include lower socio-economic background, being younger or male (Zakszeski et al., 2020), having weaker executive functioning skills (Suntheimer and Wolf, 2020), or prior experiences of peer rejection (Ladd, 2006). Furthermore, there is evidence that if not offered assistance, problematic behavior tends to cumulate toward more aggression in adolescence (Appleyard et al., 2005), which can in turn impact other learners through negative peer influence and stress contagion, and lead to distress in the learning environment (Burgess et al., 2018). Social emotional skill support, however, can be highly beneficial to those at risk of externalizing problems (Jones et al., 2011; Calhoun et al., 2020; Streimann et al., 2020).

Coherent with the findings that social-emotional (SE) and intellectual development are intertwined (Cantor et al., 2019), the integration of evidence-based programs to support SE skill development (e.g., SEL interventions) within the academic curriculum has been shown to contribute toward both: decreasing problematic behaviors, such as disruptions or aggressive behavior in the classroom; and enhancing students' academic achievement (Durlak et al., 2011; Oberle et al., 2014; Corcoran et al., 2017). SEL interventions can therefore be quite crucial in preventing educational segregation or developmental disadvantage in today's educational context. Sustainably coordinated efforts on the school level that promote a safe and cooperative environment and support the practice of SEL competences in everyday situations benefit all children and make schools optimal contexts for SE learning (SEL) (Zins et al., 2004; Jones and Bouffard, 2012).

Earlier research shows that the success of an SEL program, specifically the impact it has on student outcomes, is dependent on its implementation process (Durlak and DuPre, 2008; Cook et al., 2018; Humphrey et al., 2018), the quality of which is affected by different contextual factors, such as positive work climate or teacher predisposition (Kam et al., 2003; Durlak and DuPre, 2008). In other words, the implementation of a school-based SEL program is a process that is situated in the school context, which in turn influences the program outcomes students may obtain, through supporting or hindering that process. Therefore, looking at the qualities of the implementation process has been emphasized as a crucial research area (Durlak and DuPre, 2008; Durlak, 2015), as this makes the benefits of the programs available to students.

Despite copious research published during the past few years on SEL program implementation, there is no consensus on which contextual factors are most essential, or to which areas the highest degree of effort should be directed, in order to ensure the most supportive quality context for SEL program implementation in the school. Although related reviews have been done before (e.g., Dusenbury et al., 2003; Durlak and DuPre, 2008), to the best knowledge of the authors, none have solely concentrated on programs that support SE skills at school. The current article employs a systematic literature review process to synthesize school-based research on SEL program implementation that specifically looks at interactions between SEL program implementation quality indicators and teacher- and school-related contextual factors. The aim of the current literature review is to map the diversity in school-based contextual factors that have been explored in relation

to SEL program implementation quality and to clarify, which of them may have proven more consistently significant for implementation quality in schools. The analysis was guided by the following research questions:

(1) Which teacher- and school-related contextual factors have been operationalized in articles that focus on the relationship between implementation quality indicators and contextual factors in SEL program implementation in schools?
(2) Which contextual factors demonstrate the highest frequency of statistically significant relationships with SEL program implementation quality indicators and could therefore be more essential for ensuring the program outcomes?

In the next section, we first offer a brief overview of how the quality of SEL program implementation has been defined in previous research and shortly discuss the complexities in defining implementation quality. Afterward, we will give an overview of the systematic literature review process and data analysis. In the last paragraph of the article, we discuss the main findings and present possible future research avenues.

2. Conceptualizing the quality of SEL program implementation and its context

Domitrovich et al. (2008) have defined implementation quality as "the discrepancy between what is planned and what is actually delivered"; measures of implementation therefore also indicate high or low implementation quality in the school context (Dusenbury et al., 2003). Osher et al. (2016) also indicated that procedural fidelity to the original program design and core features is mostly reported as synonymous with implementation quality. Different aspects of the implementation process have been differentiated (Dane and Schneider, 1998; Berkel et al., 2011; Durlak, 2016), and those have been sometimes described similarly but named somewhat differently or vice versa—similar labels may have been used for different levels of the construct. For example, in implementation science "fidelity" is used as an umbrella term for implementation quality, and "adherence" is used to indicate a measure of delivering program components (e.g., Century et al., 2010; Proctor et al., 2011), whereas in other areas, such as SEL program implementation, for example, fidelity is used synonymously with adherence (Durlak, 2016). This notion is known as the "jingle-jangle fallacy" (Jones et al., 2019)—the lack of clarity in the program implementation vocabulary is an issue often pointed out in this context (Dane and Schneider, 1998; Century et al., 2010; Proctor et al., 2011; Cook et al., 2019). Of the aspects distinguished in implementation research, the critical ones (Dane and Schneider, 1998; Century et al., 2010) are:

(1) adherence/fidelity—the degree to which the major components of the program have been faithfully delivered;
(2) exposure/dosage—the amount of program delivered;
(3) quality of delivery—the qualitative aspects of program delivery: how well or in which manner the program is carried out; and
(4) participant responsiveness—the manner in which the program engages its participants (Dane and Schneider, 1998; Durlak, 2016).

Berkel et al. (2011) and Durlak (2016) also point to adaptation as an outcome measure of program implementation; the authors of this article, however, have not come across this quality indicator in SEL studies that look at implementation as an outcome. The implementation aspects most commonly studied in SEL program implementation are adherence/fidelity and exposure/dosage[1] (Durlak, 2016). In the current article, the implementation quality is treated synonymously with the four implementation process indicators listed above, as those are commonly reported in SEL program research that treats implementation as an outcome.

In addition to seeing program implementation quality as equivalent to its process characteristics, several accounts suggest the contextual dimension as a constituent part of implementation quality (Osher et al., 2016). Some conceptual models have been introduced to systematize variables that influence the implementation of SEL interventions. Durlak and DuPre (2008), for example, based their model on an extensive literature review and interactive systems framework (ISF) approach and compiled a list of 23 contextual variables on 5 levels of implementation (community; provider e.g., teacher; organizational capacity; prevention support system (such as training and coaching); and the program itself with its fit to the implementation context). The most comprehensive level in this model is the organizational context—which is divided into three subcategories: general factors, specific practices, and staffing considerations. Domitrovich et al. (2008), on the other hand, have suggested a three-level ecological framework for high-quality implementation of programs in schools that places implementation quality in the nested ecosystem of individual, school, and macro-level factors. The individual level holds teacher psychological and professional characteristics and attitudes, the school level holds both the organizational climate and culture, as well as the classroom climate; and the macrosystem refers to policies and partnerships on a larger scale. In this model, coaching and training are seen as inherent to the program implementation process itself and not as a separate contextual layer. Durlak and DuPre (2008) posit that their

1 Further in this paper referenced simply as "adherence" and "dosage".

framework might overlook some factors, and when compared to Domitrovich et al. (2008) they do not include any student-related factors. In comparison, Domitrovich et al. (2008) place classroom climate and school culture on the same level, despite evidence that classroom climate is a nested level in the school context that can vary considerably within one school (e.g., Marsh et al., 2012). In sum, both—aspects of the implementation process, and the contextual factors that support them—are part of understanding the quality of the implementation process, for creating the conditions where students reap the greatest benefit.

Looking at the two decades of SEL program implementation research from the perspective of contextual influences may allow for drawing more systematic conclusions over time. There is a growing body of research looking at the more focused pieces of the puzzle, concentrating, for example, on a few aspects of the school ecology and looking for statistically significant predictors of implementation quality indicators among contextual factors (for example, perceived school organizational health or teacher burnout; e.g., Ransford et al., 2009; Musci et al., 2019). The current article, thus, contributes to the theoretical development which attempts to illuminate the structure and dynamics of the contextual factors that can support or hinder quality SEL program implementation.

Determining the more significant contextual factors would also allow for more focused and better-informed teacher professional development for supporting students' SE skills, both in terms of inservice and preservice training and support, as well as teacher coaching. It would allow for informed consultation by teachers and school personnel toward the more effective implementation of SEL programs. It can be additionally useful for hypothesis development for quasi-experimental research designs of SEL program implementation on the school level, to ascertain whether or not the more pronounced characteristics have indeed a more distinct role to play in SEL program implementation in practical life, enabling more supportive outcomes for students, and explain SEL program implementation process quality in more detail.

3. Methods

For this study, a systematic literature review process was conducted, to identify the significant relationships between SEL program implementation indicators and teacher- and school-related contextual factors within empirical journal articles that focus specifically on those relationships.

3.1. Procedure for searching, identifying, and selecting articles

The search process followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement protocol (Moher et al., 2009). The EBSCOhost Web[2] databases were used for the search, which is the search engine with the highest volume of meta-data (over 70,000 journals, which is higher than the Web of Science or Scopus). This database was selected to avoid duplication in the search procedure due to a short supply of allocated human resources. In the following sections, an overview of the search procedure and analysis is offered.

In the first step, a list of databases was included for the search: *Academic Search Complete, APA PsycArticles, eBook Collection (EBSCOhost), ERIC, Health Source: Nursing/Academic Edition, MasterFILE Premier, MasterFILE Reference eBook Collection, MEDLINE, and Teacher Reference Center*. For the purpose of retrieving the highest number of relevant research articles, the following criteria were set for the EBSCO search: (1) apply related words, (2) apply equivalent subjects, (3) peer reviewed, (4) published between 1 January 2000 and 31 December 2021, and (5) English as publication language. Altogether 29 preliminary test searches were conducted prior to the final search on 28 April 28, 2021, in order to find the combination of search terms that would result in the highest concentration of relevant articles for the study. After several pilot searches with various topic-related search terms, resulting in either a too-narrow or a too-widespread result, it was decided to add the specific names of the recognized SEL programs (as reported by Jones et al., 2017) to the search, to find more relevant articles (counting on the parameters of related words and equivalent subjects to direct to all the relevant results). The final search strategy was carried out in the combination of the following search terms: SE *OR socio-emotional OR social and emotional OR sel OR social competence OR character OR behavior OR behavior OR intervention AND teacher* OR school* AND implementation quality OR implementation variable**. The search term *program* was excluded during the test search phase for producing too wide a range of results in heavily dispersed topics from policymaking to engineering, and substituted with the search term *intervention* which is a frequent term used for SEL programs. The search term *behavior* was added as whole school positive behavior intervention support is sometimes considered relative to SEL (e.g., Elias, 2019), and its primary level or Tier 1 activities are hard to differentiate from SEL, as they focus on teaching all students the skills for self-regulation and social success (Walker et al., 1996). The final database search on 28 April 2021, resulted in 1,883 articles. After removing duplicates, 1,281 articles remained for abstract level screening (**Figure 1**).

2 http://www.ebscohost.com

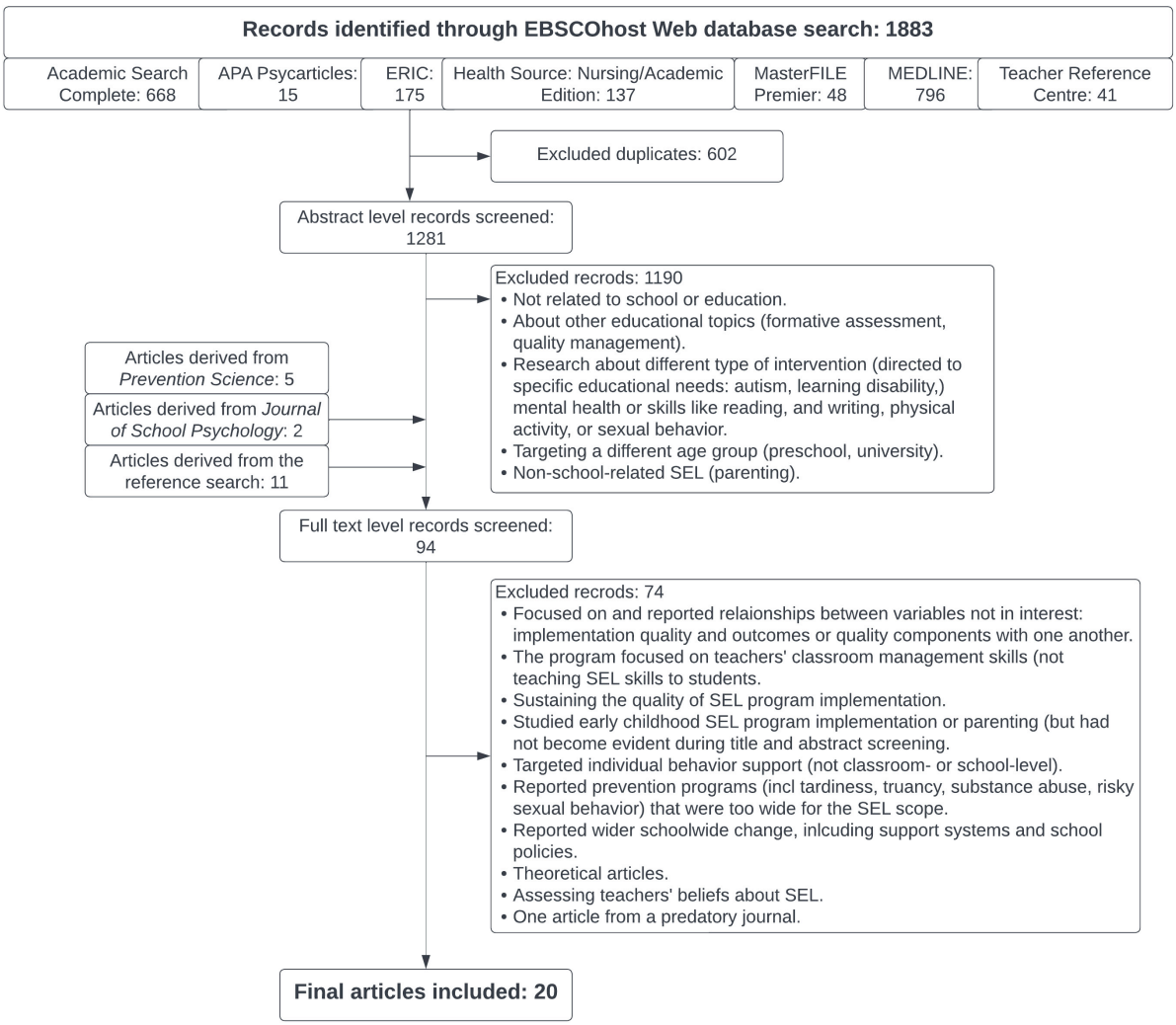


FIGURE 1
Flowchart of the literature search process.

3.2. Inclusion/exclusion criteria

The included 1,281 articles were first screened on the title and abstract level. The inclusion criterion for first level screening was: the article was about SEL programs carried out in schools, in grades 1–12. Two independent researchers assessed 10% of the articles ($n = 130$) independently on the scale of yes/no/maybe, the interrater reliability, Cohen's k was, 70 (substantial) across all labels and, 93 (excellent) between yes and no. The differences in agreements were discussed thoroughly, all occurring differences were discussed until unanimous consent was reached. Due to a high interrater agreement regarding yes/no, articles falling under both, *yes* and *maybe* were included in the second round of screening (**Figure 1**). The excluded articles were either not related to school or education; were about different educational topics (formative assessment, quality management); reported research about a different type of intervention (directed to specific educational needs, e.g., autism, learning disability, or mental health or toward other specific skills, reading, writing, physical activity, nutrition or sexual behavior); they were overall targeting a different age group (preschool or university) or different type of non-school-related SEL program (parenting). After the abstract and title-level screening, 76 articles were included for text level screening. Journal and reference sampling was carried out throughout May 2021 and additionally, in October 2022. As prevention science has been identified as the framework that drives the SEL field (Jones et al., 2019), all numbers of *Prevention Science* were screened. Similarly, all 2000–2022 issues of the *Journal of School Psychology* were screened. From the additional sampling types, 18 extra articles were found for screening (**Figure 1**).

In the second round of screening, the full texts of 94 articles were screened, 4 inclusion criteria were (1) the articles were about an evidence-based SEL approach that was applied in the school on the school level- or classroom level in grades 1–12; (2) the evidence-based SEL approach specifically included teaching SE skills to students; (3) the article was looking at the relation of some SEL program implementation indicator and some school or teacher related contextual variable, and directly aimed at determining the relationship between the two, and (4) the article reported empirical research. It must be noted that in different frameworks understandings of what an SEL program is may somewhat vary (Osher et al., 2016) and we, thus, specify what we see as the essential components of an SEL program. In the context of the current article, the SEL program is seen as a program that (a) teaches and actively practices SE skills (Zins et al., 2004; Jones et al., 2017) through (b) either kernel- or curriculum-oriented activities (Elias, 2019).

Again, two experts screened 30 full-text articles separately (with the percentages of disagreement ranging from 6.7 to 20 for the inclusion criteria) and discussed any occurring differences until a unanimous understanding was reached. The remainder of the articles was screened by the first author of the study, with the second coder double-checking all results, all emerging questions were solved in dialogue until a unanimous decision was reached. In the final sample, 20 articles were included (**Figure 1**). The articles excluded in this phase either (1) reported the implementation process as related to the outcomes of evidence based SEL programs, (but not contextual factors on implementation variables); (2) focused on: (a) a program that supported teachers' classroom management skills [but not teaching SE skills to students]; (b) sustaining or measuring quality of program implementation; (c) early childhood SEL program implementation (but had not become evident during title and abstract screening); (d) individual, not classroom- or school-level skill support, like individual behavior intervention; (e) reported a wide array of prevention programs (including tardiness and truancy, substance abuse, risky sexual behavior) and were thus too wide for the SEL scope; (f) reported wider school-wide change, including support systems and school policies, not only SEL support; (g) were theoretical in nature; (h) were addressing teachers' beliefs of SEL as such, (i) reported relationships only between different quality indicators; (j) one article had been published in a journal listed as predatory (and was, thus, excluded). An overview of the articles included in the sample is presented in **Table 1**.

Altogether 20 articles remained in the sample after the screening, of which 19 were quantitative, and 1 used a mixed design. The program most frequently studied was the PAX[3] version of the Good Behavior Game (GBG, four times), followed by Tier 1 of the school-wide positive behavioral interventions and supports (SWPBIS), and LifeSkills Training (reported 3 times); GBG, Promoting Alternative THinking Strategies (PATHS), and Positive Action were reported two times.

Dosage and quality of delivery were operationalized as implementation quality indicators in 13 studies, and adherence in 10 studies (refer to **Table 1**). Participant responsiveness was operationalized in 4 studies only. In one Dowling and Barry (2020, study 18), the study which used a mixed design, the four quality indicators (adherence, dosage, quality of delivery, and participant responsiveness) were combined into an index after quantitative assessment across implementers; interviews were then carried out and the content of the interviews compared across high and low quality implementers. The authors explained the use of an index score of four kinds of variables as "to measure implementation across multiple dimensions." The remainder of the 19 articles looked at the relationships with contextual factors and implementation quality indicators separately, and no index score was used. The rationale for using the specific implementation quality indicators was not usually offered, even if several quality indicators were assessed. For example, Mihalic et al. (2008, study 3) operationalized four implementation quality indicators (adherence, dosage, quality of delivery, and participant responsiveness) but did not provide an additional rationale for including all 4, except for referencing them as "primary elements of implementation fidelity." Ransford et al. (2009, study 4), for example, studied dosage and quality of delivery and referenced them simply as "two common measures of program fidelity." Johnson et al. (2018, study 14) stood out by explaining their choice of dosage and quality of delivery variables as structurally more relevant to the contextual factor studied (program support: coaching). In general, the provision of an explanation for the choice of implementation quality indicators was fairly uncommon.

3.3. Data analysis

In the first step, a detailed coding manual for the selected articles was created, based on Cooper's (2017) guidelines, coding was done on six separate coding sheets, which included (1) study characteristics such as purpose and research questions, study type, design, and analytic strategy, program studied, type of SEL program (curriculum or kernel), study setting and year of data collection, (2) participant and sample characteristics, (3) contextual variable name and operationalization, data collection, and measurement, including psychometric properties, (4) implementation quality indicator name and operationalization, data collection, and measurement, including psychometric properties, (5) direct interactions between contextual and implementation quality indicators, and (6) interaction effects between different contextual variables (where provided). In this step, four articles (20%) were coded

3 The PAX Good Behavior Game (PAX GBG) is a manualized version of the Good Behavior Game (GBG) that applies additional kernels and cues in comparison with the original GBG.

TABLE 1 Overview of implementation quality indicators and types of contextual factors studied.

Study ID	References	Program	Program type (curriculum vs. kernel)	Implementation quality indicators studied	Type of contextual factors studied
1	Gregory et al., 2007	Yes I can	Curriculum	Dosage	School organizational
2**	Beets et al., 2008	Positive action	Curriculum	Dosage Adherence Quality of delivery	School organizational Teacher attitudes
3	Mihalic et al., 2008	LifeSkills training	Curriculum	Dosage Adherence Quality of delivery Participant responsiveness	Program support: training Program support: coaching School organizational Teacher attitudes Teacher resources Student behavioral *Parents*
4**	Ransford et al., 2009	PATHS	Curriculum	Dosage Quality of delivery	Program support: training Program support: coaching School-organizational Teacher resources Teacher demographic Student demographic
5**	Wehby et al., 2012	GBG	Kernel	Adherence	Program support: coaching Teacher attitudes Teacher resources
6	Becker et al., 2013	PAX GBG	Kernel	Dosage Quality of delivery	Program support: coaching Teacher demographic Student demographic
7	Molloy et al., 2013	SWPBIS	Kernel	Quality of delivery	School demographic Student demographic
8	Johnson et al., 2014	GBG	Kernel	Adherence Quality of delivery	Teacher attitudes
9	Domitrovich et al., 2015	PAX GBG	Kernel	Dosage Quality of Delivery	Program support: coaching School organizational School demographic Teacher attitudes Teacher resources Teacher demographic Student demographic
10	Malloy et al., 2015	Positive action	Curriculum	Dosage Quality of delivery	School organizational Teacher attitudes
11**	Wanless et al., 2015	Responsive Classroom	Curriculum and Kernel	Adherence	Program support: training School organizational Teacher attitudes Teacher resources Teacher demographic
12	Bethune, 2017	SWPBIS	Kernel	Adherence	Program support: coaching
13	Swift et al., 2017	KiVa	Curriculum	Dosage	School organizational Teacher attitudes Teacher resources
14**	Johnson et al., 2018	PAX GBG	Kernel	Dosage Quality of delivery	Program support: coaching
15	Domitrovich et al., 2019	PATHS	Curriculum	Dosage Quality of delivery	School organizational School demographic Teacher attitudes Teacher resources Teacher demographic Student demographic

(Continued)

TABLE 1 (Continued)

Study ID	References	Program	Program type (curriculum vs. kernel)	Implementation quality indicators studied	Type of contextual factors studied
16	Musci et al., 2019	PAX GBG	Kernel	Dosage	School organizational Teacher resources Teacher demographic Student behavioral Student demographic
17	Bastable et al., 2020	SWPBIS	Kernel	Adherence	Program support: coaching
18*,***	Dowling and Barry, 2020	MindOut	Curriculum	Dosage Adherence Quality of delivery Participant responsiveness ***	Program support: coaching School organizational Teacher attitudes Student behavioral Student attitudes Student demographic
19	Combs et al., 2021	LifeSkills training	Curriculum	Dosage Adherence Quality of delivery Participant responsiveness	Program support: training School organizational School demographic Teacher attitudes Student demographic
20	Combs et al., 2022	LifeSkills training	Curriculum	Adherence Quality of delivery Participant responsiveness	School organizational School demographic Teacher attitudes Student behavioral Student demographic

*Mixed methods design; **interaction effects of contextual variables on quality indicators reported, ***four implementation quality indicators used for calculating an overall quality index.

by two coders and discussed in detail until a unanimous decision was met, thus establishing inter-rater reliability. Data from the remainder of the articles were coded by the first author of the study and afterward carefully double-checked by the second coder.

In the second step, after reviewing all articles, one single *post hoc* datasheet was formed with (a) the authors' rationale for including the contextual variables, (b) all contextual variable characteristics, (c) every relationship of the contextual variable with the implementation quality indicator, (d) the authors' interpretation of the relevance of the results, (e) rationale for further research and (f) descriptions of study limitations. This analysis sheet was compiled by the first author of the study and afterward thoroughly double-checked by the second coder.

In the third step, the inductive content analysis procedure (Vears and Gillam, 2022) was employed for all the contextual variables to be grouped into categories based on similarity. Four main categories of contextual factors emerged: program support (i.e., training and coaching), school, teacher, and student categories, with 2–3 subcategories for each that further withheld subcategories, based on similarity. The coding result of categories and subcategories can be seen in **Figure 2**. In one study, one parental support related single item was coded as a contextual variable, but as it did not yield any statistically significant effect on any implementation quality indicator, it was excluded from further analysis. All coded category and subcategory labels were then linked with the contextual variables in the *post hoc* datasheet. Relations in the mixed methods study were additionally coded if reported as particularly characteristic to the low (n = 12) or high quality implementation group (n = 8). Similarities between the two quality groups in the mixed methods study were not coded, due to their non-differential nature.

In the fourth step, in order to ascertain, which contextual factors would demonstrate the highest relative frequency of statistical significance toward SEL program implementation quality indicators, a quantitative approach was undertaken. Frequency ratios of statistically significant relationships were calculated across all quantitative articles (n = 19). Altogether 355 relationships were tested in the 19 quantitative articles and around a quarter (83) of these relationships were statistically significant ($p < 0.05$ or smaller). The largest number of tested relationships was present in the school category (113), followed by the teacher (104), program support (71), and student categories (63) (refer to **Table 2**). As an illustrating example, in Ransford et al. (2009, study 4), among other contextual variables, two different indicators of teacher resources (self-efficacy and burnout) were both studied for their relationships with dosage and quality of delivery, the latter both measured through two different indicators, which altogether presented 8 relationships tested, of which two turned out to be statistically significant in the study. Studies looked at a different number of relationships, ranging from 1 (study 12) to 51 (studies 9 and 20), and the frequency of statistically significant relations

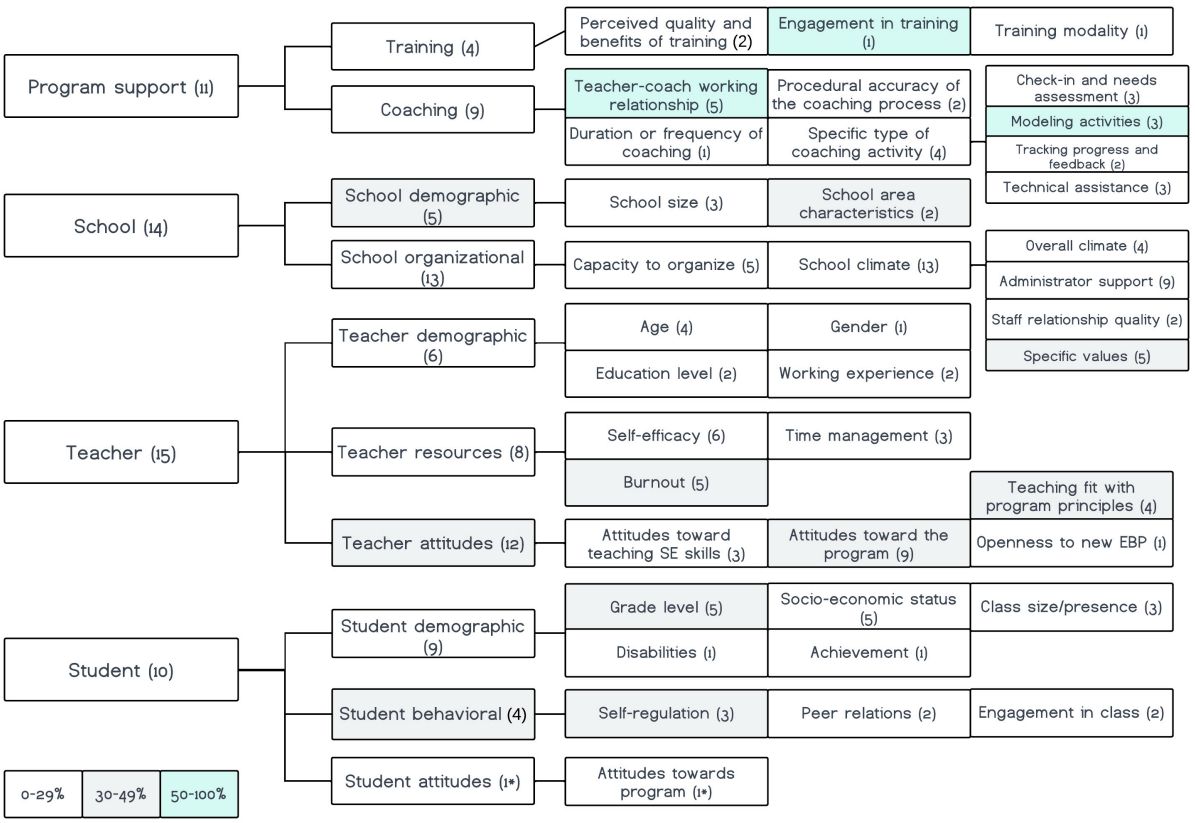


FIGURE 2
Results of the thematic coding process: an overview of school-related contextual factors, four category levels. The figure in brackets indicates the number of articles this factor was explored in. *Category variables only emerged in study 18 (Dowling and Barry, 2020) and are, therefore, not present in **Table 2**. Color codes indicate the frequency of the statistically significant relationships from tested relationships in the quantitative studies (N = 19).

from those tested, ranged from 6 (study 9) to 100% (studies 2, 7, and 12). The results of the relationship significance ratio calculations are presented in **Table 2**. Relationships with all four outcome variables of implementation quality (adherence, dosage, quality of delivery, and participant responsiveness) were studied in the context of all four categories of contextual factors, no implementation quality indicator appeared more frequently studied across any category of contextual factors.

4. Results

4.1. Categories of contextual factors

In the process of the thematic coding that was carried out for identifying unifying categories among the contextual variables, four main categories of contextual variables were revealed (refer to **Figure 2**): program support (i.e., training and coaching), school, teacher, and student category contextual variables. The different categories were present rather equally in articles, teacher category variables were looked at in 15 studies, school variables in 14, and program support variables in 11 articles. Student category factors were considered in the least number of articles (10). Four studies (1, 8, 12, and 14) looked at contextual variables in just one main category (e.g., school), whereas five studies (3, 4, 9, 18, and 19) handled background predictors in all four. The four main categories (level 1) were divided into smaller subcategories into three additional levels, revealing a diverse array of contextual variables tested (follow the subdivision of categories on the four subcategory levels both in **Figure 2** and in **Table 2**).

On level two, teacher-related factors were diversely coded into three subcategories: teacher demographics (six articles), teacher resources (eight articles), and teacher attitudes (12 articles). The demographics subcategory included factors of age (four articles), gender (one article), education level (two articles), and working experience (two articles) on level three. Teacher resources were divided on level three into such psychological resources as (a) self-efficacy (six studies), (b) burnout (five studies), and (c) time management (three articles). Teacher attitudes, in turn, were assessed either toward teaching SE skills (three studies), toward the specific SEL program

TABLE 2 Proportion of statistically significant relationships between contextual factors and implementation quality indicators by category (quantitative articles, *N* = 19).

Level 1 main categories	A	B	C	D	Level 2 subcategories	A	B	C	D	Level 3 subcategories	A	B	C	D	LEVEL 4 subcategories	A	B	C	D
Program support	71	19	26.8	4, 5, 6, 11, 12, 14, 17, 19	Training	17	4	23.5	4, 11, 19	Perceived quality and benefits of training	8	2	25.0	4					
										Engagement in training	**1**	**1**	**100**	11					
										Training modality	8	1	12.5	19					
					Coaching	54	15	27.8	4, 5, 6, 12, 14, 17	Specific type of coaching activity	34	7	20.6	6, 14, 17	Check-in and needs assessment	8	0		
										Teacher-coach working relationship	**10**	**6**	**60.0**	4, 5, 14	**Modeling activities**	**5**	**4**	**80.0**	**6, 14, 17**
										Duration or frequency of coaching	5	1	20.0	6	Tracking progress and feedback	10	2	20.0	6, 17
										Procedural accuracy of the coaching process	5	1	20.0	12	Technical assistance	11	1	9.1	17
School	113	29	25.7	1, 2, 4, 7, 10, 15, 19, 20	School organizational	92	22	23.9	1, 2, 4, 10, 15, 19, 20	School climate	56	13	23.2	1, 2, 4, 10, 15, 19	Overall climate	12	2	16.7	2
										Capacity to organize	36	9	25.0	15, 19, 20	Administrator support	23	4	17.4	1, 4, 19
					School demographic	21	7	33.3	7, 15, 19, 20	School size	7	2	28.6	7, 15	Staff relationship quality	4	1	25.0	10
										School area characteristics	14	5	35.7	19, 20	Specific values	17	6	35.3	1, 10, 15
Teacher	104	25	24.0	2, 3, 4, 5, 8, 9, 10, 13, 15, 16, 20	Teacher demographic	20	3	15.0	4, 9, 15	Age	9	2	22.2	4, 9					
										Gender	3	0	0						
										Education level	4	0	0						
										Working experience	4	1	25.0	15					
					Teacher resources	34	5	14.7	4, 9, 13, 16	Self-efficacy	13	1	7.7	4					
										Burnout	10	4	40.0	4, 9, 13, 16					
										Time management	11	0							
					Teacher attitudes	50	17	34.0	2, 3, 5, 8, 9, 10, 15, 20	Attitudes toward teaching SE skills	7	2	28.6	2, 10					

(Continued)

10.3389/feduc.2022.965538

TABLE 2 (Continued)

Level 1 main categories	A	B	C	D	Level 2 subcategories	A	B	C	D	Level 3 subcategories	A	B	C	D	LEVEL 4 subcategories	A	B	C	D
										Attitudes toward the program	29	11	37.9	3, 5, 8, 15, 20					
										Openness to new EBP	3	0	0						
										Teaching fit with program principles	11	4	36.4	9, 20					
Student	63	15	23.8	3, 4, 7, 16, 19, 20	Student demographic	47	9	19.2	4, 7, 19, 20	Grade level	10	4	40.0	4, 7					
										Socio-economic status	17	5	29.4	7, 19, 20					
										Class size/presence	14	0	0						
										Disabilities	3	0	0						
										Achievement	3	0	0						
					Student behavioral	16	6	37.5	3, 16, 20	Self-regulation	13	6	46.2	3, 16, 20					
										Peer relations	2	0	0						
										Engagement in class	1	0	0						

A, number of statistical relationships tested in this category; B, number of statistically significant relationships revealed in the category; C, percentage of statistically significant relationships from all those tested; D, list of studies where this category variable was found statistically significant in at least one relationship; ratios over 50% are boldfaced.

10.3389/feduc.2022.965538

implemented (nine studies), teaching fit with the program principles (four articles), or openness to employing a new evidence-based practice (EBP, 1 study).

The program support factors were coded into two subcategories on level two: coaching (nine articles) and training (four articles). Training variables on level 3 looked at perceived benefits of training (two articles), teacher engagement during training (one article), and training modality (online vs. face to face, one article). Coaching variables on level three were related to either (a) coach-teacher working relationship (teacher self-report, five articles), (b) the duration or frequency of coaching (one article), (c) the procedural accuracy of coaching as to abide by a certain coaching protocol (two articles), and (d) the specific type of coaching activity applied by the coach (four articles). The specific coaching activities on level four included: a) check-in and needs assessment (three articles), modeling program activities (three articles), tracking progress and feedback (two articles), and technical assistance (three articles).

School factors were coded into two categories on level two: school-organizational (13 articles) and school demographic (five articles), which again were further split into smaller subcategories of thematically grouped variables on level three. In the school demographic factors subcategory, there were two types of variables—school size (number of students, three articles), and school area characteristics, such as rural/suburban or the number of schools engaged in the SEL program implementation in the area (two articles). School organizational factors were coded into two different aspects: school climate indicators (13 articles), and the capacity to organize, which included factors related to providing facilities, materials, or cooperation structures for SEL program implementation in the school (five articles). Within the school climate category, four subcategories were distinguished on level four: (a) general school organizational climate (measured by the overall score of different organizational climate subscales, four articles); or more specific subscales of organizational culture, such as (b) administrative support and/or leadership (nine articles), (c) staff relationship quality (two articles), or (d) the level of particular organizational values such as openness to innovation, support, and respect, perceived collective responsibility, participatory decision-making, or school-wide support for SEL (five studies).

Students were the least frequently studied category and three level 2 subcategories here were student demographic, student behavioral, and student attitudes (the latter was only addressed in one article, Study 18). The demographics subcategory was coded into (a) grade level (five articles), (b) socio-economic status (five articles), (c) class size or presence in class (three articles), (d) percentage of students with disability (one article), and (e) percentage of students proficient in state achievement tests (one article). The student behavioral subcategory was coded into self-regulation (three articles), peer relations (two articles), and engagement in class (two articles).

In response to the first research question, we found that articles most frequently looked at school climate indicators (13 articles), including measures of perceived administrator support (nine articles); teacher attitudes toward the program implemented (nine articles); and teacher self-efficacy (six articles). Overall, a dispersed picture of diverse types of contextual factors emerges from **Figure 2**, as many kinds of contextual factors only surfaced in one article (such as the duration and frequency of coaching, engagement in training, or openness to new EBP).

4.2. Contextual factors demonstrating the highest frequency of statistical significance

The second research question was interested in which contextual factors would demonstrate the highest frequency of statistically significant relationships with SEL program implementation quality indicators. The relative frequency of statistically significant relationships across all four main categories was somewhat similar, ranging from 23.8% in the student category to 26.8% in the program support category (refer to **Table 2**). The proportion of statistically significant relationships found, however, varied greatly within the subcategories. Throughout most categories a statistically significant result was present between 0 and 29% of the time (**Figure 2**), indicating that these types of contextual factors showed more arbitrary statistical significance across several studies.

In 10 categories, the frequency of statistically significant relationships from those tested, ranged between 30 and 49% (marked with gray in **Figure 2**), indicating a somewhat more consistent tendency for statistical significance across studies. Only three subcategories stood out across studies by displaying a ratio of 50% or higher in the tested vs. statistically significant relationships, all of which emerged from the program support category (marked with light blue in **Figure 2**). Column D in **Table 2** lists studies where a significant relationship was found between that category of contextual factor and an implementation quality indicator; articles that tested, yet failed to find a statistically significant relationship in that category were excluded from **Table 2**.

Within the program support category, both, training and coaching appear to be significant contextual factors for implementation quality. Training-related contextual variables were not copiously operationalized in the studies. Of special interest was the observer-rated teacher engagement in training, which was the only contextual variable of seven in Wanless et al. (2015, study 11), that yielded any statistical significance toward their single implementation quality indicator of interest (adherence). Coaching-related factors were assessed more frequently. In the specific type of coaching activity category, four

different subcategories were distinguished, where only one type of activity was frequently statistically significant toward quality implementation indicators. In all three studies that looked at specific coaching activities' influence on implementation quality indicators (studies 6, 14, and 17) modeling program-specific activities (i.e., demonstrating how to implement the program practices) was the specific activity that yielded a consistent relationship with implementation quality indicators (statistically significant in 80% of the tested relationships, refer to **Table 2**). Furthermore, in the mixed methods study, the low-quality implementers' group was distinctive from the high-quality implementers by suggesting that an external person would deliver the lessons altogether. The teacher-coach working relationship was measured with a single item in study 4 and with a variant of a *teacher-coach alliance scale* in three studies (5, 9, and 14), all collected as teacher self-report of the working relationship and its benefits with the coach. Throughout articles, this was tested through 10 quantitative relationships in four studies and found statistically significant in six relationships in three articles ($p < 0.01$), this makes the teacher-coach working relationship the second more consistent variable in the coaching subcategory, as connected to implementation quality indicators. Additionally, in Dowling and Barry (2020, study 18), high quality implementers expressed more openness to implementation support.

Factors in the school category mostly indicated a modest number of statistically significant relationships. School climate was the most common contextual variable category (explored in 13 articles). Despite its popularity, especially by exploring the influence of perceived administrator support on implementation quality (nine studies) the school organizational climate variables, just as demographic variables did not indicate any consistent statistical significance toward implementation quality indicators.

In the teacher category, no types of variables yielded statistically significant relationships across 50% of the tested relationships. Teacher resources were a rather frequent target of inquiry, altogether 34 quantitative relationships were assessed, of which only five were found statistically significant. Teacher self-efficacy indicators, which are often suggested as relevant contextual factors for implementation quality, were statistically significant in only 7.7% of the relationships tested. Teacher burnout (as measured by the Maslach Burnout Inventory, Maslach et al., 1996) however, was found to be negatively related to dosage (and not to other implementation quality indicators) in four studies (4, 9, 13, and 16), which sets the relative frequency of statistical significance to 66.7% for relationships tested only between teacher burnout and dosage. This was the only regularity in the current sample where some type of contextual variable would be systematically connected to only one kind of implementation quality indicator.

In the student category, similarly, no types of variables were statistically significant more than 50% of the tested relationships. Student-related factors were primarily assessed as demographic factors (nine articles). Student self-regulation was assessed as a contextual factor for implementation quality in three quantitative articles (studies 3, 16, and 20) and showed the most promise (statistically significant in 46.2% of the tested relationships). More specifically, in Mihalic et al. (2008, study 3) observer-rated quality of student behavior in class predicted adherence and quality of delivery; and in Musci et al. (2019, study 16) observer-rated aggressive behavior in class was negatively related to program dosage. In Combs et al. (2022, study 20) observer-rated student misbehavior negatively predicted all three of their implementation quality indicators ($p < 0.001$). Additionally, Dowling and Barry (2020, study 18) was the only article in the sample where both teachers and students were included as informants (*via* interviews) and where student attitudes about the program had a chance to surface, as well as were revealed as distinctive of the quality of program implementation. For example, students in the high implementer group listed more personal benefits of the program, claimed more frequently to have enjoyed the learning experience and could offer more specific examples of SEL benefits to them. Students in the low quality group reported negative program experiences more frequently and could bring less concrete examples of SEL benefits to themselves.

Five articles (studies 2, 4, 5, 11, and 14) viewed the interaction effects of contextual factors on implementation quality indicators. In four of them, interaction effects between variables in different contextual categories were revealed. The type of interaction effect, which was similarly revealed throughout two studies, appeared in the context of perceived burnout and coach-teacher working relationship. Both in Wehby et al. (2012, study 5) and Ransford et al. (2009, study 4), a strong coach—teacher working relationship was seen to reduce the impact of teacher burnout on implementation quality indicators.

5. Discussion

A systematic literature process with a final sample of 20 articles, was carried out to map, what kind of teacher- and school-related contextual factors have been studied in relation to SEL implementation quality indicators; as well as, to see whether any of those contextual factors showed more consistent statistical significance across the articles in the sample.

The current study offers confirmation that different levels of contextual variables are relevant to ensuring implementation quality and further supports an ecological understanding of implementation quality context. Based on the articles in the current sample, four different categories of contextual variables were exposed: student, teacher, school, and program support categories, which, in turn, were divided into quite heterogeneous subcategories into three sublevels (demographic, attitudinal,

behavioral, etc., refer to **Figure 2** and **Table 2**), revealing the diversity in contextual factors studied across SEL program implementation quality. It is important to note that all four broader contextual categories also emerged in the mixed methods study (Dowling and Barry, 2020, study 18) where they surfaced through qualitative interviews, not variables operationalized beforehand.

The category of program support factors was the only one in the current study where more consistent statistically significant contextual factors were revealed; and which, based on the current analysis, may thus prove more essential for ensuring the program outcomes for students. The contextual factor with the highest frequency of statistical significance (80%) across articles, was a single kind of coaching activity: modeling program activities to teachers, revealed in studies looking at coaching activities for implementing GBG (Becker et al., 2013, study 6), PAX GBG (Johnson et al., 2018, study 14) and SWPBIS (Bastable et al., 2020, study 17). Modeling has previously been shown as something that supports teacher self-efficacy, especially during the beginning phases of the profession (Bandura, 1977; Johnson, 2010), and can thus be seen as supportive for teachers in adopting new practices. Second, teacher self-report of the working relationship and its benefits with the coach proved more consistently significant (in 60% of the tested relationships in the sample) for ensuring implementation quality for PATHS (Ransford et al., 2009, study 4), GBG (Wehby et al., 2012, study 5), and PAX GBG (Johnson et al., 2018, study 14). It is noteworthy that none of the "procedural" qualities (procedural accuracy of the process, progress feedback, check-in, or time spent coaching) managed this. Study 5 describes the role of the coach as the link between teachers and project staff, offering feedback and assistance and providing program materials; study 15 describes coaching in more collaborative and tailored terms in supporting teachers' program implementation skill development. Study 4 does not offer a description of the coaching principles applied in implementation. Studies 5 and 14 used a teacher-coach alliance scale (10 and 23 items, respectively), whereas study 4 managed to obtain statistically significant results through the use of a single item. "Overall, how useful was the consultation time with your PATHS coordinator." Issues of measuring the transactional nature of the coach-teacher working relationship have been discussed in Johnson et al. (2016); study 4, however, indicates, that this contextual factor may also be captured by a single item. The positive impact of coaching on program implementation quality is not a large surprise, as coaching has been shown as an efficient measure for teacher professional development and desired classroom impact (Kraft et al., 2018); the current study, suggests an emphasis on the cooperative or relational aspect of this working alliance, as opposed to its technical aspects. However, as coaching is also costly and the benefits may be short-lived after the program implementation coaching phase had ended (Pas et al., 2022)—the question about longer-term student benefits through quality of teacher implementation practice remains.

A frequently significant relationship between teacher burnout and dosage (and no other implementation quality indicators) was revealed in the current article, suggesting a more systematic pattern between the exhaustion of psychological resources and the amount of program delivered. Additionally, it was shown in studies 4 and 5 that a quality coach-teacher working relationship had the potential to reduce the effects of teacher burnout on program implementation. Research by Ghasemi (2021, 2022) has shown that individual motivational and empowering interventions (similar to coaching) can effectively reduce teachers' burnout levels. Even though among high-coping teachers, the effects of burnout on everyday practice may not be detrimental, burnout combined with the effects of teacher stress has an evident impact on student outcomes (Herman et al., 2018)—thus coaching in the form of teacher support may also alleviate the risks teacher burnout presents to classroom practice.

The student self-regulation subcategory did not pass the 50% frequency mark in statistical significance (it stayed at 46.2%), but showed potential, as it was statistically significant in all three studies that addressed it. It must be added that In Combs et al. (2021, study 19), a "capacity to organize" indicator[4] was measured with a checklist of observer-rated technical and organizational difficulties (e.g., lack of materials, poor facilities), which also included a student disruptive behavior item. This composite factor was predictive of three implementation quality indicators on the teacher level, but the contribution of the observed misbehavior to that effect can only be hypothesized. In their second similar study, Combs et al. (2022, study 20) assessed all observer-rated technical and behavioral difficulty factors as separate contextual variables, and both, capacity to organize and student self-regulation factors were statistically significant contextual predictors in the study (especially, again, on the teacher-, as opposed to the school district level). This might also suggest a hidden effect of observed student self-regulation in study 19. We, thus, suggest that student baseline self-regulation deserves additional attention as a contextual factor in implementation quality research. Musci et al. (2019) pointed out that too few studies had considered the students' behavioral influences on program implementation by teachers and suggested that there could be a "tipping point" where negative student behavior could present a challenge to implementation. Farmer et al. (2016) have pointed to "correlated constraints" as the network of synergistic associations between individual and social factors in group functioning; which makes specific student contingencies in the

4 The only mixed contextual indicator in the sample. This was coded as "capacity to organize" due to the nature of the large majority of items in the observer list.

classroom and their interaction also an important contextual factor for program implementation, and these should be looked at in further research.

In the current study, some levels of school ecology that have been previously regarded as relevant for ensuring program implementation quality did not prove as consistently significant as might have been anticipated (such as a teacher or school level contextual factors). As an example, the school organizational climate was the most common level 2 background factor category in the current sample (13 articles), assessed with general organization culture measures, as well as more specific assessments of organizational values or administrator support (both generally and to program implementation), that did not yield any consistently significant results. It can be seen that the *school* (Domitrovich et al., 2008) and *organizational capacity* (Durlak and DuPre, 2008) levels in previous models also contain a dispersed array of organization-related factors (from organizational norms to ways of decision making) to influence implementation quality. The tradition of including organizational variables can be traced to the ISF which sees different organizational characteristics play an active role in the program implementation process (Wandersman et al., 2008). Despite the undisputed importance of such hygiene factors for school daily life, the current article could not confirm the consistent relevance of those factors for SEL program implementation quality across studies.

On the contrary, the program support category was revealed as a more consistently convincing contextual layer for supporting the quality implementation of SEL programs, in the current study. Even though teachers are frequently referred to as central players in program implementation (e.g., Brackett et al., 2012; Schonert-Reichl, 2017), based on the current study, the quality of program implementation may rely less heavily on teachers' demographic characteristics, attitudes, or personal resources, and more on something that happens in the "zone" of their professional development. Coaching has a long history of being viewed as an essential part of teacher professional development, that enables the transfer of acquired skills and knowledge into practice (Joyce and Showers, 1981). In Domitrovich et al. (2008) *support system* is seen as more inherent to the intervention and its implementation process. It may be worth considering treating program support through training and coaching as a separate contextual layer of teacher professional development, designed to induce the desired change in teacher everyday practice.

Furthermore, there is an additional contextual category that has not been suggested explicitly in previous theoretical models and has also been included more infrequently in previous research—the student level. The relative scarcity of studies examining student-level contextual factors could be explained by the tradition of evidence-based programming, where students' behavior has rather been seen as a result, not the context of program implementation. Students, however, bring baseline behavioral qualities to the interaction that may impact the implementation process. Students' behavioral factors that are nested in classrooms could also be regarded on the individual level in the school ecology, to interact with the teacher level, who in turn is influenced by the professional development (program support). Tolan et al. (2020) provide support for this idea, as in their integration trial of GBG and My Teaching Partner^{TM} interventions, an effect on student outcomes was observed in the interaction of student behavior, teacher personal resources, and professional development variables.

5.1. Study limitations

There were also several limitations to the study. The first limitation was the relatively small number of articles in the study that remained in the sample after screening. The second limitation is that the current article is affected by both: publication and journal bias. First, only journal articles were considered for the sample, and only *Prevention Science* and *Journal of School Psychology* were sampled separately. This could indicate that some relevant information may still have remained unexposed for the sample, also it is possible, that research that concentrates implicitly on the contextual factors' influence on SEL program implementation is still relatively limited to this day. Third, we see that there is unclarity in SEL program terminology: programs may be called universal prevention programs or classroom management programs in different contexts, which may or may not include teaching and practice of SEL skills. This terminological unclarity may have contributed to the failure to include all relevant articles for the sample. In comparison to an overview of 20 articles, the current study only offers an initial map of the contextual factors and their relevance to implementation quality, several contextual factors were only assessed in just one article and more confirmation would be needed for the relevance of such factors across studies. One such factor was teacher engagement in training that predicted implementation quality. Even though this was a convincing contextual factor in Wanless et al. (2015, study 11), more studies would be needed to confirm the relevance of this factor across studies. The fourth limitation is that it was decided not to carry out any additional quality appraisal of the articles in the sample and that publication in peer-reviewed journals was considered a sufficient benchmark for research quality. It was, however, noticed that many indicators in the frame of the same study may have been assessed with varying degrees of quality (e.g., self-reported single item with low variance, as opposed to observer ratings with high interrater agreement), which could account for the nature of some statistical relationships reported in the sample and should be considered in more detail in further studies. The jingle-jangle fallacy could be seen as one limitation for interpreting the results of this articles, Dane and Schneider (1998) have pointed out that "inconsistencies in the conceptualization of fidelity reduce interpretability of studies." However, implementation quality indicators have been

proposed to be considerably interconnected (Beets et al., 2008; Berkel et al., 2011) and inferences were generally not made, based on the type of implementation variable measured, in the current article.

5.2. Implications of the study

The current article offers an overview, a more organized road map of studies looking into the contextual factors of program implementation which support students' SE skill development in school, and as such, it presents an original contribution in the context of SEL program implementation research and discussion. The analytical frame of the study has suggested four broad contextual categories to support or hinder quality SEL implementation, as such, contributing to the theoretical development of the field. The results of the current study bolster the relevance of program support factors for implementation quality and reveal a link between teacher burnout and program dosage. Based on the current analysis, student factors emerge as a separate contextual level in school, with special attention to student baseline self-regulation that may influence SEL program implementation quality. Scientists interested in practical research assuring quality contexts or SEL implementation can utilize this knowledge as a navigation tool toward more or less promising avenues for program implementation support.

The current study suggests an emphasis on teacher professional development and support in SEL program implementation, concurring with Cook et al. (2019) emphasis on *educator support* in the guidelines for effective school-based implementation. Despite many SEL programs providing coaching during initial implementation (e.g., Hershfeldt et al., 2012; Becker et al., 2013), these effects may not be sustained over time (Pas et al., 2022). Longer term coaching initiatives may be considered internally for schools, such as peer coaching or professional learning communities (Timperley et al., 2007; Cook et al., 2019; Elias, 2019). Effective coaching partnerships require teachers to have skills like the reflection of professional practice, which may receive little attention in initial teacher training (Pas et al., 2014) or be learned and practiced in a dubious manner (Marcos et al., 2011). Initial teacher education should, thus, find efficient ways of promoting teacher reflective practice and inquiry mindset (Muijs et al., 2014) so that teachers would already be more equipped with those skills when incorporating new evidence-based practices into their work.

Studying SEL implementation is a tortuous area of study, therefore, such issues as the intricateness of implementation research (Durlak, 2015), the need for theoretical integrity (Jones et al., 2019), or measurement challenges (McKown, 2019) have been previously discussed. Further research is needed that examines both individual and organizational factors on different ecological levels (e.g., teacher, school, and school district), that impact implementation as an outcome (e.g., Domitrovich et al., 2019; Combs et al., 2022). However, research should also address, what accounts for the discrepancies of such findings across different studies and SEL implementation contexts: could they be accounted to the different programs implemented, differing conceptualization and operationalization of implementation quality indicators, or, instead, varying ways of operationalizing or assessing contextual factors. Based on the current study, we would also like to support the further application of mixed methods research on the field, as it may allow for additional contextual factors (such as student attitudes in Dowling and Barry, 2020) to surface, which could have been underrepresented in some previous multi-level models that often guide quantitative research.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

TU was responsible for the systematic literature review process and the analysis and recruiting a second expert as the second rater. KP-V performed the literature review and analysis. Both authors contributed to the article and approved the submitted version.

Funding

This study was supported by the EEA Financial Mechanism 2014–2021 and Higher Education in Baltic Research Programme (36.1-3.4/289).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Appleyard, K., Egeland, B., van Dulmen, M. H. M., and Sroufe, L. A. (2005). When more is not better: the role of cumulative risk in child behavior outcomes. *J. Child Psychol. Psychiatr.* 46, 235–245. doi: 10.1111/j.1469-7610.2004.00351.x

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychol. Rev.* 84, 191–215. doi: 10.1037/0033-295X.84.2.191

Bastable, E., Massar, M. M., and McIntosh, K. (2020). A Survey of Team Members' Perceptions of Coaching Activities Related to Tier I SWPBIS Implementation. *J. Pos. Beh. Inter.* 22, 51–61. doi: 10.1177/1098300719861566

Becker, K. D., Bradshaw, C. P., Domitrovich, C., and Ialongo, N. S. (2013). Coaching Teachers to Improve Implementation of the Good Behavior Game. *Adm. Policy Ment Health* 40, 482–493. doi: 10.1007/s10488-013-0482-8

Beets, M. W., Flay, B. R., Vuchinich, S., Acock, A. C., Li, K.-K., and Allred, C. (2008). School Climate and Teachers' Beliefs and Attitudes Associated with Implementation of the Positive Action Program: a Diffusion of Innovations Model. *Prev. Sci.* 9, 264–275. doi: 10.1007/s11121-008-0100-2

Berkel, C., Mauricio, A. M., Schoenefelder, E., and Sandler, I. N. (2011). Putting the Pieces Together: an Integrated Model of program Implementation. *Prev. Sci.* 12, 23–33. doi: 10.1007/s11121-010-0186-1

Bethune, K. S. (2017). Effects of Coaching on Teachers' Implementation of Tier I School-Wide Positive Behavioral Interventions and Support Strategies. *J. Pos. Beh. Inter.* 19, 131–142. doi: 10.1177/1098300716680095

Brackett, M. A., Reyes, M. R., Rivers, S. E., Elbertson, N. A., and Salovey, P. (2012). Assessing Teachers' Beliefs about Social and Emotional Learning. *J. Psychoeduc. Assess.* 30, 219–236. doi: 10.1177/0734282911424879

Burgess, L. G., Riddell, P. M., Fancourt, A., and Murayama, K. (2018). Contagion Within Education: a Motivational Perspective. *Mind Brain Educ.* 12, 164–174. doi: 10.1111/mbe.12178

Calhoun, B., Williams, J., Greenberg, M., Domitrovich, C., Russell, M. A., and Fishbein, D. H. (2020). Social Emotional Learning Program Boosts Early Social and Behavioral Skills in Low-Income Urban Children. *Front. Psychol.* 11:561196. doi: 10.3389/fpsyg.2020.561196

Cantor, P., Osher, D., Berg, J., Steyer, L., and Rose, T. (2019). Malleability, plasticity, and individuality: how children learn and develop in context. *Appl. Dev. Sci.* 23, 307–337. doi: 10.1080/10888691.2017.1398649

Century, J., Rudnick, M., and Freeman, C. (2010). A Framework for Measuring Fidelity of Implementation: a Foundation for Shared Language and Accumulation of Knowledge. *Am. J. Eval.* 31, 199–218. doi: 10.1177/1098214010366173

Combs, K. M., Buckley, P. R., Lain, M. A., Drewelow, K. M., Urano, G., and Kerns, S. E. U. (2022). Influence of Classroom-Level Factors on Implementation Fidelity During Scale-up of Evidence-Based Interventions. *Prev. Sci.* 23, 969–981. doi: 10.1007/s11121-022-01375-3

Combs, K. M., Drewelow, K. M., Habesland, M. S., Lain, M. A., and Buddey, P. R. (2021). Does Training Modality Predict Fidelity of an Evidence-based Intervention Delivered in Schools? *Prev. Sci.* 22, 928–938. doi: 10.1007/s11121-021-01227-6

Cook, C. R., Low, S., Buntain-Ricklefs, J., Whitaker, K., Pullmann, M. D., and Lally, J. (2018). Evaluation of Second Step on Early Elementary Students' Academic Outcomes: a Randomized Controlled Trial. *Sch. Psychol. Q.* 33, 561–572. doi: 10.1037/spq0000233

Cook, C. R., Lyon, A. R., Locke, J., Waltz, T., and Powell, B. J. (2019). Adapting a compilation of implementation strategies to advance school-based implementation research and practice. *Prev. Sci.* 20, 914–935.

Cooper, H. (2017). *Research Synthesis and Meta-Analysis. A Step-by-Step Approach Fifth Edition*. Thousand Oaks, CA: SAGE Publications INC. doi: 10.4135/9781071878644

Corcoran, R. P., Cheung, A. C. K., Kim, E., and Xie, C. (2017). Effective universal school-based social and emotional learning programs for improving academic achievement: a systematic review and meta-analysis of 50 years of research. *Educ. Res. Rev.* 25, 56–72. doi: 10.1016/j.edurev.2017.12.001

Dane, A. V., and Schneider, B. H. (1998). Program Integrity in Primary and Early Secondary Prevention: are Implementation Effects out of Control? *Clin. Psychol. Rev.* 18, 23–45. doi: 10.1016/S0272-7358(97)00043-3

Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., et al. (2008). Maximizing the Implementation Quality of Evidence-Based Preventive Interventions in Schools: a Conceptual Framework. *Adv. Sch. Ment Health Promot.* 1, 6–28. doi: 10.1080/1754730X.2008.9715730

Domitrovich, C. E., Pas, E. T., Bradshaw, C. P., Becker, K. D., Keperling, J. P., Embry, D. D., et al. (2015). Individual and School Organizational Factors that Influence Implementation of the PAX Good Behavior Game Intervention. *Prev. Sci.* 16, 1064–1074. doi: 10.1007/s11121-015-0557-8

Domitrovich, C. E., Yibing, L., Mathis, E. T., and Greenberg, M. (2019). Individual and organizational factors associated with teacher self-reported implementation of the PATHS curriculum. *J. Sch. Psychol.* 76, 168–185. doi: 10.1016/j.jsp.2019.07.015

Dowling, K., and Barry, M. M. (2020). Evaluating the Implementation Quality of a Social and Emotional Learning Program: a Mixed Methods Approach. *Int. J. Environ. Res. Public Health* 17, 3249–3265. doi: 10.3390/ijerph17093249

Durlak, J. A. (2015). Studying Program Implementation Is Not Easy but It Is Essential. *Prev. Sci.* 16, 1123–1127. doi: 10.1007/s11121-015-0606-3

Durlak, J. A. (2016). Programme implementation in social and emotional learning: basic issues and research findings. *Cambridge J. Educ.* 46, 333–345. doi: 10.1080/0305764X.2016.1142504

Durlak, J. A., and DuPre, E. (2008). Implementation Matters: a Review of Research on the Influence of Implementation on Program Outcomes and the Factors Affecting Implementation. *Am. J. Community Psychol.* 41, 327–350. doi: 10.1007/s10464-008-9165-0

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

Dusenbury, L., Brannigan, R., Falco, M., and Hansen, W. B. (2003). A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Educ. Res.* 18, 237–256. doi: 10.1093/her/18.2.237

Elias, M. J. (2019). What If the Doors of Every Schoolhouse Opened to Social-Emotional Learning Tomorrow: reflections on How to Feasibly Scale Up High-Quality SEL. *Educ. Psychol.* 54, 233–245. doi: 10.1080/00461520.2019.1636655

Farmer, T. W., Dawes, M., Alexander, Q., and Brooks, D. S. (2016). "Challenges Associated with Applications and Interventions: correlated Constraints, Shadow of Synchrony, and Teacher/ Institutional Factors that Impact Social Change," in *Handbook of Social Influences in School Contexts. Social-Emotional, Motivation, and Cognitive Outcomes*, eds K. R. Wentzel and G. B. Ramani (New York, NY: Routledge), 435–449.

Ghasemi, F. (2021). EFL teachers' burnout and individual psychology: the effect of an empowering program and cognitive restructuring techniques. *Curr. Psychol.* doi: 10.1007/s12144-021-01368-5 [Epub ahead of print].

Ghasemi, F. (2022). An Adlerian-Based Empowering Intervention Program with Burned-Out Teachers. *J. Educ.* 202, 355–364. doi: 10.1177/0022057421998331

Gregory, A., Henry, D. B., and Schoeny, M. E. (2007). School Climate and Implementation of a Preventive Intervention. *Am. J. Community Psychol.* 40, 250–260. doi: 10.1007/s10464-007-9142-z

Herman, K. C., Hickmon-Rosa, J., and Reinke, W. M. (2018). Teacher Stress, Burnout, Self-Efficacy, and Coping and Associated Student Outcomes. *J. Pos. Beh. Inter.* 20, 90–100. doi: 10.1177/1098300717732066

Hershfeldt, P. A., Pell, K., Sechrest, R., Pas, E. T., and Bradshaw, C. P. (2012). Lessons Learned Coaching Teachers in Behavior Management: the PBISplus Coaching Model. *J. Educ. Psychol. Consult.* 22, 280–299. doi: 10.1080/10474412.2012.731293

Humphrey, N., Barlow, A., and Lendrum, A. (2018). Quality Matters: implementation Moderates Student Outcomes in the PATHS Curriculum. *Prev. Sci.* 19, 197–208. doi: 10.1007/s11121-017-0802-4

Johnson, D. (2010). Learning to Teach: the Influence of a University-School Partnership Project on Pre- Service Elementary Efficacy for Literacy Instruction Teachers'. *Read. Horizon.* 50, 23–48. .

Johnson, L. A., Wehby, J. H., Symons, F. J., Moore, T. C., Maggin, D. M., and Sutherland, K. S. (2014). An Analysis of Preference Relative to Teacher Implementation of Intervention. *J. Spec. Educ.* 48, 214–224. doi: 10.1177/0022466913475872

Johnson, S. R., Pas, E. T., and Bradshaw, C. P. (2016). Understanding and Measuring Coach-Teacher Alliance: a Glimpse Inside the Black Box. *Prev. Sci.* 17, 439–449. doi: 10.1007/s11121-016-0633-8

Johnson, S. R., Pas, E. T., Bradshaw, C. P., and Ialongo, N. S. (2018). Promoting Teachers' Implementation of Classroom-Based Prevention Programming Through Coaching: the Mediating Role of the Coach-Teacher Relationship. *Adm. Policy Ment Health* 54, 404–416. doi: 10.1007/s10488-017-0832-z

Jones, S. M., and Bouffard, S. M. (2012). Social and emotional learning in schools: from programs to strategies. *Soc. Policy Rep.* 26, 3–22. doi: 10.1002/j.2379-3988.2012.tb00073.x

Jones, S. M., Brown, J. L., and Aber, J. L. (2011). The longitudinal impact of a universal school-based social-emotional and literacy intervention: an experiment in translational developmental research. *Child Dev.* 28, 533–554. doi: 10.1111/j.1467-8624.2010.01560.x

Jones, S. M., Brush, K. E., Bailey, R., Brion-Meisels, G., McIntyre, J., Kahn, J., et al. (2017). *Navigating SEL from the Inside out: Looking Inside & Across 25 Leading SEL Programs: A Practical Resource For Schools And Ost Providers*. New York, NY: The Wallace Foundation.

Jones, S. M., McGarrah, W., and Kahn, J. (2019). Social and Emotional Learning: a Principled Science of Human Development in Context. *Educ. Psychol.* 54, 129–143. doi: 10.1080/00461520.2019.1625776

Joyce, B. R., and Showers, B. (1981). Transfer of Training: the Contribution of "Coaching". *J. Educ.* 163, 163–172. doi: 10.1177/002205748116300208

Kam, C.-M., Greenberg, M., and Walls, C. T. (2003). Examining the Role of Implementation Quality in School-Based Prevention Using the PATHS Curriculum. *Prev. Sci.* 4, 55–63. doi: 10.1023/A:1021786811186

Kochenderfer-Ladd, B., and Ladd, G. W. (2016). "Integrating academic and social-emotional learning in classroom interactions" in: handbook of Social Influences," in *School Contexts. Social-Emotional, Motivation, and Cognitive Outcomes*, eds K. R. Wentzel and G. B. Ramani (New York, NY: Routledge), 361–378.

Kraft, M. A., Blazar, D., and Hogan, D. (2018). The Effect of Teacher Coaching on Instruction and Achievement: a Meta-Analysis of the Causal Evidence. *Rev. Educ. Res.* 88, 547–588. doi: 10.3102/0034654318759268

Ladd, G. W. (2006). Peer Rejection, Aggressive or Withdrawn Behavior, and Psychological Maladjustment from Ages 5 to 12: an Examination of Four Predictive Models. *Child Dev.* 77, 822–846. doi: 10.1111/j.1467-8624.2006.00905.x

Malloy, M., Acock, A., DuBois, D. L., Vuchinich, S., Silverthorn, N., Ji, P., et al. (2015). Teachers' Perceptions of School Organizational Climate as Predictors of Dosage and Quality of Implementation of a Social-Emotional and Character Development Program. *Prev. Sci.* 16, 1086–1095. doi: 10.1007/s11121-014-0534-7

Marcos, J. M., Sanchez, E., and Tillema, H. H. (2011). Promoting teacher reflection: what is said to be done. *J. Educ. Teach.* 37, 21–36. doi: 10.1080/02607476.2011.538269

Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A. J. S., Abduljabbar, A. S., et al. (2012). Classroom Climate and Contextual Effects: conceptual and Methodological Issues in the Evaluation of Group-Level Effects. *Educ. Psychol.* 47, 106–124. doi: 10.1080/00461520.2012.670488

Maslach, C., Jackson, S. E., and Schwab, R. E. (1996). *Maslach Burnout Inventory Manual*, 3rd Edn. Palo Alto, CA: Consulting Psychologists Press.

McKown, C. (2019). Challenges and Opportunities in the Applied Assessment of Student Social and Emotional Learning. *Educ. Psychol.* 54, 205–221. doi: 10.1080/00461520.2019.1614446

Mihalic, S. F., Fagan, A. A., and Argamaso, S. (2008). Implementing the LifeSkills drug prevention program: factors related to implementation fidelity. *Implement. Sci.* 3:5. doi: 10.1186/1748-5908-3-5

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and The Prisma Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J. Clin. Epidemiol.* 62, 1006–1012. doi: 10.1016/j.jclinepi.2009.06.005

Molloy, L. E., Moore, J. E., Trail, J., Van Epps, J. J., and Hopfer, S. (2013). Understanding Real-World Implementation Quality and "Active Ingredients" of PBIS. *Prev. Sci.* 14, 593–605. doi: 10.1007/s11121-012-0343-9

Muijs, D., Kyriakides, L., van der Werf, G., Creemers, B., Timperley, H., and Earl, L. (2014). State of the art – teacher effectiveness and professional learning. *Sch. Eff. Sch. Improv.* 25, 231–256. doi: 10.1080/09243453.2014.885451

Musci, R. J., Pas, E. T., Bettencourt, A. F., Masyn, K. E., Ialongo, N. S., and Bradshaw, C. P. (2019). How do collective student behavior and other classroom contextual factors relate to teachers' implementation of an evidence-based intervention? A multilevel structural equation model. *Dev. Psychopathol.* 31, 1827–1835. doi: 10.1017/S095457941900097X

Oberle, E., Schonert-Reichl, K. A., Hertzman, C., and Zumbo, B. D. (2014). Social-emotional competences make the grade: predicting academic success in early adolescence. *J. Appl. Dev. Psychol.* 35, 138–147. doi: 10.1016/j.appdev.2014.02.004

OECD (2021a). *Beyond academic learning: First results from the survey of social and emotional skills*. Paris: OECD Publishing.

OECD (2021b). *Building the future of education*. Available online at: https://www.oecd.org/education/future-of-education-brochure.pdf

Osher, D., Kidron, Y., Brackett, M., Dymnicki, A., Jones, S., and Weissberg, R. P. (2016). Chapter 17. Advancing the Science and Practices of Social and Emotional Learning: looking Back and Moving Forward. *Rev. Educ. Res.* 40, 644–681. doi: 10.3102/0091732X16673595

Pas, E. T., Bradshaw, C. P., and Cash, A. H. (2014). "Coaching Classroom-Based Preventive Interventions," in *Handbook of School Mental Health. Research, Training, Practice and Policy*, eds M. D. Weist, N. A. Lever, C. P. Bradshaw, and J. S. Owens (New York, NY: Springer), 255–268. doi: 10.1007/978-1-4614-7624-5_19

Pas, E. T., Kaihoi, C. A., Debnam, K. J., and Bradshaw, C. P. (2022). Is it more effective or efficient to coach teachers in pairs or individually? A comparison of teacher and student outcomes and coaching costs. *J. Sch. Psychol.* 92, 346–359. doi: 10.1016/j.jsp.2022.03.004

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., et al. (2011). Outcomes for Implementation Research: conceptual Distinctions, Measurement Challenges, and Research Agenda. *Adm. Policy Ment Health* 38, 65–76. doi: 10.1007/s10488-010-0319-7

Rabiner, D. L., Godwin, J., and Dodge, K. A. (2016). Predicting Academic Achievement and Attainment: the Contribution of Early Academic Skills, Attention Difficulties, and Social Competence. *Sch. Psych. Rev.* 45, 250–267. doi: 10.17105/SPR45-2.250-267

Ransford, C. R., Greenberg, M. T., Domitrovich, C. E., Small, M., and Jacobson, L. (2009). The Role of Teachers' Psychological Experiences and Perceptions of Curriculum Supports on the Implementation of a Social and Emotional Learning Curriculum. *Sch. Psych. Rev.* 38, 510–532.

Schonert-Reichl, K. A. (2017). Social and Emotional Learning and Teachers. *Future Child* 27, 137–155. doi: 10.1353/foc.2017.0007

Streimann, K., Selart, A., and Trummal, A. (2020). Effectiveness of a Universal, Classroom-Based Preventive Intervention (PAX GBG) in Estonia: a Cluster-Randomized Controlled Trial. *Prev. Sci.* 21, 234–244. doi: 10.1007/s11121-019-01050-0

Suntheimer, N. M., and Wolf, S. (2020). Cumulative risk, teacher-child closeness, executive function and early academic skills in kindergarten children. *J. Sch. Psychol.* 78, 23–37. doi: 10.1016/j.jsp.2019.11.005

Swift, L. E., Hubbard, J. A., Bookhout, M. K., Grassetti, S. N., Smith, M. A., and Morrow, M. T. (2017). Teacher factors contributing to dosage of the KiVa anti-bullying program. *J. Sch. Psychol.* 65, 102–115. doi: 10.1016/j.jsp.2017.07.005

Timperley, H., Wilson, A., Barrar, H., and Fung, I. (2007). *Teacher Professional Learning and Development. Bet Evidence Synthesis Iteration [BES]*. New Zealand: Iterative Best Evidence Synthesis Programme.

Tolan, P., Molloy Elreda, L., Bradshaw, C. P., Downer, J. T., and Ialongo, N. (2020). Randomized trial testing the integration of the Good Behavior Game and My Teaching Partner^{TM}: the moderating role of distress among new teachers on student outcomes. *J. Sch. Psychol.* 78, 75–95. doi: 10.1016/j.jsp.2019.12.002

Vears, D. F., and Gillam, L. (2022). Inductive content analysis: A guide for beginning qualitative researchers. *FoHPE* 23, 111–127.

Walker, H. M., Horner, R. H., Sugai, G., Bullis, M., Sprague, J. R., Bricker, D., et al. (1996). Integrated approaches to preventing antisocial behavior patterns among school-age children and youth. *J. Emot. Behav. Disord.* 4, 194–209. doi: 10.1177/106342669600400401

Wandersman, A., Duffy, J., Flaspohler, P., Noonan, R., Lubell, K., Stillman, L., et al. (2008). Bridging the Gap Between Prevention Research and Practice: the Interactive Systems Framework for Dissemination and Implementation. *Am. J. Community Psychol.* 41, 171–181. doi: 10.1007/s10464-008-9174-z

Wanless, S. B., Rimm-Kaufman, S. E., Abry, T., Larsen, R. A., and Patton, C. L. (2015). Engagement in Training as a Mechanism to Understanding Fidelity of Implementation of the Responsive Classroom Approach. *Prev. Sci.* 16, 1107–1116. doi: 10.1007/s11121-014-0519-6

Wehby, J. H., Maggin, D. M., Moore Partin, T. C., and Robertson, R. (2012). The Impact of Working Alliance, Social Validity, and Teacher Burnout on Implementation Fidelity of the Good Behavior Game. *Sch. Ment Health* 4, 22–33. doi: 10.1007/s12310-011-9067-4

Weissberg, R. P., Durlak, J. A., Domitrovich, C. E., and Gullotta, T. P. (2015). "Social and emotional learning: past, present and future," in *Handbook of Social and Emotional Learning: Research and Practice*, eds J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, and T. P. Gullotta (New York, NY: Guilford Press), 3–19.

Zakszeski, B., Hojnoski, R. L., Dever, B. V., DuPaul, G. J., and McClelland, M. M. (2020). Early Elementary Trajectories of Classroom Behavior Self-Regulation: prediction by Student Characteristics and Malleable Contextual Factors. *Sch. Psych. Rev.* 49, 161–177. doi: 10.1080/2372966X.2020.1717373

Zins, J. E., Bloodworth, M. R., Weissberg, R. P., and Walberg, H. (2004). "The scientific base linking social and emotional learning to school success," in *Building Academic Success On Social And Emotional Learning: What Does The Research Say?*, eds J. E. Zins, R. P. Weissberg, M. C. Wang, and H. J. Walberg (New York, NY: Teachers College Press), 3–22.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 25 January 2023
DOI 10.3389/fpsyg.2022.973184

OPEN ACCESS

EDITED BY
Andres Molano,
University of the Andes, Colombia

REVIEWED BY
Li Zhang,
East China Normal University, China
Sarah Pila,
Northwestern University, United States
Jonathan Seiden,
Harvard University, United States

*CORRESPONDENCE
Zezhen Wu
✉ zw1429@nyu.edu

SPECIALTY SECTION
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

RECEIVED 19 June 2022
ACCEPTED 20 December 2022
PUBLISHED 25 January 2023

CITATION
Wu Z, Brown L, Kim HY, Yoshikawa H and Aber JL (2023) Measuring the dosage of brief and skill-targeted social-emotional learning (SEL) activities in humanitarian settings. *Front. Psychol.* 13:973184. doi: 10.3389/fpsyg.2022.973184

Measuring the dosage of brief and skill-targeted social-emotional learning (SEL) activities in humanitarian settings

Zezhen Wu*, Lindsay Brown, Ha Yeon Kim, Hirokazu Yoshikawa and J. Lawrence Aber

Department of Applied Psychology, New York University, New York, NY, United States

Introduction: In humanitarian settings, social-emotional learning (SEL) programs for children are often delivered using a field-feasible approach where the programs are more easily deployable and adaptable in the field, require minimal training, and depend less on the strict sequence and structure of the program components to elicit the intended treatment effect. However, evidence is lacking on what aspects of this implementation approach enable the SEL programming to be more beneficial to children's SEL development.

Method: In this study, we propose and evaluate measures for three dimensions of dosage (quantity, duration, and temporal pattern) of two sets of brief and skill-targeted SEL activities (*Mindfulness* and *Brain Games*) implemented in 20 primary schools in two low-income chiefdoms of Sierra Leone.

Results: We find preliminary evidence of predictive validity that these dosage measures could predict children's attendance and classroom adaptive behavior.

Discussion: This study is the first to develop procedures to measure the dimensions of dosage of brief SEL activities in humanitarian settings. Our findings illuminate the need for future research on optimizing the dosage and implementation design of SEL programming using brief SEL activities.

KEYWORDS

social-emotional learning, skill-targeted activity, implementation, dosage, humanitarian setting

Introduction

Wars and diseases have shattered many children's lives. As conflicts and crises continue, promoting children's learning and well-being through schooling becomes even more challenging in humanitarian settings (Winthrop and Kirk, 2008; UNICEF, 2021). One way to tackle this challenge is to develop and deliver programs that foster

students' social-emotional learning (SEL) in classrooms (Greenberg et al., 2003). Not only do SEL programs improve individual students' abilities to cope with social and emotional challenges (Durlak et al., 2010; Weissberg et al., 2015), but classrooms infused with SEL-principled practices also provide a nurturing environment for students to socialize and to "cope and hope" (Winthrop and Kirk, 2008).

Over the past 20 years, numerous studies conducted in Western, high-income countries have demonstrated that school-based SEL programs can positively impact children's social-emotional skills and academic outcomes over time (Durlak et al., 2010, 2011; Weissberg et al., 2015). However, there is a dearth of evidence on what SEL programs could support children's development in humanitarian settings. SEL programs developed in Western contexts are typically comprehensive, pre-packaged, and lesson-based curricula to be implemented in a formal school setting with extensive support from research- and practice-oriented organizations (Durlak et al., 2011; Jones et al., 2021). These can hardly be achieved in countries plagued with conflicts and crises. First, fragile formal education systems common in humanitarian settings are unlikely to support comprehensive SEL curricula (Global Education Monitoring Report., 2018). Second, many students are unable to attend the programs regularly due to various risk factors in their lives (Kearney et al., 2019). For instance, in a set of large-scale SEL-infused remedial programs by the International Rescue Committee (IRC), the average monthly attendance rate in the program was only 50% in Lebanon and 64% in Niger (Aber et al., 2021). Third, more comprehensive programs are often difficult to "implement with fidelity" at scale (Durlak, 1998; Carroll et al., 2007)—an umbrella term for the degree to which the program is implemented as intended by the developer (Dusenbury et al., 2003). Evidence suggests that low fidelity is often associated with a loss in program effectiveness (Dane and Schneider, 1998; Dusenbury et al., 2003), and it is especially difficult to maintain fidelity for comprehensive SEL programs in humanitarian settings due to factors such as lack of trained and well-supported personnel, attrition of personnel over time, and under-resourced facilities (Murray et al., 2014).

To address these challenges, many international organizations have attempted to build their own SEL programs using a *field-feasible* approach. This approach ensures that the programs are more easily deployable and adaptable in the field, require minimal training, and depend less on the strict sequence and structure of the program components to elicit the intended treatment effect (Jones et al., 2017). One notable effort in building such an approach is the brief and skill-targeted *SEL activities* developed by the IRC and implemented through the Education in Emergencies—Evidence for Action (3EA) initiative in multiple countries in the Middle East and Africa. Specifically, these brief activities are simple, small, and essential elements of larger, more complex evidence-based SEL practices that are designed to be flexibly implemented daily in classrooms by the teachers. These activities are being proposed as additions to IRC's SEL-infused academic programs to provide key support for children's growth and behavioral change (Embry and Biglan, 2008). Teachers are trained to use a menu of locally-adapted activities and can flexibly choose any activity that best suits their student's needs in each session. Although these features allow the implementation of the program to be field-feasible, the drawback to this approach is that there are likely as many specific versions of the program as teachers due to the variations in actual quantity, duration, and repetition patterns of the implemented activities. As a result, it is unclear what aspects of this implementation approach enable the SEL programming to be more beneficial to children's SEL development, if at all (Kim et al., in press).

Our study aims to explore and illustrate one possible way to better understand and improve the effectiveness of these brief and skill-targeted SEL activities by examining their implementation dosage in two low-income chiefdoms of Sierra Leone. Specifically, we focus on measuring three dimensions of dosage: "how much, how often, and for how long" the activities were implemented (Dolan, 2018), and examining the relationship between these fine-grained measures and children's outcomes—attendance and classroom adaptive behavior—to provide preliminary evidence of predictive validity for these measures.

Defining and measuring the dosage of brief and skill-targeted SEL activities

Because skill-targeted SEL activities are designed to be brief and repeatable, they have the potential benefit of being implemented frequently over time. However, as teachers choose what they deem suitable to the children's needs and preferences, it necessarily leads to great variations in the implementation *dosage* across teachers and classrooms.

Broadly, dosage describes how much of the program is delivered (Durlak, 1998). Common measures of program dosage as delivered include the program duration, number of program components, and comprehensiveness of the content [see Durlak and DuPre (2008) for a review of implementation factors]. Recent dosage frameworks of educational programs and behavioral interventions further expand the definition of dosage to encompass "how much, how often, and for how long" each set of activities in a program is implemented (Voils et al., 2012; Dolan, 2018). For brief SEL activities, these frameworks help us distinguish the dimensions of dosage at the *activity* level (e.g., number of brief SEL activities) from the ones at the *temporal* level (e.g., minutes to deliver one SEL activity session). Building on these past frameworks, we develop detailed measures to specifically capture three dimensions of dosage: *quantity* (how much), *duration* (for how long), and *temporal pattern* (how often).

First, *quantity* is measured as (1) the number of implemented activities (*amount*) and (2) the number of unique activities (*variety*) in a given period (e.g., per day, week, or month), regardless of what SEL domains the activities are targeting. An intuitive rationale for focusing on *amount* is that more activities may indicate more SEL skills that the children could receive and have the opportunity to practice if they attend the program regularly. This *amount* measure is also the most commonly used one in other implementation frameworks to represent the dosage of behavioral interventions (Voils et al., 2012; Dolan, 2018). The rationale to focus on *variety* is that more variety may indicate more diversity in the implemented activities, and less variety may indicate more repetition. On the one hand, diverse activities may provide children with more varied opportunities to expand their skill sets, while they also risk overwhelming the children with too many skills to acquire or may be more difficult for the teachers to implement. On the other hand, a smaller set of activities may provide children with more opportunities to regularly practice the targeted skill sets but reduce opportunities for them to try new and varied activities.

Next, *duration* is measured as the length of an activity session. It should not be confused with the duration of the entire program, which is more useful for cross-program comparisons but less so when schools implement the program under a relatively similar time frame (e.g., one school year). Although the activities are designed to be brief, a very short average implementation duration may indicate less adherence to the intended duration or potentially less engagement from teachers to implement the activities comprehensively. In general, we would expect that the activities best engage children when their duration is close to the duration intended by the program developer.

Finally, *temporal pattern* refers to the longitudinal repetition patterns in implementing different activities. In this paper, we focus on the *domain-specific temporal pattern* in implementing activities from groups targeting different SEL domains. Specifically, we are interested in measuring (1) how often activities targeting the same SEL domain are repeated and (2) how many activities are implemented before at least one activity is attempted from each available SEL domain. The former reflects the tendency to implement activities targeting the same skills consecutively, and the latter reflects the tendency to implement activities targeting various skills consecutively.

How activities targeting the same or different types of skills are repeated is central yet unique to the frequently implemented brief activities and thus has been less studied in past implementation frameworks. However, the concept of meaningful repetition of classroom practices is deeply rooted in the literature on establishing norms and regularities over time (Seidman, 1988; Sarason, 1996). In educational settings, norms can be created through "intentional, deliberate, frequent actions" (Jones and Bouffard, 2012). By engaging students in everyday SEL activities, teachers create norms that shape and routinize their SEL practices and habits. In a minimally-resourced classroom in humanitarian settings, these norms become especially important to create a sense of stability and predictability for students (Rawlings Lester et al., 2017) and increase their feelings of "security and control" (Cummings, 2000; Winthrop and Kirk, 2006). Furthermore, meaningful iterations of multiple SEL activities may create a "spiral curriculum" (Harden, 1999) when previously learned SEL skills are reinforced and deepened in a patterned and structured way. Therefore, it is important to capture what routines or patterns are created in repeating activities targeting various SEL domains.

The current study

In the current study, we operationalize the three measures of dosage—quantity, duration, and temporal pattern—for two sets of brief SEL activities conducted in an IRC program called *Learn Safe in Bo* in Sierra Leone. We also take a descriptive and exploratory approach to "identify and narrow the universe of (dosage) values" (Voils et al., 2014) by analyzing the implementation data. That is, we use the measures to predict children's outcomes in the program to acquire evidence of predictive validity (Spear, 2014).

First, we examine whether dosage measures of brief SEL activities have enough variations across the classrooms that participated in the program. Second, we examine the relationship between the measures of dosage and children's later classroom attendance rate, adjusting for their current attendance rate.[1] Third and finally, we examine the relationship between the measures of dosage and changes in children's classroom adaptive behavior (concentration problems, disruptive behavior, and prosocial behavior) from the beginning to the end of the school year.

Research questions

RQ1: Do measures of the dimensions of dosage—(a) quantity, (b) duration, (c) temporal pattern of brief SEL activities have enough variations across classrooms?

RQ2: Do measures of the dimensions of dosage—(a) quantity, (b) duration, (c) temporal pattern—of brief SEL activities predict children's later attendance (next day, week, month), adjusting for concurrent attendance and baseline child characteristics?

1 The direct and indirect outcomes of the brief SEL activities implemented in *Learn Safe in Bo* are associated negatively with several known risk factors for school absenteeism, such as anxiety and depression, negative school attitude and low academic self-concept (Gubbels et al., 2019), consistent with the mounting evidence on the "added value" of SEL programs in children's schooling outcomes (Gjicali et al., 2020). Furthermore, preliminary qualitative evidence suggests that a higher dosage of SEL activities might even *lead to* a higher attendance rate among children in *Learn Safe in Bo*, as indicated in the interviews for pedagogical coaches [Brown, L. (in preparation). *Attendance patterns and predictors of attendance among primary school children in Sierra Leone*. New York, NY].

RQ3: Do measures of the dimensions of dosage—(a) quantity, (b) duration, (c) temporal pattern—of brief and skill-targeted SEL activities predict children's classroom adaptive behaviors at the end of the school year, adjusting for children's behavior and characteristics at baseline?

Materials and methods

Context

For many decades, sub-Saharan African countries have faced tremendous challenges due to armed conflict (Moe, 2009) and public health crises. Among those countries, Sierra Leone experienced an 11-year civil war (Gberie, 1998), and later an Ebola pandemic affected the lives of tens of thousands of people (World Health Organization., 2015). Large-scale studies found a high prevalence of mental health and developmental problems among Sierra Leonean children, even many years after the armed conflict (Behrendt, 2008; Yoder et al., 2016; Thulin et al., 2020).

Despite their potential to buffer Sierra Leonean children from their social and emotional challenges, SEL programming was not introduced to the country's education system until very recently (Boisvert, 2017). Among the efforts to introduce SEL to such humanitarian settings, the IRC, in collaboration with the Ecological Approaches to Social Emotional Learning (EASEL) Lab at Harvard University and Global TIES for Children at New York University (NYU), developed and adapted several brief and skill-targeted SEL activities in Sierra Leone and other countries such as Niger and Lebanon (Brown et al., 2019, 2022; Dolan et al., 2021). In 2017–2018, the IRC implemented an SEL-infused academic program called *Learn Safe in Bo* in 20 primary schools in two chiefdoms (Baoma and Niawa Lenge) in Bo Town, Bo district, the second-largest city in Sierra Leone with a population of over 200,000, which was severely affected by the Ebola pandemic in 2014–2015.

Program characteristics

Sample

Data collection was conducted in 20 schools in Baoma and Niawa Lenge. Each school had one classroom in each grade, and the study sample included all children (N = 1,414, 52.5% female) from all classrooms from grades one to three (J = 60). 40.6% of the sampled children were in the 1st grade, 33.2% in the 2nd grade, and 26.2% in the 3rd grade. There were altogether 74 teachers on record, but data collection challenges prevented reliable tracking of their names and IDs.

Intervention

The program had multiple teacher training and coaching components on literacy curricula and SEL activities for classroom use, material provision and facility improvement in school, and community mobilization. Two sets of brief SEL activities were implemented as part of *Learn Safe in Bo*. The first set included 24 teacher-led *Mindfulness* activities that involved various brief breathing techniques and self-regulatory strategies to help children down-regulate and relieve their stress and overwhelming emotions (Scholastic, 2011; Kim et al., 2019). The IRC developed these activities, drawing references from existing practices of mindfulness (Greenberg and Harris, 2012) and the activities in *Mindup* (Scholastic, 2011)—a comprehensive mindfulness-based SEL program for children from pre-kindergarten to eighth grade. Two recent studies in sub-Saharan countries have also found positive effects of mindfulness-based SEL programs on reducing sadness dysregulation and aggressive responses in social conflict situations for children in grades two to four in Niger[2] (Kim et al., 2019) and more empathic behaviors and better grades for children in grade five to seven in Uganda (Matsuba et al., 2020).

There were three types of targeted skills among the 24 Mindfulness activities: (1) *discovering* (students discover what is happening around them and in their bodies), (2) *experimenting* (students build an understanding of belly breathing and the purpose of mindfulness), and (3) *accepting* (students remain still and quiet for a longer time and learn to accept the different feelings and sensations in their bodies, as well as what is happening around them).

The second set of activities included 20 teacher-led *Brain Games* activities (Jones et al., 2019). These activities were developed based on the core activities in a larger SEL program called SECURe (Social, Emotional, and Cognitive Understanding and Regulation in education) (Jones et al., 2014). These activities aim to improve children's executive function and self-regulation. A recent cluster-randomized study delivered 40 weeks of Brain Games (five games per week) to low-income Latinx children from pre-K through fourth grade in the U.S. Although the actual quantity of implementation was not ideal (from 72 in fourth grade to 157 games in pre-K), the study still yielded marginal positive effects on regulation-related behaviors, attention control, and impulsivity (Barnes et al., 2021).

There were also three types of targeted skills (*Brain Games Power*) among the 20 *Brain Games* activities: (1) *focus* (attention skills; e.g., "The teacher says, "I spy with my little eyes something that is —" (choose a color or shape to describe an object in the room) and children look and point at what they think the object is."), (2) *remember* (working memory; e.g., "Students stand in a circle. One by one, each student says their name and does a motion along with it. The rest of the class then repeats the name with the motion as a group, ultimately trying to remember and repeat all names and motions."), and (3) *stop and think* (inhibitory control; e.g., "Students follow the teacher's directions

2 Kim et al. (2019) implemented in Niger the same *Mindfulness* activities that were implemented in *Learn Safe in Bo*.

and movements, but only when the teacher says "Simon says" first.") (Jones et al., 2019).

After conducting each activity, teachers were instructed to conclude by asking children to briefly reflect on (1) what they noticed, (2) what they felt compared to before the activity, and (3) how and when they could use the activity in their daily life. This post-game debrief was intended to draw children's awareness of any changes experienced by the activities and increase the probability of using them outside of a classroom setting.

As a part of their training, teachers were told to conduct at least one *Mindfulness* activity and one *Brain Games* activity per day throughout the school year (see **Supplementary Appendix**: SEL activity list). The suggested schedule was one *Mindfulness* activity before the first class in the morning and one *Brain Games* activity before the first class in the afternoon. Teachers were instructed to conduct these activities in English, the language in which they were originally designed. Both activity sessions were expected to take around 10 min, including the brief reflection period. For both sets of activities, teachers were also instructed to try as many activities from all groups as they deemed fit.

Teacher training and material adaptation

Six face-to-face teacher training workshops on *Mindfulness* (two workshops) and *Brain Games* activities (four workshops) were delivered by IRC pedagogical coaches and NYU research staff in 2016–2017 in Baoma and Niawa Lenge. The teacher training workshops explained the rationale of brief SEL activities and demonstrations from the training staff on their processes.

Before holding those workshops, the research staff paid visits to five schools (three for *Mindfulness* and two for *Brain Games*) to pilot the activities among small groups of children, each with approximately 20 children. The purpose of these contextualization sessions was to refine materials for the population and ensure the materials were more readily accessible to the children. The final adaptations to the materials include (1) clarifications in the names and meaning of activities (e.g., to put on a "Mindfulness Hat" was changed to wear a "Mindfulness Cap"; The Pickler game, where one child had to attempt to make another child laugh was renamed 'The Comedian' since the children did not have familiarity with clowns) and (2) refinement in the prompts (e.g., new questions were added to help with reflection: "What are the differences between before you started the activity and now?").

Measures

Dosage measure: Quantity

To better understand which activities were conducted and at what frequency, teachers recorded every day whether they conducted a *Mindfulness* or *Brain Games* activity, which one it was, and why they selected that activity. Quantity was thus measured using data from these daily activity trackers. Specifically, we created two measures of quantity using this information: (1) the total number of any SEL activities (*amount*) and (2) the number of unique SEL activities (*variety*) implemented within each time frame (week, month, or school year). We calculated these measures for any SEL activity instead of each one separately because teachers implemented one *Mindfulness* and one *Brain Games* activity in 94.5% of the days (as compared to 2.7% where there was one, and 2.8% where there was none).

Dosage measure: Duration

The duration of the activities was collected as a part of the teacher observation protocol, with records of the length of each SEL activity session as observed by the pedagogical coaches during their monthly mentoring visits.

Dosage measure: Temporal pattern

We also used data from the daily activity trackers to calculate two measures of temporal patterns across groups of activities. Individual activities fell into two pre-determined activity groups by the IRC, each with three levels: (1) type of *Mindfulness* activities (discovering, experimenting, accepting) and (2) type of *Brain Games* powers (focus/attention, remember/working memory, stop and think/inhibitory control). Hence, we calculated measures of temporal patterns for these activity groups that clustered the activities by design. All measures were calculated for the entire school year to reflect the change and temporal patterns over the full implementation period. In doing so, we considered the large gaps of no schooling for all schools during the Christmas holiday (December–January) and the presidential election (March–April).[3]

The first measure is termed "Average Exhaustive Gap" (AEG), such that it refers to the number of activity gaps before all activity groups are exhausted. AEG was derived from the *Coupon* measure (Ginsburg and Karpiuk, 1994), named after the "coupon collector's problem," and originally developed as a simple measure of randomness in randomly generated numbers by a human. By calculating an AEG score for a given set of activity groups, we could use it to represent the average number of activity groups implemented before all the possible alternatives were implemented. For example, in a hypothetical three-activity-group set where each group targets a different SEL skill (e.g., for Brain Games, 1 = focus, 2 = remember, 3 = stop and think), "1, 1, 2, 3, 1, 1, 1, 3, 2, 1, 2, 3" would produce a (4 + 5 + 3)/3 = 4 AEG score (Towse and Neil, 1998; see **Table 1** for a visual explanation). It means that, on average, four

3 Schools were closed in March for about a month due to security issues resulting from the national elections.

TABLE 1 Measures of dosage (quantity, duration, and temporal pattern).

Measures of dosage	Sub-category	Operational definition	Examples/Illustrations
Quantity	Amount	Number of implemented SEL activities	☆△□O□ △O☆O□ Amount = 10
	Variety	Number of unique SEL activities implemented in a certain period	☆△□O□ △O☆O□ Variety = 4 (☆△□O)
Duration	–	Average length of an activity session	Activity session duration is observed and recorded by learning coaches (e.g., 10 min).
Temporal pattern	Average exhaustive gap (AEG)	Number of activity gaps before all activity groups are exhausted	1, 1, 2, 3, 1, 1, 1, 3, 2, 1, 2, 3 4 5 3 AEG = (4+5+3)/3 = 4 1–3: three different activity groups A larger AEG score means that more repetitions of certain activity groups happened before all possible activity groups were implemented.
	Average repetition gap (ARG)	Average activity gap between repeated activity groups	1, 1, 2, 3, 1, 1, 1, 3, 2, 1, 2, 3 5 5 $ARG_{group\ 3}$ = (5+5)/2 = 5 1–3: three different activity groups A larger ARG score indicates a lower tendency to repeat a given group of activities.

activities are implemented before teachers conduct activities from all groups. Hence, a larger AEG score would mean that more repetitions of certain activity groups happened before all possible activity groups were implemented.

We created three average AEG scores (fall semester before Christmas 09/25/2017–12/03/2017; spring semester until election 01/08/2018–03/01/2018; spring semester after election 04/16/2018–06/01/2018) and then averaged these scores in each classroom. We could not calculate an AEG score among individual activities because only 6 (out of 60) classes went through all 24 *Mindfulness* activities, and 16 (out of 60) went through all 20 *Brain Games* activities.

The second measure is called Average Repetition Gap (ARG; Ginsburg and Karpiuk, 1994). This simply denotes the average gap or lag between repeated activity groups. For example, in a hypothetical three-activity-group set where each group targets a different SEL skill (e.g., for Brain Games, 1 = focus, 2 = remember, 3 = stop and think), "1, 1, 2, 3, 1, 1, 1, 3, 2, 1, 2, 3" would produce a (5 + 5)/2 = 5 ARG score for group 3 (see **Table 1** for a visual explanation). In our case, it made less sense to create an overall ARG across all activities because the choice to implement one activity more frequently necessarily led to the non-implementation of others. Therefore, we created ARGs separately for the three activity groups in each classroom. For each activity group, we created three ARG scores based on the school recessions and then averaged them in each classroom. In general, higher ARG scores would indicate a lower tendency to repeat a given group of activities.

Attendance

Attendance is recorded from the school administrative data as reported by classroom teachers. They contained daily binary records for each child in each classroom over the school year. We also calculated the average attendance rates (sum of attended days divided by the total number of days intended) for each child per week, month, and school year.

Child SEL outcomes

Child SEL outcomes were collected using the Teacher Observation of Classroom Behavior-Checklist (TOCA-C) (Koth et al., 2009). TOCA-C is a teacher-report assessment of children's socially adaptive classroom behavior. This measure contains 21 items on a six-point scale across three subscales (see **Supplementary Appendix**: TOCA-Checklist): Concentration Problems (seven items), Disruptive Behavior (nine items), and Prosocial Behavior (five items). The original measure was developed using a group of teachers and students from the U.S., and the reliability of the measure was good for all three subscales (Cronbach's αs > 0.80). In our current sample, we had acceptable to good internal consistency (Concentration Problems: $\alpha_{baseline}$ = 0.85, $\alpha_{endline}$ = 0.87; Disruptive Behavior: $\alpha_{baseline}$ = 0.76, $\alpha_{endline}$ = 0.72; Prosocial Behavior: $\alpha_{baseline}$ = 0.68, $\alpha_{endline}$ = 0.66).[4] Importantly, to minimize the reporting burden on teachers, only a randomly

4 Item 13 and 14 were excluded in analysis for both baseline and endline because they greatly lowered the Cronbach's α. Baseline data collection was in September 2017, and endline data collection was in June 2018.

selected sub-sample of children (N = 597; around 10 children per classroom) were rated on TOCA-C by their teachers around both the baseline and the endline of the study. We calculated an average score per classroom for each of the three subscales of the TOCA-C at both time points. Of the three subscales, concentration problem and prosocial behavior are the targeted outcome of both *Mindfulness* and *Brain Games*, while disruptive behavior is not an immediate target of these activities but a medium-transfer outcome that is expected to change as children get more attentive and prosocial in classrooms.

Covariates

Several characteristics of the children and their households were measured using child reports in their home language at either baseline or endline (because they were assumed to be time-invariant) collected by locally-trained enumerators. These covariates could influence children's average attendance rate and thus confound the relationship between dosage and our outcomes. The covariates include demographic characteristics (age, gender, religion, number of minutes to travel to school, number of adults in the household, number of children in the household), material well-being (material that the house was made of, number of mobile phones at home, how often hunger was felt, whether electricity was available at home), and household educational assets (parents' job,[5] whether the parent can read or write, how often English (the language of instruction) was spoken at home, number of books at home, how often parents talked about schoolwork, whether children helped with chores at home, and whether children helped with work outside the home).

Analytical plans

All analyses were conducted in R (R Core Team, 2020). To answer RQ1, we examined the distribution of each measure of dosage (quantity, duration, and temporal pattern) to see if there was adequate variation across classrooms.

To answer RQ2, we examined the relationship between the measures of dosage and children's attendance. Without a prior hypothesis on the time frame that the relationship was established, we conducted exploratory analyses using four time frames: days nested in children in classrooms, weeks nested in children in classrooms, months nested in children in classrooms, and years (no time frame) nested in children nested in classrooms. We averaged both the predictors and the outcomes at the level of each time frame and examined the coefficients on the dosage predictors (π_{001} for measures of quantity and duration; γ_{001} for measures of temporal sequence) to see if the results were sensitive to different nesting. In the analyses with time nested in classrooms, we built three-level models to account for the variation within children over time, within classrooms across children, and across classrooms (see Eqs 1–3).

In all models presented below (Eqs 1–5), t is the time frame (i.e., day, week, or month), and j is the classrooms. X denotes the measure of dosage for classroom j,[6] Z denotes the k child-level covariates, and Y denotes the average attendance rate at t. π, β, and γ denote the random and fixed intercept coefficients at each level. ϵ, u, and ζ denote the error terms at each level. σ_1^2, τ_1^2, and φ_1^2 denote the corresponding variances of the random effects. For models with time frames, we also included attendance lagged by one time frame as a predictor. For the models with no time frame, we built two-level hierarchical linear models with children nested in classrooms. In addition, we tested both separate models with each one of the dosage measures as a predictor and joint models with all measures as predictors. This was to understand whether there was any added value that the measures brought to explain the variance in attendance beyond each single measure, especially the most-commonly used *amount* measure in past literature. We calculated the explained variance from our linear mixed effects models using Omega-squared (Ω^2) by Xu (2003), where $\Omega^2 = 1 - \frac{\sigma^2}{\sigma_0^2}$ (σ^2 is the variance of the residuals in the model, and σ_0^2 is the variance of the response variable in the data).

Level 1: Time

$$\begin{gathered} Y_{(t+1)ij} = \pi_{0ij} + (\pi_{001} X_{tij}) + \pi_{002} Y_{tij} + \epsilon_{tij} \\ \epsilon_{tij} \sim N(0, \sigma_1^2) \end{gathered} \tag{1}$$

Level 2: Child

$$\begin{gathered} \pi_{0ij} = \beta_{00j} + \sum_k \beta_k Z_{0ij} + u_{0ij} \\ u_{0ij} \sim N\left(0, \tau_1^2\right) \end{gathered} \tag{2}$$

Level 3: Classroom

$$\begin{gathered} \beta_{00j} = \gamma_{000} + (\gamma_{001} X_{00j}) + \zeta_{00j} \\ \zeta_{00j} \sim N\left(0, \varphi_1^2\right) \end{gathered} \tag{3}$$

To answer RQ3, we examined the relationship between the measures of dosage and children's classroom adaptive behavior. We used two-level hierarchical linear models with children nested in classrooms and examined the coefficients on the classroom-level dosage predictors predicting endline TOCA subscale scores (i.e., γ_{01}), adjusting for baseline scores and covariates (see Eqs 4, 5). Importantly, children in the current study are nested in classrooms rather than teachers or schools. As mentioned above, this is because teachers were assigned to multiple grades and flexibly deployed on any given day, as well

5 Parents' job is categorically coded: 1 = "Farming", 2 = "Mining", 3 = "Businessperson (seller or vendor in market)", and 4 = "Formal employment (teacher, driver, carpenter, mechanic, etc.)".

6 For measures of temporal patterns, we built one model that included all measures instead of separate models with each of the measures as the only predictor.

10.3389/fpsyg.2022.973184

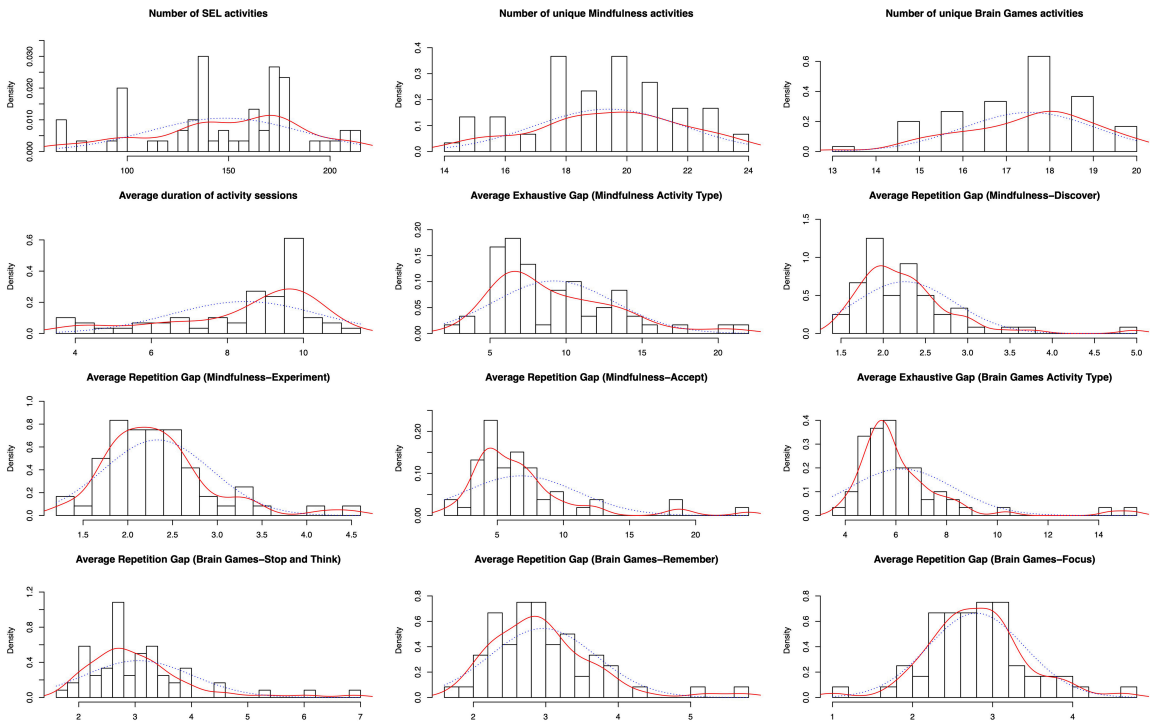


FIGURE 1
Histograms of measures of dosage across classrooms (solid lines indicating density distributions and dotted lines indicating normal fits).

as staff absenteeism and turnover that were not reliably tracked. Hence, the measures of dosage were treated as classroom-level characteristics rather than being tied to each teacher's practices. Again, we tested both separate models and joint models to examine the added value that the measures brought to explain the variance in adaptive behavior beyond each single measure.

Level 1: Child

$$Y_{end\_ij} = \beta_{0j} + \beta_{01} Y_{base\_ij} + \sum_k \beta_k Z_{0j} + u_{0j} \qquad (4)$$
$$u_{0j} \sim N\left(0, \tau_2^2\right)$$

Level 2: Classroom

$$\beta_{0j} = \gamma_{00} + \gamma_{01} X_{0j} + \zeta_{0j} \qquad (5)$$
$$\zeta_{0j} \sim N\left(0, \varphi_2^2\right)$$

Results

RQ1: Do the measures of dosage vary across classrooms?

Figure 1 shows the distributions of the measures of dosage across classrooms along with their density distributions (solid lines) and the normal fits (dotted lines). All measures show substantial variations across their values. AEG—the average number of activity groups implemented before all possible alternatives were attempted—was generally right-skewed. This was expected because teachers were encouraged to try a variety of activities from pre-determined groups. Therefore, classrooms are expected to produce short cycles of attempting activities from all groups. Other measures all had a wide range of values and had distributions close to normal. Pedagogical coaches reported an average activity duration of 3.67–11.20 min (M = 8.45, SD = 1.93). Because the expected duration was about 10 min, the distribution of the actual duration shows that teachers typically took less time to finish a session, and certain teachers might have implemented the activities too quickly (e.g., 4 min), possibly skipping the reflection session. The correlations are low to moderate across the measures of dosage (see Table 2).

RQ2a: Was the quantity of SEL activities associated with higher children's attendance?

The average monthly attendance rate across all classrooms from September to June was 80.7%, ranging from 70.9% in

TABLE 2 A correlation matrix across all measures of dosage (aggregated at the classroom level).

	1. Amount	2. Variety (Mindfulness)	3. Variety (Brain games)	4. Duration	5. AEG (Mindfulness)	6. ARG (Discovering)	7. ARG (Experimenting)	8. ARG (Accepting)	9. AEG (Brain games)	10. ARG (Stop and think)	11. ARG (Remember)	12. ARG (Focus)
1	–											
2	0.341**	–										
3	0.200	0.439***	–									
4	0.595***	0.250	0.068	–								
5	0.090	−0.219	−0.039	0.217	–							
6	−0.034	−0.058	0.038	0.210	−0.147	–						
7	0.196	0.229	0.148	−0.058	−0.044	−0.443***	–					
8	0.053	−0.313*	−0.19	0.085	0.330*	−0.002	−0.083	–				
9	0.467***	−0.082	0.149	0.191	0.162	−0.069	0.094	−0.025	–			
10	0.258*	0.070	0.083	0.253	−0.107	−0.035	0.060	0.069	0.141	–		
11	0.319*	−0.006	0.026	0.121	0.129	0.104	0.074	0.071	0.052	−0.111	–	
12	0.152	−0.215	0.077	0.123	0.034	0.098	0.098	0.138	0.246	−0.16	0.064	–

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

March (due to the election) to 90.5% in October. However, monthly attendance rates varied greatly across classrooms and over time (see **Supplementary Figure 1**).

Table 3 displays how the number of SEL activities predicted children's attendance rate aggregated at different time frames. In models with a time frame, more SEL activities consistently predicted higher children's attendance rate at $t + 1$ adjusting for the current attendance rate at t, such that one more SEL implementation was significantly associated with a 1.9% increase in daily attendance, 1.1% increase in weekly attendance, and 0.2% increase in monthly attendance. A similar relationship was also found when adjusting for other time-invariant covariates.

Table 4 displays the relationship between the variety of SEL activities and children's attendance aggregated at different time frames. In models with a time frame, more variety in SEL activities consistently predicted higher children's attendance rate at $t + 1$ adjusting for the current attendance rate at t, such that one more unique SEL activity is associated with small (0.2–1.1%) but significant increase in attendance. A similar relationship was also found when adjusting for other time-invariant covariates.

RQ2b: Was the duration of SEL activities associated with higher children's attendance?

There was little evidence that the average duration of SEL activity sessions predicted children's attendance rate aggregated at any time frame. The signs of the relationship were inconsistent across different time frames, and no relationship was shown after adjusting for covariates (see **Supplementary Table 1**).

RQ2c: Was the temporal pattern of SEL activities associated with higher children's attendance?

In our models with measures of temporal pattern (AEG or ARG) as predictors, a smaller average repetition gap (ARG) in the *accepting* Mindfulness activity group significantly predicted higher attendance (γ = 0.004, SE = 0.002, 95% CI = [−0.008, 0.000], p = 0.038; that is, implementing one fewer activity targeting other skills in between two *Mindfulness* activities in the *accepting* group was associated with 0.4% increase in attendance), even after controlling for the number of unique individual *Mindfulness* activities (i.e., the variety; γ = 0.005, SE = 0.002, 95% CI = [−0.009, 0.000], p = 0.034). This might indicate that more frequent repetition of a variety of activities related to the *accepting* skill was related to higher attendance. None of the temporal pattern measures for any of the two activity groups yielded a statistically significant relationship with

children's attendance (see **Supplementary Table 2**). In addition, the explained variance was not substantially different in the model with only the *amount* measure ($\Omega^2 = 0.288$) vs. the model with all dosage measures ($\Omega^2 = 0.280$).

RQ3a: Was the quantity of SEL activities associated with children's classroom behavior?

There was little evidence that the amount of implemented SEL activities predicted any of the subscales of children's endline classroom behavior, adjusting for baseline classroom behavior (see **Supplementary Table 3** for the model and **Supplementary Table 9** for summary statistics of the TOCA measure). Meanwhile, more variety in SEL activities predicted fewer concentration problems ($\gamma = -0.063$, $SE = 0.027$, 95% $CI = [-0.115, -0.010]$, $p = 0.022$) and more prosocial behavior ($\gamma = 0.060$, $SE = 0.023$, 95% $CI = [0.014, 0.106]$, $p = 0.013$), adjusting for their baseline scores. That is, more variety in SEL activities were associated with decreased teacher report of children's concentration problems and increased report of prosocial behavior. These findings persisted when adjusting for child-level covariates (concentration problems: $\gamma = -0.056$, $SE = 0.027$, 95% $CI = [-0.109, -0.002]$, $p = 0.045$; prosocial behavior: $\gamma = 0.057$, $SE = 0.025$, 95% $CI = [0.009, 0.105]$, $p = 0.025$) (see **Supplementary Table 4**).

RQ3b: Was the duration of SEL activities associated with children's classroom behavior?

We found that a larger average duration of SEL activity was associated with increased endline prosocial behavior, adjusting for baseline prosocial behavior ($\gamma = 0.101$, $SE = 0.042$, 95% $CI = [0.018, 0.184]$, $p = 0.020$). This finding persisted when adjusting for child-level covariates ($\gamma = 0.092$, $SE = 0.045$, 95% $CI = [0.005, 0.180]$, $p = 0.042$). That is, a longer average activity duration was associated with larger positive changes in prosocial behavior (see **Supplementary Table 5**).

RQ3c: Was the temporal pattern of SEL activities associated with children's classroom behavior?

In our model with measures of temporal pattern (AEG or ARG) as predictors, a larger average exhaustive gap among the three groups of *Brain Games* activities significantly predicted positive changes in prosocial behavior ($\gamma = 0.091$, $SE = 0.039$, 95% $CI = [0.014, 0.168]$, $p = 0.025$; that is, implementing one more activity targeting the same *Brain Games* skill before trying other skills was associated with 0.091 unit of increase in prosocial behavior on a six-point scale), even after controlling for the number of unique individual *Brain Games* activities (i.e., the variety; $\gamma = 0.094$, $SE = 0.038$, 95% $CI = [0.019, 0.168]$, $p = 0.018$). This might indicate that the tendency to implement a variety of activities targeting the same SEL skill before trying all types of skills was related to more prosocial behavior. This finding also persisted when adjusting for child-level covariates ($\gamma = 0.088$, $SE = 0.041$, 95% $CI = [0.009, 0.167]$, $p = 0.037$). Furthermore, no other measures of the temporal pattern for the other groups yielded a statistically significant relationship with children's classroom behavior (see **Supplementary Tables 6–8**). In addition, the explained variance was not substantially different in the model with only the *amount* measure (prosocial behavior: $\Omega^2 = 0.601$; disruptive behavior: $\Omega^2 = 0.591$; concentration problems: $\Omega^2 = 0.586$) vs. the model with all dosage measures (prosocial behavior: $\Omega^2 = 0.590$; disruptive behavior: $\Omega^2 = 0.614$; concentration problems: $\Omega^2 = 0.586$).

Discussion

There are many challenges in measuring the dosage of brief SEL activities and testing its relationship to program effectiveness in humanitarian settings. What we provided in this paper was a novel theory-informing analytical solution to these challenges by developing and testing both manifested (quantity and duration) and hidden (temporal pattern) measures of dosage in implementing brief SEL activities. These measures reflect a wide variety of information embedded in the implementation data commonly collected in SEL interventions (i.e., "how much, how often, and for how long" are the interventions conducted). Although we did not find consistent support for the claim that the new measures explained substantially more variance than the *amount* measure in our data, the new measures could still be conceptually useful for other research projects with similar structures of implementation data, depending on the measures' practical relevance with the research project. In addition, our correlational findings also shed light on potential directions to improve the implementation of brief SEL activities, at least in the Sierra Leonean context. These results suggest that program developers and implementers who wish to improve children's attendance and classroom adaptive behavior should consider increasing the amount and variety of SEL activities and the duration of each session.

Dosage of SEL activities and children's school attendance

In our exploratory analysis, we found some evidence in support of the relationship between measures of dosage and children's school attendance rate at $t + 1$, adjusting for the

TABLE 3 The relationship between the number of SEL activities and children's attendance at time t + 1 (both aggregated at different time frames).

	Day	Week	Month	School year	Day	Week	Month	School year
(Intercept)	0.627* [0.604; 0.650]	0.367* [0.347; 0.387]	0.179* [0.168; 0.190]	0.530* [0.465; 0.594]	0.556* [0.303; 0.808]	0.428* [0.263; 0.592]	0.206* [0.069; 0.344]	0.557* [0.432; 0.682]
Attendance rate at time *t*	0.192* [0.185; 0.198]	0.306* [0.293; 0.319]	0.607* [0.590; 0.623]	–	0.194* [0.187; 0.202]	0.349* [0.334; 0.364]	0.686* [0.668; 0.704]	–
Number of SEL activities	0.019* [0.011; 0.027]	0.011* [0.009; 0.013]	0.002* [0.001; 0.002]	−0.000 [−0.001; 0.000]	0.022* [0.014; 0.031]	0.010* [0.008; 0.012]	0.001* [0.001; 0.001]	−0.000 [−0.001; 0.000]
Added covariates?	**No**	**No**	**No**	**No**	**Yes**	**Yes**	**Yes**	**Yes**
AIC	53431.765	−9539.613	−7556.549	−2285.434	45827.670	−8165.567	−6936.349	−2273.657
BIC	53488.586	−9491.910	−7513.901	−2264.420	46097.719	−7939.485	−6734.844	−2136.293
Log likelihood	−26709.882	4775.806	3784.274	1146.717	−22884.835	4111.783	3497.174	1163.829
$N_{observations}$	95,813	20,961	9,027	1,413	81,802	17,960	7,696	1,197
$N_{children}$	1,394	1,398	1,401	–	1,180	1,185	1,188	–
$N_{classrooms}$	60	60	60	60	60	60	60	60
$Variance_{children}$	0.033	0.014	0.009	–	0.034	0.013	0.009	–
$Variance_{classrooms}$	0.003	0.001	0.000	0.004	0.002	0.001	0.000	0.004
$Variance_{residual}$	0.098	0.032	0.020	0.010	0.098	0.032	0.019	0.006

95% confidence intervals (CI) are also displayed (* indicates that the CI does not contain zero). Model 1–3 and 5–7 have time frames nested children and children nested in classrooms and include attendance rate (led by one time frame) as an outcome. Model 4 and 8 directly nests children in classrooms with no time frame.

TABLE 4 The relationship between the number of unique SEL activities and children's attendance at time t + 1 (both aggregated at different time frames).

	Week	Month	School year	Week	Month	School year
(Intercept)	0.373* [0.354; 0.393]	0.174* [0.162; 0.187]	0.375* [0.201; 0.550]	0.433* [0.268; 0.598]	0.209* [0.071; 0.347]	0.405* [0.207; 0.604]
Attendance rate at time *t*	0.306* [0.293; 0.319]	0.620* [0.604; 0.637]	–	0.350* [0.335; 0.365]	0.697* [0.679; 0.714]	–
Number of unique SEL activities	0.011* [0.009; 0.012]	0.002* [0.001; 0.002]	0.003 [−0.002; 0.008]	0.009* [0.008; 0.011]	0.001* [0.001; 0.002]	0.003 [−0.001; 0.008]
Added covariates?	**No**	**No**	**No**	**Yes**	**Yes**	**Yes**
AIC	−9528.451	−7499.264	−2290.162	−8159.504	−6909.389	−2278.923
BIC	−9480.748	−7456.616	−2269.148	−7933.423	−6707.884	−2141.559
Log likelihood	4770.225	3755.632	1149.081	4108.752	3483.695	1166.462
$N_{observations}$	20,961	9,027	1,413	17,960	7,696	1,197
$N_{children}$	1,398	1,401	–	1,185	1,188	–
$N_{classrooms}$	60	60	60	60	60	60
$Variance_{children}$	0.014	0.009	–	0.013	0.009	–
$Variance_{classrooms}$	0.001	0.000	0.004	0.001	0.000	0.003
$Variance_{residual}$	0.032	0.021	0.010	0.032	0.019	0.006

95% confidence intervals (CI) are also displayed (* indicates that the CI does not contain zero). Model 1–2 and 4–5 have time frames nested children and children nested in classrooms and include attendance rate (led by one time frame) as an outcome. Model 3 and 6 directly nest children in classrooms with no time frame.

present attendance rate at *t*. To begin with, more SEL activities were associated with higher average children's attendance. This finding provided partial support for the hypothesis that more SEL implementation could *lead to* more attendance, consistent with the pedagogical coach's reflection on the program: "*Some pupils are actually coming to school because of those [SEL] games.*" Furthermore, a 0.2% increase in the monthly attendance rate was associated with just one more activity. Therefore, an

addition of ten more activities per month (approximately two more activities per 5-day school week) would associate with a 2% increase in monthly attendance rate (0.10 increase in standard deviation), on top of an average rate of 80.7%. While exploratory, this result provides a promising strategy to increase attendance in humanitarian settings where student attendance fluctuates greatly (Brown et al., 2019). Further studies are needed to examine whether it was increased engagement in the classroom, improved SEL skills, or other factors that explained the relationship between implementing a certain amount of activities and increased attendance.

We also found that more variety in SEL activities was associated with higher attendance. This could be because classrooms with more diverse activities, instead of those repeating a smaller set of activities, motivated children to attend more or that children attended more often and more regularly, allowing the teachers to try new activities confidently. Moreover, implementing fewer activities targeting other skills in between *Mindfulness* activities in the *accepting* group was associated with an increase in attendance after controlling for the variety of individual *Mindfulness* activities. This further indicates that more frequent repetition of activities designed to target the same *accepting* skill might be especially related to higher attendance, even when individual activities varied in their specific content and format. Again, further studies are needed to examine how teachers' and children's motivation and behavior change due to the diversity of the implemented activities in a dynamic process.

Dosage of SEL activities and children's classroom adaptive behavior

Besides children's attendance, we also found evidence supporting the relationship between several of the measures of dosage and children's classroom adaptive behavior. First, more variety in SEL activities was associated with more prosocial behavior and fewer concentration problems. This generally matches the findings in Western countries that *Mindfulness* and *Brain Games* activities could promote children's adaptive behavior by boosting their self-regulation and executive functions (Viglas and Perlman, 2018; Barnes et al., 2021). Thus, by implementing a variety of activities, children may have been exposed to and ultimately learned more skills related to these developmental skills. This could also be because classrooms with more diverse SEL activities created a better atmosphere for children to focus on learning and develop prosocial behavior.

Second, a longer average duration of SEL activities was associated with more prosocial behavior. This might be because teachers who spent more time on one SEL session implemented the post-game reflection or implemented it with higher quality, thereby producing more of the developer's intended effects. Even though certain activities might take less time to implement on average (e.g., Belly breathing), very short sessions (e.g., around 3–5 min) may be less engaging for children or may omit key components of the activity. Although we could not make predictions beyond the sample space in which the range of duration was from 3 to 11 min, we did find that more time on each activity may bring out more intended effects, at least in the observed sample sessions in this program.

Third, implementing more activities targeting the same *Brain Games* skill before trying other skills was associated increase in prosocial behavior after controlling for the variety of individual *Brain Games* activities. This finding might be informative for developers of brief SEL activities to construct detailed instructions for teachers to repeat activities targeting the same skills in a more intentionally sequenced manner while allowing diversity among the content and format of individual activities. Nevertheless, we still need to be cautious about this finding, as it is not necessarily generalizable beyond the sample and context in the current study. In addition, we did not find any relationship between dosage and disruptive behavior. This might be because disruptive behavior was not an explicit immediate target behavior by either set of activities.

Limitations and future directions

This study is one of the first in West Africa to study the dosage of brief SEL activities; thus, our study is limited in its generalizability. It also has various other limitations due to programming and methodological challenges. First, although we emphasized the exploratory nature of the analyses and reported all hypotheses that we tested, the statistically significant results should be interpreted with discretion. Specifically, we tested a total of 52 hypotheses and found 11 statistically significant results with a 5% Type I error rate, which was about 4 times the random chance.

Second, we could not separate the effect of *Mindfulness* activities from that of *Brain Games* activities. Because both activities were implemented on 94.5% of the days, we could only estimate their joint effect as brief SEL activities. However, we acknowledge that these two types of activities target different developmental domains, and the findings presented here might not be generalizable to stand-alone *Mindfulness* or *Brain Games* programs.

Third, the brief SEL activities in *Learn Safe in Bo* were only parts of this larger and more holistic climate-targeted intervention, with other components of teacher training, facility improvement, material provision, and community mobilization. Therefore, our results might not generalize to *Mindfulness* or *Brain Games* programs that are not a part of a climate-targeted intervention.

Fourth, all the activity trackers and outcome measures were reported by teachers. This posed major threats to the validity of the inferences we made using these measures. Activity trackers and attendance records were subject to errors as no other

data source could verify their accuracy. The TOCA-C only measured teachers' observation of children's behavior, which might be biased by teachers' knowledge and perceptions about their own SEL implementation with the children they were observing. We suggest that future studies create a data collection system to track the implemented daily activities more accurately and take a multi-informant measurement approach to measure children's outcomes.

Fifth, all covariates used in our analysis were reported by children. The accuracy of these self-reports might be questionable. For instance, young children might not be familiar with "the material that the house was made of" or accurately remember "how often parents talked about schoolwork." Again, we suggest that future studies take a multi-informant measurement approach to measure this personal and household-related information from both the children and their parents.

Sixth, the findings in this study were neither generalizable to other contexts nor other ranges of dosage values. We were also unable to ascertain the optimal dosage for a lack of information on the generalizability of the study and out-of-sample predictions. To do so, we would need data from other contexts and times. We also encourage future studies to validate these measures against more detailed measures of teacher decision-making in selecting activities and the quality of each activity session. We also encourage studies to explore what the *optimal* value of the dimensions of dosage (Voils et al., 2014)—quantity, duration, temporal pattern—should be to produce a "detectable effect" or the "best effect" (Carroll et al., 2007) of brief SEL activities on program-intended outcomes. As future implementation and experimental studies gather more evidence about the range of dosage values in different contexts and for children at different developmental stages, we can build evidence-based suggestions for teacher training.

Finally, we could not examine teachers' role in implementing activities in our study. We could not reliably link teacher information to the implementation data due to data inconsistencies caused by flexible teacher deployment, teacher absenteeism, and teacher turnover. Therefore, the measures describe classroom-level implementation characteristics only, even if some might be linked to teachers' characteristics, motivation, and classroom-management style, as suggested in past studies (Jones et al., 2014). We were also unable to investigate whether some teachers implemented the activities more consistently and patiently than others or if new teachers had trouble implementing the activities when they substituted for previous teachers. We were also unable to know if teacher bias resulted in high ratings of classroom adaptive behavior for certain groups of children with historically excluded backgrounds. In future studies, we need to collect more information on teachers to determine whether and how variation in dosage patterns relates to teachers' characteristics and practices. Although we could not make strong inferences about teachers in this study, their perceptions of the brief SEL activities and their attitudes and strategies in implementing them matter greatly to the quality delivery of the activities, especially for those that require more teacher-child interactions (Donohue et al., 2000; Hromek and Roffey, 2009).

Conclusion

In conclusion, this study is the first to develop procedures to measure the dimensions of dosage of brief SEL activities and explore their relationship with children's outcomes in humanitarian settings. While exploratory, our findings provide a set of concrete and promising strategies that we can implement and further test to improve the implementation and effectiveness of such SEL programming. It also illuminates the need for future research on developing and validating measures of dosage to provide an evidence-based strategy for implementing these easily trainable, less costly, and effective brief SEL activities in these contexts. Brief and skill-targeted SEL activities hold promise as a feasible programming approach that can improve children's learning and development, especially in crisis-affected low-resource settings, and context- and population-specific implementation dosage, such as those explored in this study, can be used to strengthen its impact.

Data availability statement

Replication data supporting the conclusions of this article will be made available by the authors under a CC-BY 4.0 (Creative Commons Attribution 4.0 International license).

Ethics statement

The studies involving human participants were reviewed and approved by International Rescue Committee Institutional Review Board # 0009752 and New York University Institutional Review Board # IRB-FY2016-1174. Written informed consent to participate in this study was provided by the participants or their legal guardian/next of kin.

Author contributions

ZW conceived the main research questions, carried out secondary data analysis to answer these questions using data from a project led by LB at NYU Global TIES, and wrote the manuscript. LB, HK, HY, and JA provided the critical feedback and revisions for the manuscript. All authors read and approved the final manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.973184/full#supplementary-material

References

Aber, J. L., Tubbs Dolan, C., Kim, H. Y., and Brown, L. (2021). Children's learning and development in conflict- and crisis-affected countries: Building a science for action. *Dev. Psychopathol.* 33, 506–521. doi: 10.1017/S0954579420001789

Barnes, S. P., Bailey, R., and Jones, S. M. (2021). Evaluating the impact of a targeted approach designed to build executive function skills: A randomized trial of Brain Games. *Front. Psychol.* 12:4163. doi: 10.3389/fpsyg.2021.655246

Behrendt, A. (2008). *Psychosocial needs of children without parental support in a post-conflict area: A cross section study in the district of Kailahun in sierra leone*. Dakar: AWARE HIV/AIDS, USAID, FHI, Plan International.

Boisvert, K. (2017). *Case study report: Save the children Sierra Leone*. Available online at: https://scholarworks.umass.edu/cie_eccn/1 (accessed March 20, 2021).

Brown, L., Kim, H. Y., Annan, J., and Aber, J. L. (2019). *SEL functioning in crisis-contexts: The role of gender, grade level, and school attendance in Sierra Leone*. Washington D.C: Society for Research in Educational Effectiveness.

Brown, L. E., Kim, H. Y., Tubbs Dolan, C., Brown, A., Sklar, J., and Aber, J. L. (2022). Remedial programming and skill-targeted SEL in Low-Income and crisis-affected contexts: Experimental evidence From Niger. *J. Res. Educ. Effect.* 1–32. doi: 10.1080/19345747.2022.2139785

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., and Balain, S. (2007). A conceptual framework for implementation fidelity. *Implement. Sci.* 2:40. doi: 10.1186/1748-5908-2-40

Cummings, C. B. (2000). *Winning strategies for classroom management*. Alexandria, VA: ASCD.

Dane, A. V., and Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clin. Psychol. Rev.* 18, 23–45. doi: 10.1016/S0272-7358(97)00043-3

Dolan, C. T. (2018). *Methods for rigorously and responsively improving children's holistic learning and development in crisis contexts: Towards an evidence-based education in emergencies field* [PhD Thesis]. New York, NY: New York University.

Dolan, C. T., Kim, H. Y., Brown, L., Gjicali, K., Borsani, S., Houchaimi, S. E., et al. (2021). Supporting Syrian Refugee children's academic and social-emotional learning in national education systems: A cluster randomized controlled trial of nonformal remedial support and mindfulness programs in Lebanon. *Am. Educ. Res. J.* 59:00028312211062911. doi: 10.3102/00028312211062911

Donohue, K. M., Weinstein, R. S., Cowan, P. A., and Cowan, C. P. (2000). Patterns of teachers' whole-class perceptions and predictive relationships between teachers' and parents' perceptions of individual child competence. *Early Childhood Res. Quart.* 15, 279–305. doi: 10.1016/S0885-2006(00)00068-5

Durlak, J. A. (1998). Why program implementation is important. *J. Prevent. Int. Commun.* 17, 5–18. doi: 10.1300/J005v17n02_02

Durlak, J. A., and DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am. J. Commun. Psychol.* 41:327. doi: 10.1007/s10464-008-9165-0

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

Durlak, J. A., Weissberg, R. P., and Pachan, M. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *Am. J. Commun. Psychol.* 45, 294–309. doi: 10.1007/s10464-010-9300-6

Dusenbury, L., Brannigan, R., Falco, M., and Hansen, W. B. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Educ. Res.* 18, 237–256. doi: 10.1093/her/18.2.237

Embry, D. D., and Biglan, A. (2008). Evidence-based kernels: Fundamental units of behavioral influence. *Clin. Child and Fam. Psychol. Rev.* 11, 75–113. doi: 10.1007/s10567-008-0036-x

Gberie, L. (1998). *War and state collapse: The case of Sierra Leone*. Waterloo, ON: Wilfrid Laurier University.

Ginsburg, N., and Karpiuk, P. (1994). Random generation: Analysis of the responses. *Percept. Mot. Skills* 79, 1059–1067. doi: 10.2466/pms.1994.79.3.1059

Gjicali, K., Wu, Z., Kim, H. Y., and Tubbs Dolan, C. (2020). *Psychometric evidence on the child friendly school questionnaire for syrian children in Lebanon (CFSQ-SL): A measurement tool of student-perceived school climate [Technical working paper]*. New York, NY.

Global Education Monitoring Report. (2018). *Global education monitoring report 2019: Migration, displacement, and education: Building bridges, not walls*. Paris: UNESCO.

Greenberg, M. T., and Harris, A. R. (2012). Nurturing mindfulness in children and youth: Current state of research. *Child Dev. Perspect.* 6, 161–166. doi: 10.1111/j.1750-8606.2011.00215.x

Greenberg, M. T., Weissberg, R. P., O'Brien, M. U., Zins, J. E., Fredericks, L., Resnik, H., et al. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *Am. Psychol.* 58, 466–474. doi: 10.1037/0003-066X.58.6-7.466

Gubbels, J., van der Put, C. E., and Assink, M. (2019). Risk factors for school absenteeism and dropout: A meta-analytic review. *J. Youth Adolesc.* 48, 1637–1667. doi: 10.1007/s10964-019-01072-5

Harden, R. M. (1999). What is a spiral curriculum? *Med. Teach.* 21, 141–143. doi: 10.1080/01421599979752

Hromek, R., and Roffey, S. (2009). Promoting social and emotional learning with games: "It's fun and we learn things.". *Simulation Gaming* 40, 626–644. doi: 10.1177/1046878109333793

Jones, S. M., Bailey, R., Brush, K., and Kahn, J. (2017). *Kernels of practice for SEL: Low-cost, low-burden strategies*. New York, NY: The Wallace Foundation.

Jones, S. M., Bailey, R., Brush, K., and Nelson, B. (2019). *Introduction to the taxonomy project: Tools for selecting & aligning SEL frameworks. Establishing practical social-emotional competence assessments work group, collaborative for academic*. Cambridge, MA: Harvard Graduate School of Education.

Jones, S. M., Bailey, R., and Jacob, R. (2014). Social-emotional learning is essential to classroom management. *Phi Delta Kappan* 96, 19–24. doi: 10.1177/0031721714553405

Jones, S. M., and Bouffard, S. M. (2012). Social and Emotional Learning in Schools: From Programs to Strategies and commentaries. *Soc. Policy Rep.* 26, 1–33. doi: 10.1002/j.2379-3988.2012.tb00073.x

Jones, S. M., Brush, K. E., Ramirez, T., Xinyi, M., Marenus, M., Wettje, S., et al. (2021). *Navigating Social and Emotional Learning from the Inside Out: Looking Inside and Across 25 Leading SEL Programs: A Practical Resource for Schools and OST Providers (Elementary School Focus)*. Cambridge, MA: Harvard Graduate School of Education.

Kearney, C. A., Gonzálvez, C., Graczyk, P. A., and Fornander, M. J. (2019). Reconciling contemporary approaches to school attendance and school absenteeism: Toward promotion and nimble response, global policy review and implementation, and future adaptability (Part 1). *Front. Psychol.* 10:2222. doi: 10.3389/fpsyg.2019.02222

Kim, H. Y., Brown, L., Ferrans, S. D., and Weiss-Yagoda, J. (2019). *The impact of IRC's healing classrooms tutoring on children's learning and social-emotional outcomes in Niger (March 2019 3EA Policy Brief)*. New York, NY: New York University.

Kim, H. Y., Brown, L., Tubbs Dolan, C., Gjicali, K., Deitz, R., Prieto Bayona, M., et al. (in press). Testing the impact of a skill-targeted social and emotional learning curriculum and its variation by pre- and post-migration conflict experiences: A cluster randomized trial with Syrian refugee children in Lebanon. *J. Educ. Psychol.*

Koth, C. W., Bradshaw, C. P., and Leaf, P. J. (2009). Teacher observation of classroom adaptation—Checklist: Development and factor structure. *Measur. Evalu. Counsel. Dev*. 42, 15–30. doi: 10.1177/0748175609333560

Matsuba, M. K., Schonert-Reichl, K. A., McElroy, T., and Katahoire, A. (2020). Effectiveness of a SEL/mindfulness program on Northern Ugandan children. *Int. J. Sch. Educ. Psychol.* 9, S113 - S131. doi: 10.1080/21683603.2020.1760977

Moe, A. T. (2009). *The causes and dynamics of conflict in Sub-Saharan Africa. Army War Coll Carlisle Barracks PA*. Available online at: https://apps.dtic.mil/sti/citations/ADA508284 (accessed March 20, 2021).

Murray, L. K., Tol, W., Jordans, M., Zangana, G. S., Amin, A. M., Bolton, P., et al. (2014). Dissemination and implementation of evidence based, mental health interventions in post conflict, low resource settings. *Int.* 12(Suppl 1), 94–112. doi: 10.1097/WTF.0000000000000070

R Core Team (2020). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Rawlings Lester, R., Allanson, P. B., and Notar, C. E. (2017). Routines are the foundation of classroom management. *Education* 137, 398–412.

Sarason, S. B. (1996). *Revisiting "The culture of the school and the problem of change*. New York, NY: Teachers College Press.

Scholastic (2011). *The mindup curriculum*. New York, NY: Scholastic.

Seidman, E. (1988). Back to the future, community psychology: Unfolding a theory of social intervention. *Am. J. Commun. Psychol.* 16, 3–24. doi: 10.1007/BF00906069

Spear, C. F. (2014). *Examining the relationship between implementation and student outcomes: The application of an implementation measurement framework*. Eugene, OR: Department of Special Education and Clinical Sciences and the Graduate School of the University of Oregon.

Thulin, E. J., McLean, K. E., Sevalie, S., Akinsulure-Smith, A. M., and Betancourt, T. S. (2020). Mental health problems among children in Sierra Leone: Assessing cultural concepts of distress. *Trans. Psychiatry* 59:1363461520916695. doi: 10.1177/1363461520916695

Towse, J. N., and Neil, D. (1998). Analyzing human random generation behavior: A review of methods used and a computer program for describing performance. *Behav. Res. Methods Instrum. Comput.* 30, 583–591. doi: 10.3758/BF03209475

UNICEF (2021). *UNICEF global COVID-19 final report*. Available online at: https://www.unicef.org/media/96151/file/Corrigendum-UNICEF-Final-2020-Global-Covid-19-Situation-Report-Feb-Dec-2020.pdf (accessed March 20, 2021).

Viglas, M., and Perlman, M. (2018). Effects of a mindfulness-based program on young children's self-regulation, prosocial behavior and hyperactivity. *J. Child Fam. Stud.* 27, 1150–1161. doi: 10.1007/s10826-017-0971-6

Voils, C. I., Chang, Y., Crandell, J., Leeman, J., Sandelowski, M., and Maciejewski, M. L. (2012). Informing the dosing of interventions in randomized trials. *Contemp. Clin. Trials* 33, 1225–1230. doi: 10.1016/j.cct.2012.07.011

Voils, C. I., King, H. A., Maciejewski, M. L., Allen, K. D., Yancy, W. S., and Shaffer, J. A. (2014). Approaches for informing optimal dose of behavioral interventions. *Ann. Behav. Med.* 48, 392–401. doi: 10.1007/s12160-014-9618-7

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

Winthrop, R., and Kirk, J. (2006). *Creating healing classrooms: Tools for teachers and teacher educators*. New York, NY: International Rescue Committee, Child and Youth Protection and Development Unit.

Winthrop, R., and Kirk, J. (2008). Learning for a bright future: Schooling, armed conflict, and children's well-being. *Compar. Educ. Rev.* 52, 639–661. doi: 10.1086/591301

World Health Organization. (2015). *WHO | Ebola in Sierra Leone: A slow start to an outbreak that eventually outpaced all others*. Geneva: WHO; World Health Organization.

Xu, R. (2003). Measuring explained variation in linear mixed effects models. *Stat. Med.* 22, 3527–3541. doi: 10.1002/sim.1572

Yoder, H. N. C., Tol, W. A., Reis, R., and de Jong, J. T. V. M. (2016). Child mental health in Sierra Leone: A survey and exploratory qualitative study. *Int. J. Ment. Health Syst.* 10:48. doi: 10.1186/s13033-016-0080-8

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 15 February 2023
DOI 10.3389/fpsyg.2023.977680

OPEN ACCESS

EDITED BY
Celene Domitrovich,
Georgetown University,
United States

REVIEWED BY
Luisa Fernanda Estrada,
Pontifical Bolivarian University,
Colombia
Lucia Bonassi,
Bolognini Hospital,
Italy
Alexis Harris,
University of Virginia,
United States

*CORRESPONDENCE
Gwyne W. White
✉ whitegw@merrimack.edu

SPECIALTY SECTION
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

RECEIVED 24 June 2022
ACCEPTED 23 January 2023
PUBLISHED 15 February 2023

CITATION
White GW, Hatchimonji DR, Vaid E, Simmons CC, Yuan M, Wang A and Elias MJ (2023) Mechanisms for change: A theoretical pathway for a school-wide social–emotional learning initiative in an urban middle school.
Front. Psychol. 14:977680.
doi: 10.3389/fpsyg.2023.977680

Mechanisms for change: A theoretical pathway for a school-wide social–emotional learning initiative in an urban middle school

Gwyne W. White[1]*, Danielle R. Hatchimonji[2], Esha Vaid[3], Christopher C. Simmons[1], May Yuan[4], Angela Wang[4] and Maurice J. Elias[4]

[1]Department of Psychology, Merrimack College, North Andover, MA, United States, [2]Center for Healthcare Delivery Science, Nemours Children's Hospital, Delaware, Wilmington, DE, United States, [3]Department of Psychology, The Pennsylvania State University (PSU), University Park, PA, United States, [4]Rutgers, The State University of New Jersey, New Brunswick, NJ, United States

Introduction: Investment in academic instruction without complementary attention to the social–emotional environment of students may lead to a failure of both. The current study evaluates a proposed mechanism for change, whereby academic achievement occurs as a result of the social–emotional learning environment impacting behavioral (discipline) outcomes.

Methods: We tested the hypothesized model during each year of a 3-year intervention to determine whether the relations among these constructs held potential as a pathway for targeted improvement.

Results: Path analysis for each year demonstrated excellent fit [Year 1: $\chi^2(19)=76.16$, $CFI=0.99$, $RMSEA=0.05$,$TLI=0.98$; Year 2: $\chi^2(19)=70.68$, $CFI=0.99$, $RMSEA=0.048$, $TLI=0.98$; Year 3: $\chi^2(19)=66.59$, $CFI=0.99$, $RMSEA=0.05$, $TLI=0.98$] supporting the theoretical model for change. For each year the effect of the SEL Environment construct on discipline was significant, as was the effect of discipline on Academic Performance. Further, the indirect effect of SEL Environment on Academic Performance was significant across all years.

Discussion: The consistency of these relationships supports the proposed logic model as a potential mechanism for change and has the potential to guide interventions for whole school improvement.

KEYWORDS

SEL, school climate, bullying, discipline, academic achievement, middle school

1. Introduction

As an institution, schools are often tasked with improving the lives of young people through access to support, resources, and services, in addition to academic instruction. Indeed, research has shown that investment in academic instruction without complementary attention to the social and emotional needs of students may lead to failure in both areas (Aygün and Taşkın, 2022). The transition from the final years of elementary school to the next level of schooling is typically when average academic performance falls, particularly for Black and Latinx students (Felmlee et al., 2018; Seeskin et al., 2018). Given that the completion of high school is a critical predictor of future success and overall well-being (De Witte et al., 2013; Rocque et al., 2017), identifying factors that can be modified to support student academic achievement is a valuable target for intervention research.

Programs under the mantle of Social–Emotional Learning (SEL) have been developed in school settings as a means to promote positive social, emotional, and academic growth. These interventions set out to improve student abilities related to a broad set of social and emotional skills in the domains of self-awareness, social awareness, self-management, relationship skills and responsible decision-making (Weissberg et al., 2015; Dermody et al., 2022). A number of different theories have contributed to the development of SEL programs including models of emotional intelligence (Salovey and Mayer, 1990; Goleman, 1995) and emotional consciousness (Damasio, 1999). Emotional intelligence (EI) posits that cognitive abilities and personal characteristics (e.g., emotional abilities, self-regulatory qualities, characteristics of self-awareness, and social skills) are critical for successful interpersonal and goal-oriented outcomes. Key to this construct is the idea that emotional intelligence is an acquired skill, and thus can be enhanced by training and learning (Kanesan and Fauzan, 2019), making EI focus point for intervention work, particularly in the education setting. Relatedly, the research on consciousness and role the feelings has provided biological evidence for the power and role of emotion identification and interpretation for our ability to successfully self-regulate (Damasio and Carvalho, 2013). In tandem with these developments in our understanding of the important role of emotional skills, school-based-prevention experts and educators developed programmatic guidelines to support educational 'Social Emotional Learning' interventions for children and youth (Elias et al., 1997).

The results of meta-analysis and large-scale reviews indicate that SEL interventions can result in positive effects in youth behavior, attitudes, and school performance (Taylor et al., 2017; Murano et al., 2020). When SEL is combined with efforts that foster universal values such as compassion, mutual support, and community service, the degree of distress and disconnection students experience in schools may be reduced (Elias, 2014; Linsky et al., 2018; Wortham et al., 2020). However, the mechanism by which student academic outcomes are improved is complex, shaped by a wide variety of factors both intrinsic to students and existing in their external environment.

Findings from interventions in schools seeking improved academic outcomes indicate that whole school improvement may first begin through a positive change in school culture and climate (Wang and Degol, 2016; Darling and Cook-Harvey, 2018; Hamlin, 2021). A positive social–emotional learning environment can provide an atmosphere of support for students to acquire and grow the individual competencies needed for effective participation in classroom learning and school life. The logic model for change would propose that, as a result of these shifts in environment and expectations, classroom behaviors and peer interactions become less disruptive and more positive resulting in fewer disciplinary referrals (Lacoe and Steinberg, 2018; Reaves et al., 2018). The ultimate outcome of these changes would then be seen in academic improvement at the student and school level. The current study explores the validity of this logic model (i.e., that student perception of their social–emotional learning environment would impact discipline referrals which, in turn, would positively impact academic achievement/grades) in the context of a school-wide effort to improve the social–emotional learning environment of the school.

1.1. Social–emotional learning environment

The environment in which students learn comprises a diverse range of categories and characteristics, including relationships between students and staff, the norms and values in the school, promotion of culture and ethnic traditions, and the physical structure of the building (Wang and Degol, 2016; Del Toro and Wang, 2021; Grazia and Molinari, 2021). It has been shown that students' perceptions of school environment are related to students' academic achievement (Maxwell et al., 2017; Eugene, 2020; Barksdale et al., 2021), students' behavior, and students' decisions to remain in or drop out of school (Gage et al., 2016; Jia et al., 2016). Additionally, research has found that the ability of social–emotional programs to be implemented successfully is related to the culture and climate of the school (Osher et al., 2016). This suggests that, to make a difference in academic achievement, interventions that target academic outcomes must contend with various facets of how students perceive their school environment.

1.1.1. School climate

One aspect that research has identified as key to student perceptions of the school environment, and critical for overall school success, is school climate. Thapa et al. (2013) identified five dimensions of school climate: (1) safety (including social–emotional safety), (2) relationships, (3) teaching and learning, (4) institutional environment, and (5) the school improvement process. Broadly, when students perceive these dimensions of their educational environment positively, the literature indicates a wealth of positive outcomes at both the school and individual level, including an influence on the motivation to learn (Wang et al., 2020); supporting less aggression, violence, and disorder (Bryson and Childs, 2018; O'Connor et al., 2020), and less bullying (Espelage and Hong, 2019). A positive perception of school climate can also mitigate the negative impact of the socioeconomic context on academic success (Eugene, 2020), acting as a protective factor for the learning and positive life development of young people (Lester and Cross, 2015; Steinmayr et al., 2018). A positive school perception of a school's climate by its students is also linked to better overall psychological well-being (Zullig et al., 2018; Capp et al., 2020; Wang et al., 2020). Additionally, and critical for the logic model of the current study, a positive school climate can lower rates of student suspension (Gage et al., 2016). Thus, perceptions of school climate are a key component of an overall positive social–emotional learning environment.

1.1.2. Bullying

There is evidence to suggest that student perceptions of safety (i.e., prevalence of bullying) may be also key aspect of the social–emotional learning environment, and benefit from evaluation distinct from general school climate. Research has identified that feeling safe from harassment and bullying in the school setting is necessary for the promotion of student learning and development (Bradshaw et al., 2021). In schools where students do not experience the supportive norms, structures, and relationships that promote this sense of safety, students are more likely to experience violence and victimization (Williams et al., 2018). Adolescence is a developmental time period during which peer influence is highly formative and peers have been shown to affect each other's behavior, including acceptance of bullying (Dahl et al., 2018; Halliday et al., 2021). In settings where bullying is perceived as a normative part of the school environment, evidence suggests that there are higher levels of absenteeism and reduced academic achievement (Kim et al., 2020). In sum, students who perceive their environment as safe from bullying are more likely to succeed both academically and socially (Juvonen et al., 2011; Bouman et al., 2012; Thompson, 2019; Huang, 2022), suggesting that perception of bullying is another key component of students' social–emotional learning environment.

1.1.3. Peer expectations

Finally, there is evidence to suggest that social-normative expectations, or the expectations one has for the achievements of one's peers, can have an impact on the learning environment (Bell et al., 2019; Vaid et al., 2023). Peer norms have been found to be an important factor in shaping students' academic behaviors (Dijkstra and Gest, 2015; Gremmen et al., 2017). Research indicates that students who have positive expectations about their educational attainment develop optimistic ideas about their potential and achieve in accordance with these notions (Anderson et al., 2018; Saadat et al., 2019). Positive educational expectations are not only critical for promoting achievement, but these expectations may also be a protective asset for vulnerable, at-risk youth (Herrenkohl et al., 2012; Gerard and Booth, 2015; Stoddard and Pierce, 2015; Brumley et al., 2017). While self-expectations are valuable to understand, social-normative expectations may assess a similar construct while reducing potential biases (self-serving bias theory; Miller and Ross, 1975; Shepperd et al., 2008) and provoke students to also think about potential environmental support and barriers (Vaid et al., 2023). In fact, Sommerfeld (2016) found that social-normative expectations explained educational outcomes above and beyond accounting for self- and parental-expectations. Indeed, research has found that group beliefs or attitudes about academic achievement may have a more substantial influence on academic achievement than expectations about oneself (Bell et al., 2019). These findings suggest that the social-normative expectations students hold may influence student behaviors as well as academic outcomes, and may be another core component of a students' perceptions of their social–emotional learning environment.

1.2. Discipline

There is evidence to suggest that, as a result of improvement in student perceptions of the environment and increased social–emotional skills, classroom behavior and instruction can become less disruptive and more positive (Lacoe and Steinberg, 2018; Reaves et al., 2018). This is particularly important in light of the finding that exclusionary school discipline rates in the United States are high, with nearly 2.7 million K-12 students received one or more out-of-school suspensions and over 100,000 students expelled (Gerlinger et al., 2021). Disciplinary actions also have been found to be tied to the race of the student, with racially minoritized students disproportionately affected (Skiba et al., 2011; Anyon et al., 2014; Anyon et al., 2018; Cruz et al., 2021). Notably for the current study, during the middle school grades, there appears to be an increase in both disciplinary rates and racial disparities in discipline and achievement (Skiba et al., 2011; Anyon et al., 2014; White et al., 2016; Gerlinger et al., 2021). Literature further, and unsurprisingly, indicates that high discipline rates tend to be related to negative academic and behavioral outcomes (Anderson et al., 2019; Sorensen et al., 2021). Critical for our understanding of how student perceptions can influence student behaviors, there is also evidence that repeated discipline referrals may trigger a cycle of negative adult-student interaction and may contribute to a student's psychological disengagement (Gregory et al., 2021). The environment created by the teacher and the school can thus be seen as in a cycle with negative student behavior, whereby students are apt to act out in environments where they feel disrespected and disengaged, and teachers' response (e.g., discipline referrals) further that alienation (Cook et al., 2018). Conversely, the proposed change model here suggests that, when students feel a positive connection to their school environment, it serves to make them less likely to engage in acting out behavior, and teachers are, in-turn, less likely to respond harshly to minor perceived infractions supporting a cycle of support and engagement (Valente et al., 2019). Thus, a model of interaction could be proposed in either direction – does the students' perception of their social–emotional learning environment impact discipline, or does discipline impact the perception of the social emotional learning environment? Thus, our study also sought to explore an alternative path between factors, whereby discipline is the first in the cascade, rather than environment. Regardless of direction, however, the evidence suggests that overall school improvement may be related to improving student behaviors, as indicated by disciplinary referrals.

1.3. Student academic achievement

The challenge for American public education is to improve student achievement both broadly and, specifically, for those deemed in need of additional educational support. The importance of academic achievement is long-term and self-reinforcing, as academic success confers many long-term benefits. Indeed, research has consistently found that individuals with higher levels of education are less likely to be unemployed and more likely to earn higher incomes than those with lower levels of education (U.S. Census Bureau, 2019).

Academic achievement has been found to be related to a significant number of factors, which have also been historic targets for intervention. For example, the affective qualities of teacher–student relationships have been found to impact students' as well as teachers' school engagement and achievement (Roorda et al., 2011; Spilt et al., 2011). Student perceptions of competence and relatedness have also been linked to academic outcomes, particularly in the context of students with social and behavioral difficulties (Olivier et al., 2020; Buzzai et al., 2021). Additionally, teachers that demonstrate higher levels of professional competence have been found to engage in more effective teaching, resulting in improved student learning (Fauth et al., 2019; Kyriakides et al., 2020). At the student level, factors such at childhood intelligence (McCoach et al., 2017), executive functioning (Samuels et al., 2016), and perseverance/grit (Credé et al., 2017) all have an impact on academic achievement. In academic achievement outcomes and interventions, the literature suggests there are many pathways to success.

Of concern regarding issues of equity, is that immigrant, and racial/ethnic minoritized children from low-income families face greater barriers to academic success resulting in reduced chances to earn a high school diploma in comparison to their more affluent, White peers (McKinley Yoder et al., 2022). Further, and related, teacher perceptions of children's achievement, whether accurate or not, impact students' grades and scores on standardized achievement tests (Jussim et al., 2009; Liang and Zhang, 2009; Zajda, 2021). These expectancy effects appear strongest for non-White and for low SES youth (McKown and Weinstein, 2008; Fitzpatrick et al., 2015), which may explain the increasing impact that race has on achievement scores from elementary to middle and high school (White et al., 2016). This achievement gap was exacerbated during the COVID-19 pandemic as communities of color continue to face disproportionate detrimental health and economic impacts (U.S. Bureau of Labor Statistics, 2019). Unfortunately, low resourced schools, (e.g., those with high-poverty populations) have historically experienced challenges in implementing effective interventions aimed at achievement due to range of factors (Herman, 2012; Strunk et al., 2016).

Evidence suggests that the social and emotional needs of students are an important component of this overall achievement goal, if not a gatekeeper of academic progress (Corcoran et al., 2018). Thus, it is critical to approach the academic needs of all student populations from a strengths-based, whole-child approach. A youth mindset of perseverance, a construct that has been empirically linked to academic success (Farrington et al., 2012; Yeager and Dweck, 2012), can be fostered in a supportive social–emotional learning environment where interpersonal resilience is scaffolded by intrapersonal engagement (Corcoran et al., 2018). The ongoing and long-term consequences of the COVID-19 pandemic suggest that it is even more important than ever to explore ways to support student resilience and academic achievement.

1.4. The present study

Pathways to sustained improvement in academic achievement are a multidimensional and multistep process, and the mechanisms by which change can occur benefit from validation. The present study seeks to evaluate the theoretical model for change, that hypothesized that, by addressing the social–emotional learning environment, student behaviors would improve, resulting in fewer disciplinary referrals and, ultimately, allowing for overall improved academic achievement. These factors were explored because they were employed by a 3-year school-based intervention ("School of Character") and the current study seeks to assess the value of the model for change imbedded within that active intervention. Our study additionally tested an alternative pathway, to see if the pathway for change alternatively occurred by addressing student problem behaviors improved the social–emotional learning environment, allowing for overall improved academic achievement.

The current study explores the relationships among the domains targeted by the School of Character program to provide support for the logic model proposed and implemented by this intervention. The theory proposed was that, by positively impacting the social–emotional learning environment, there would be a resulting cascading impact on academic achievement. The expectation was to see an impact on disciplinary referrals as function of this pathway. In order to understand the success or failure of SEL intervention programs like the School of Character program, it is critical to evaluate if the proposed mechanisms of change, the pathways by which the intervention hopes to achieve outcomes, are valid. The current study explores the structural pathways between the target constructs to test the underlying theory for the hypothesized change model proposed in the School of Character intervention, as well as an alternative pathway where discipline impacts the social–emotional learning environment. Analyses, therefore, focused on a cross-sectional analysis of each year of the 3 years of the intervention to assess if the underlying theoretical model holds true across time and student population, irrespective of external factors, including intervention impacts.

2. Method

Data for this project were drawn from a 3-year school improvement (School of Character) initiative that assessed school climate and indicators of the school's functioning in an urban middle school in New Jersey. This study was approved by the University Institutional Review Board.

2.1. Participants

This urban middle school generally reflected a student population of approximately 1,300–1,400 students, grades 6 through 8. Students were majority Latinx. The student population also reflected a lower income lower socio-economic status based on percent of students qualifying for free or reduced lunch (a proxy variable for parent income due to the federal income standards required for student receipt of free/reduced lunch price). During Year 1, the mean age of the students at the time of the survey was 12.84, SD = 1.16 (range = 10–16), at Year 2, the mean age of the students was 12.83, SD = 1.12 (range = 11–17), and at Year 3, the mean age of the students was 12.67, SD = 1.02 (range = 11–16). Demographic data for the school at each year of the study are presented in Table 1. Across all 3 years, the school population consistently reflected a majority Latinx population (Year 1: 88%; Year 2: 90%; Year 3: 92%). The student population also

TABLE 1 Demographics characteristic by year.

	Year 1		Year 2		Year 3	
	N	%	N	%	N	%
Grade						
6th	431	37.8%	435	35.8%	433	42.8%
7th	390	34.2%	413	34.0%	311	30.7%
8th	319	28.0%	367	30.2%	268	26.5%
Gender						
Male	587	51.5%	633	52.1%	498	49.2%
Female	553	48.5%	582	47.9%	514	50.8%
Lunch status						
Full price	57	5.0%	51	4.2%	46	4.5%
Reduced	60	5.3%	55	4.5%	44	4.3%
Free	1,023	89.7%	1,109	91.3%	922	91.1%
Classification (LEP or IEP)						
None	890	78.1%	928	76.4%	741	73.2%
Support	250	21.9%	287	23.6%	271	26.8%
Ethnicity (according to School)						
White	7	0.6%	3	0.2%	2	0.2
Black	121	10.6%	107	8.8%	79	7.8
Hispanic	1,005	88.2%	1,097	90.3%	926	91.5
Asian	5	0.4%	5	0.4%	2	0.2
Multi-Ethnic	2	0.2%	3	0.2%	3	0.3
Latinx						
Not Latinx	135	11.8%	118	9.7%	86	8.5
Latinx	1,005	88.2%	1,097	90.3%	926	91.5%
Country of origin						
Not US Born	269	23.6%	267	22.0%	207	20.5%
US Born	871	76.4%	948	78.0%	805	79.5%
Total	1,140		1,215		1,012	

predominately met federal criteria to receive Free Lunch (Year 1: 90%; Year 2: 91%; Year 3: 91%). Due to the homogeneity of these results, further analyses by ethnicity and income-status were not conducted.

2.2. Procedures

The three-year SEL "School of Character" intervention engaged a whole-school intervention model, including several initiatives to impact all students and staff in the school. The methodology of the project followed community-based participatory action research guidelines. The district targeted by this intervention had one of the lowest graduation rates in the state (under 60%) and reading and math testing scores below the 15th percentile in the state. The school in which the intervention was implemented was designated as a "priority" school, an iteration of language used to denote a "failing" school and was publicly known as the "worst" middle school in its entire county. Preliminary work by the School of Character intervention identified that both the culture/climate of the school and the number of disciplinary incidents/referrals were of significant concern to teachers and administration. The intervention program therefore, co-developed with school staff, was intended to help build the positive adult climate and then, by being a source of both engagement and value to students, improve students' perceptions of the climate and greater engagement in the school through reduced disruptive behaviors and increase attention to academics. Research team members partnered directly with administrators and teachers to summarize discipline data and infuse SEL practices into the school discipline system, particularly in the context of In-School Suspension. To simultaneously address staff culture and climate and student discipline concerns, the research team and school staff formed several staff-led committees. One overarching "Climate and Culture" committee, led by one of the school's guidance counselors, met regularly to set an overall strategy for improving school climate and culture, using aggregated school data (e.g., climate surveys, discipline data). Initiatives included opportunities to provide positive feedback amongst staff in a monthly newsletters and hosting community-building events. Subcommittees included a team tasked with the co-creation of an SEL curriculum implemented in daily advisory periods. This team also monitored and supported implementation of the daily curriculum, with one-on-one coaching and modeling for teachers who requested support. Additional subcommittees focused on youth empowerment initiatives.

Data for the current study were collected as part of that school-wide intervention during the 2012–2013, 2013–2014, and 2014–2015 academic years. Survey data were collected from all students during the Fall and Spring for all 3 years, with the exception of Fall of the 2014–2015 school year. Due to administrative concerns regarding the logistics and time–cost associated with a school-wide survey, student data were only collected for 6th graders in the Fall of 2014. In the Spring of 2015, survey data was again collected school-wide for all students. Students and their parents were given the opportunity to opt out of the data collection both through a passive consent form sent home to the parents and an assent form given to the students prior to survey administration. Less than 1% of students or parents opted out. In addition, school records were reviewed to obtain student demographic and academic data.

2.3. Measures

2.3.1. Survey data: School climate

School climate was measured using an adaptation of the School as a Caring Community Profile-II (Lickona and Davidson, 2003), a 42-item measure of perception of school climate. In order to reduce item redundancy and administration time, 22 items from the original measure, with factor loadings below 0.40 or cross factor loadings, were eliminated. The shortened questionnaire consisted of 20 items, for which students rated their agreement on a 5-point Likert Scale, ranging from "Disagree A LOT!" to "Agree A LOT!" Survey included items evaluating student perception of their peers, with questions such as: "students treat classmates with respect;" perception of their teachers, with questions such as: "Teachers in this school like to come here;" and student perception of the student-teacher relationships, with questions such as: "Teachers are unfair in their treatment of students." A total score for this scale was created by summing the items with a higher score equating a more positive sense of school climate. At the time of this study, the SCCP-II was the only empirically supported scale with parallel items for all grade levels, an important consideration to the school district in adopting a school climate measure. Cronbach alphas for each of 3 years, Fall and Spring, ranged from 0.83 to 0.88, suggesting good reliability.

2.3.2. Survey data: Perceptions of bullying

Student perceptions of bullying were evaluated using an 8-item scale created by the research team. The items were developed based on existing assessments of bullying (Williams and Guerra, 2007; Swearer et al., 2010; Espelage and Hong, 2019). Each item used a 5-point Likert Scale, ranging from "Disagree A LOT!" to "Agree A LOT!" Survey items included questions evaluating students' sense of general school safety, including items such as: "Students at this school feel safe," and 'When students see another student being picked on or put down, they try to stop it." Questions also assessed student perceptions of individual level bullying with items such as "Students are often bullied or teased in my school" and "My classmates use computers, videos, smart phones, and other technology to harass other students." Negative items were reverse coded and a total score for this scale was created by summing the items with a higher score, indicating a more positive perception of school safety (less bulling). Cronbach alphas for each of 3 years, Fall and Spring, ranged from 0.69 to 0.75.

2.3.3. Survey data: Social normative expectations

Social-Normative Expectations (SNE), asked students to rate their peers on six items adapted from a study on educational attainment in the Chicago Public Schools (Ou and Reynolds, 2008) and was evaluated as a construct using pilot data from the current project (Bell et al., 2019). Declarative statements were rated on a 5-point Likert Scale ranging from "Disagree A LOT!" to "Agree A LOT!" Questions included items such as: "In the future, most students from this school will graduate from high school" and "In the future, most students in this school will have a happy family life." A single total score for this scale was created by summing the items. Higher scores indicated more favorable ratings of social-normative expectations, i.e., a belief that peers expected to attain success across the six areas. Cronbach alphas for each of 3 years, Fall and Spring, ranged from 0.87 to 0.92.

10.3389/fpsyg.2023.977680

2.3.4. Discipline referrals

Disciplinary data for students were provided by the school. Examples of discipline referrals include such minor misbehavior as 'dress code violation,' 'in the halls without a pass,' and 'tardies to class,' as well as major discipline referrals such as 'harassment/bullying,' 'threatening a staff member/student,' and 'serious disruptive/inappropriate behavior.' In Year 1, the total number of discipline referrals per student ranged from 0 to 121 with a Mean of 5.95 (SD = 13.05; Median = 1.0) with approximately 43% of students evidencing no referral. In Year 2, the total number of discipline referrals ranged from 0 to 134 with a Mean of 8.10 (SD = 15.11; Median = 2.0) with approximately 31% of students evidencing no referral. In Year 3, the total number of discipline referrals ranged from 0 to 65 with a Mean of 2.90 (SD = 5.92; Median = 1.0) with approximately 44% of students evidencing no referral. In order to identify the sample into a relatively even distribution, discipline referrals were recoded into a 0–5 scale for each year, with 0 coded as no discipline referrals, 1 coded as a single discipline referral, 2 coded as 2–3 discipline referrals, 3 coded as 4–7 discipline referrals, 4 coded as 8–20 discipline referrals and 5 coded as greater than 21 discipline referrals (see Table 2).

2.3.5. Academic achievement

Academic grades were obtained from official school records and used in their numeric form, rather than as letter grades (i.e., 95, not "A"), in order to preserve the continuous nature of the data. Academic achievement was measured using the mean of the four quarters for each of the four core subject areas (Language Arts, Math, Science, and Social Students). An average final overall grade was created from these grades that represented a synthesis of the year's academic efforts. Grades, rather than standardized tests, were used as the indicator of academic achievement because of literature supporting grades as better predictors of high school graduation, college performance, and longer-term life outcomes than standardized tests (Geiser and Santelices, 2007). Students who received a grade in 3 out of 4 quarters for 3 out of 4 core subject areas (Language Arts, Math, Science, and Social Students) were considered to have a valid final grade for data analysis. Academic achievement data by year of study are presented in Table 3.

TABLE 2 Discipline referrals by year.

	Year 1	Year 2	Year 3
	n(%)	*n*(%)	*n*(%)
No Discipline referrals	484(42.5%)	375(30.9%)	448(44.3%)
1 Discipline referral	145(12.7%)	180(14.8%)	182(18.0%)
2–3 Discipline referrals	151(13.2%)	168(13.8%)	158(15.6%)
4–7 Discipline referrals	126(11.1%)	161(13.3%)	114(11.3%)
8–20 Discipline referrals	141(12.4%)	185(15.2%)	85(8.4%)
Greater than 21 Discipline referrals	93(8.2%)	146(12.0%)	25(2.5%)

2.3.6. Covariates

In order to control for the known effects of demographic information on academic achievement, we explored four covariates, grade level, gender, if a student received academic support (Individualized Education Plan or Limited English Proficiency) and country of origin (US born or not). Data were obtained from official school records. These covariates were explored due to their documented impacts on academic achievement and discipline (e.g., Porter, 2000; Hubbard, 2005; La Salle et al., 2013; Moreno and Segura-Herrera, 2013; Santiago et al., 2014; Gašević et al., 2016; Morris and Perry, 2017; Daily et al., 2019).

2.4. Missing data

Students without demographic information from the school, who completed less than 3 core classes (language arts, mathematics, science, social studies) were excluded from analysis. Further, the analysis sample was reduced to those who had responded to, at minimum, half of items on each of the 3 social–emotional learning environment survey measures (Climate, Bullying, SNE). Finally, the preferred data point for survey analysis was spring, however, to reduce the bias from missing survey data, if a student had completed a fall survey but not spring, the fall data was substituted (see Table 4). Due to having Fall of 2014 student data for 6th grade students only, any 7th or 8th grade students missing Spring of 2015 survey data were not included in analyses for that year. As a result, a relative reduction in analysis sample size (approximately 100 less students than previous years) occurred.

TABLE 3 Academic achievement data by year.

	Year 1		Year 2		Year 3	
	M	SD	M	SD	M	SD
Year 2012–2013						
Language Arts	75.63	9.81	--	--	--	--
Mathematics	74.67	10.84	--	--	--	--
Science	77.75	9.80	--	--	--	--
Social Studies	78.15	10.36	--	--	--	--
Overall Grade	76.55	8.86	--	--	--	--
Year 2013–2014						
Language Arts	--	--	75.43	9.26	--	--
Mathematics	--	--	75.99	10.32	--	--
Science	--	--	77.47	9.82	--	--
Social Studies	--	--	78.30	10.83	--	--
Overall Grade	--	--	76.80	8.78	--	--
Year 2014–2015						
Language Arts	--	--	--	--	75.22	9.13
Mathematics	--	--	--	--	75.15	10.77
Science	--	--	--	--	77.82	9.76
Social Studies	--	--	--	--	75.98	9.53
Overall Grade	--	--	--	--	76.04	8.43

2.5. Data analyses

Preliminary analyses were conducted to understand the relationships among study variables. All predictor variables were standardized before being entered into the modeling analyses. T-tests and One-Way Analysis of Variance were run to examine differences between the potential demographic variables (gender, grade level, country of birth and if the student received support such as a 504 plan or LEP) and predictor (school climate, social-normative expectations, bullying) and outcome variables (grades).

We tested the hypothesized pathway model, whereby student perception of their social–emotional learning environment (climate, bullying and social normative expectations) impacted discipline referrals, which in turn impacted final grade over three timepoints. This model involved three points (Year 1, Year 2, and Year 3 of the School of Character intervention) examined independently rather than a change model assessing the impact of the intervention on the constructs from Year 1 to Year 3. While the School of Character intervention proposed to improve academic achievement by its implementation, the current study does not explore the efficacy of that program in a longitudinal model of change. Our hypothesis is that the theoretical mechanism of change employed by this intervention (that the social–emotional learning environment impacts discipline which impacts academic achievement) has conceptual validity, with the constructs and variables interacting in such a way that positive academic outcomes could theoretically result from improvement in student perceptions of the social–emotional learning environment. The model explored here is that the proposed pathways between variables are significant, and that another model does not better explain the relationship between the study variables. The efficacy of the School of Character intervention itself must be examined separately so as to accuracy reflect the strengths, weaknesses, successes and failures of a program implemented within a complex community sample and academic system. If the underlying theoretical model for change utilized by the School of Character program has support, future intervention work can then potentially utilize the theoretical model proposed here.

Covariates were not included in analysis model as demographic factors were not predicted to differentiate the proposed mechanism for change being tested. All variables were entered into the sample model and path analysis was used to test a "structural model" (Cohen et al., 1993). For all models, the continuous variables were centered to reduce multicollinearity. Path analysis, while similar to regression analyses, is considered to be more powerful as it examines linear relationships with path coefficients calculated simultaneously for all endogenous variables, rather than sequentially as in multiple regression models, as well as accounting for measurement error. Path analysis has been used to support identifying causal relationships, however, the current study design is not a causal model. Our analyses seek to identify whether the hypothesized path relationships between the study variables were significant, or a different path model would offer a better explanation. Both direct and indirect effects are estimated in the structural model (Kline, 2011). Good fitting models generally have non-significant chi-square values, TLI at or above 0.90, CFI at or above.95, and RMSEA at or below.06. Parameters were established as statistically significant with alpha <0.05. All preliminary analyses were conducted using SPSS software, version 27 (IBM Corporation, 2021) and the modeling analyses was conducted with AMOS software (Arbuckle and Wothke, 2006).

3. Results

3.1. Preliminary analyses

Pearson's correlations were conducted between academic achievement variables (i.e., LA, Math, Science, Social Studies, and Overall Grade) across all 3 years of the study. The relationships between all achievement variables were established to be highly significant and generally consistent across the 3 time points ($r = 0.62–0.77$; $p < 0.001$; See Table 5). Greater variability was identified in the relationship between academic achievement and other study variables (i.e., discipline, climate, bullying, and social normative expectations) across time points (see

TABLE 4 Student perceptions of social emotional learning environment by year.

	Year 1			Year 2			Year 3		
Student-reported measures	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
School climate									
Fall*	984	63.04	12.88	910	64.72	12.33	414	72.02	12.31
Spring	1,044	59.15	11.78	1,154	62.98	11.67	909	67.03	12.67
Analysis sample	1,140	59.30	11.92	1,215	64.13	12.06	1,012	67.36	12.73
Perceptions of bullying									
Fall*	979	24.02	5.73	896	25.30	5.79	390	27.53	5.78
Spring	1,039	23.32	5.65	1,142	25.38	5.56	947	27.22	5.72
Analysis sample	1,140	23.30	5.62	1,215	25.33	5.55	1,012	27.16	5.72
Social normative expectations									
Fall*	986	19.38	5.70	904	20.38	5.51	403	22.47	5.38
Spring	1,046	18.82	5.51	1,153	19.44	5.21	977	20.98	5.46
Analysis sample	1,140	18.89	5.50	1,215	19.49	5.22	1,012	20.98	5.49

*Fall of Year 3 was completed by 6th grade students only.

Table 6; Supplementary material). Notably, academic achievement and discipline referrals were consistently, significantly negatively correlated across all 3 years ($r = -0.67$–-0.51; $p < 0.001$), and all SEL environment measures were significantly positively correlated ($r = 0.41$–0.66; $p < 0.001$). Additionally, the number of discipline referrals and student perceptions of school climate were consistently, significantly negatively related across all 3 years [$r(1140) = -0.16$–-0.12; $p < 0.001$].

Independent t-tests and ANOVAs were also conducted to examine the impact of demographic covariates on academic achievement and discipline (Table 7; Supplementary material) and on measures of the SEL environment (Table 8; Supplementary material). Grade level appeared to have some impact on measures of the social–emotional learning environment, with 6th graders evidencing a better perception of school climate across all 3 years [$F(2,1,137) = 7.11$, $p = 0.001$; $F(2,1,212) = 44.90$, $p < 0.001$; $F(2,1,009) = 17.41$, $p < 0.001$ respectively]. Grade level appeared to have a varying impact on perceptions of bullying, with a significant relationship during year 1 and 2 [$F(2,1,137) = 3.88$, $p = 0.021$; $F(2,1,212) = 4.71$, $p = 0.009$ respectively] with 7th graders reporting the least positive perceptions of student bullying culture during both years. Positive social normative expectations were most consistently reported by 6th graders across all 3 years [$F(2,1,137) = 25.81$, $p < 0.001$; $F(2,1,212) = 15.28$, $p < 0.001$; $F(2,1,009) = 9.30$, $p < 0.001$ respectively]. Gender had no impact on any social emotional learning variable across any of the 3 years (see Table 8; Supplementary material).

3.2. Mechanism for change: The social–emotional learning environment model

To test the reliability of the hypothesized conceptual model on the relationship between student perceptions of the social–emotional learning (SEL) environment, disciplinary action, and academic performance across the three samples tested, we conducted a path analysis in AMOS using each year's sample. Following our *a priori* model, we tested the impact of student perceptions of SEL Environment on academic performance by way of a path through student disciplinary action. Our SEL Environment latent variable was comprised of the Social Normative Expectations, Perceived Bullying, and Climate Survey, and our Academic Performance latent variable was comprised of the Language Arts, Math, Science, and Social Studies grades, in accordance with the *a priori* model's goals. As preliminary findings did not suggest a consistent pattern across measures or time, and due to the low variability of some factors (e.g., overrepresentation of 70% or more), demographic factors can be further explored in future research as part of individual level analysis rather than in the context of the hypothesized mechanism for change.

The path model for Year 1 of the study (Figure 1) demonstrated excellent fit [$\chi^2(19) = 76.16$, $CFI = 0.99$, $RMSEA = 0.05$, $TLI = 0.98$]. The effect of SEL Environment on discipline was significant ($\beta = -0.13$, $p < 0.001$) as was the effect of discipline on Academic Performance ($\beta = -0.60$, $p < 0.001$). Most importantly, the indirect effect of SEL Environment on Academic Performance was significant, if small ($\beta = 0.08$, $p < 0.001$). The path model for Year 2 of the study (Figure 2) likewise demonstrated excellent fit [$\chi^2(19) = 70.68$, $CFI = 0.99$, $RMSEA = 0.048$, $TLI = 0.98$]. The effect of SEL Environment on discipline was significant ($\beta = -0.11$, $p < 0.001$) as was the effect of discipline on Academic Performance ($\beta = -0.64$, $p < 0.001$). As in the first year, the indirect effect of SEL Environment on Academic Performance was significant ($\beta = 0.07$, $p < 0.001$). Finally, the path model for Year 3 of the study (Figure 3) also demonstrated excellent fit [$\chi^2(19) = 66.59$, $CFI = 0.99$, $RMSEA = 0.05$, $TLI = 0.98$]. The effect of SEL Environment on discipline was significant ($\beta = -0.16$, $p < 0.001$) as was the effect of discipline on Academic Performance ($\beta = -0.54$, $p < 0.001$). Once again, the indirect effect of SEL Environment on Academic Performance was

TABLE 5 Pearson's correlations among academic achievement variables.

	1			2			3			4		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
1. LA	--	--	--									
2. Math	0.70***	0.63***	0.62***	--	--	--						
3. Science	0.66***	0.68***	0.67***	0.67***	0.71***	0.67***	--	--	--			
4. Social Studies	0.66***	0.68***	0.72***	0.64***	0.62***	0.57***	0.70***	0.77***	0.68***	--	--	--
5. Overall Grade	0.87***	0.85***	0.87***	0.87***	0.85***	0.84***	0.87***	0.90***	0.88***	0.87***	0.87***	0.86***

***$p < 0.001$.

TABLE 6 Pearson's correlations among continuous study variables.

	1			2			3			4		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
1. Overall grade	--	--	--									
2. Discipline	−0.57***	−0.61***	−0.51***	--	--	--						
3. Climate	0.04	0.08**	−0.08*	−0.12***	−0.13***	−0.16***	--	--	--			
4. Bullying	0.06	0.04	0.08*	−0.10*	−0.05	−0.12***	0.61***	0.47***	0.66***	--	--	--
5. SNE	−0.07*	−0.03	0.00	−0.08**	−0.01	−0.05	0.56***	0.47***	0.60***	0.41***	0.42***	0.58***

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

TABLE 7 Impact of demographic covariates on overall grade and discipline.

	Overall grade						Discipline					
	Year 1		Year 2		Year 3		Year 1		Year 2		Year 3	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Grade												
6th	77.60**	8.97	78.17***	8.02	76.56	8.58	1.53	1.77	1.86	1.76	1.39*	1.48
7th	75.92	8.15	75.31	8.71	75.29	8.41	1.89**	1.78	2.19*	1.85	1.31	1.47
8th	75.90	9.41	76.84	9.44	76.09	8.18	1.44	1.62	2.07	1.74	1.10	1.38
Gender												
Male	74.55	8.80	75.00	8.67	73.57	8.02	1.80***	1.78	2.24***	1.78	1.56***	1.54
Female	78.68***	8.42	78.75***	8.48	78.44***	8.14	1.44	1.68	1.81	1.78	1.03	1.32
Classification												
None	76.67	9.13	77.23**	8.80	76.25	8.82	1.58	1.76	1.92	1.79	1.18	1.40
Support	76.11	7.79	75.39	8.55	75.47	7.24	1.79	1.67	2.38***	1.76	1.59***	1.56
Country of Origin												
Not US	76.15	8.89	76.89	9.19	77.11	7.92	1.61	1.62	1.99	1.70	1.24	1.35
US Born	76.67	8.85	76.77	8.66	75.77*	8.54	1.63	1.78	2.04	1.82	1.30	1.48

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

significant ($\beta = 0.09$, $p < 0.001$). The consistency of the model across the 3 years supports the proposed relationship between SEL Environment, Discipline, and Academic Performance.

Best practices in SEM recommend contrasting a path analysis model with an alternative model using the same data but based on competing theories or alternative explanations. Based upon the literature reviewed here, we developed an alternative model from that of the School of Character intervention, which predicted academic performance as a result of disciplinary action mediated by student perceptions of the social–emotional learning environment (Figure 4). Fit was worse in the alternative model across all three time periods; Time 1 exhibited poor fit [$\chi^2(57) = 1081.86$, $CFI = 0.87$, $RMSEA = 0.09$, $TLI = 0.81$], Time 2 exhibited the worst fit of any model in the study [$\chi^2(57) = 1258.2$, $CFI = 0.86$, $RMSEA = 0.09$, $TLI = 0.79$], and Time 3 had the best fit of the alternative models, but still showed worse fit than our theoretical model [$\chi^2(57) = 734.48$, $CFI = 0.9$, $RMSEA = 0.08$, $TLI = 0.86$]. Given these results, we can conclude that our initial model is a better fitting model than the alternative model.

4. Discussion

The present study evaluated a theoretical model of change that hypothesized a relationship between student perception of social–emotional learning environment, discipline and academic achievement. Our results found that the relationships among the constructs are significant and directionally appropriate to provide support for a mechanism of change. Our data suggests that student perceptions of the social–emotional learning environment impact their disciplinary behaviors which impacts their academic achievement. Further, our results do not support an alternative theory that places disciplinary behavior as the first in the mediational cascade lending further support to the mechanism as proposed. Path analysis for each year demonstrated excellent fit ([Year 1: $\chi^2(19) = 76.16$, $CFI = 0.99$, $RMSEA = 0.05$, $TLI = 0.98$; Year 2: $\chi^2(19) = 70.68$, $CFI = 0.99$, $RMSEA = 0.048$, $TLI = 0.98$; Year 3: $\chi^2(19) = 66.59$, $CFI = 0.99$, $RMSEA = 0.05$, $TLI = 0.98$]. Further, the effect of the social–emotional learning environment construct on discipline was significant during each of the 3 years, as was the effect of discipline on Academic Performance. Finally, the indirect effect of student perceptions of the social–emotional learning environment on Academic Performance was significant across all years. The consistency of the model across the 3 years supports the proposed relationship between student perceptions of SEL Environment, Discipline, and Academic Performance.

These findings suggest that the logic model behind the School of Character Intervention, which proposed a relationship between student perceptions of social–emotional learning environment, student discipline and academic achievement, broadly held true and holds the potential to be an area to target as a mechanism for change. This model of change was implemented as an intervention in a "failing" middle school and the theory hypothesized by the intervention program was that improvement to the school as a whole begins through a positive shift in school culture and climate, and that student perceptions of the social–emotional learning environment has an impact on behavior as evidenced by disciplinary referrals. The School of Character intervention proposed that the mechanism for change proceeded along this pathway to result in student academic achievement outcomes. The current study found, in a cross-sectional analysis of each year, that the relationships between the variables proposed by the theorized logical model were related as hypothesized.

4.1. Limitations

This study faced several limitations that must be considered. The sample utilized may impact generalizability as it reflects a single school. The school district also had one of the lowest graduation rates in the state of New Jersey (under 60%) and reading and math testing scores ranking below the 15th percentile, suggesting a particularly high needs sample. Further, the sample reflects particular demographic characteristics (e.g., majority Latinx, majority of lower SES as evidenced by over 80% of students qualifying for free lunch). As preliminary findings did not suggest a consistent pattern across measures or time, and due to the low

TABLE 8 Impact of demographic covariates on social emotional learning environment factors.

	Climate						Bullying						SNE					
	Year 1		Year 2		Year 3		Year 1		Year 2		Year 3		Year 1		Year2		Year3	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Grade																		
6th	60.48**	12.31	68.26***	12.33	70.01***	12.71	23.19	5.61	25.87	5.65	27.55	5.83	20.32***	5.27	20.49***	5.49	21.84***	5.70
7th	57.49	11.40	60.99	11.74	65.83	12.91	22.85*	5.58	24.70**	5.69	26.93	5.65	17.74	5.67	19.35	4.86	20.34	5.43
8th	59.92	11.77	62.75	10.67	64.85	11.76	24.01	5.65	25.39	5.22	26.80	5.59	18.34	5.16	18.48	5.08	20.36	5.01
Gender																		
Male	59.94	12.01	64.71	11.53	67.34	12.09	23.60	5.60	25.50	5.46	27.05	5.26	18.95	5.33	19.48	4.93	20.91	5.13
Female	58.62	11.79	63.50	12.59	67.38	13.33	22.98	5.63	25.14	5.65	27.28	6.13	18.82	5.67	19.51	5.53	21.06	5.81
Classification																		
None	58.14	11.47	63.23	11.83	66.56	12.72	23.08	5.66	25.10	5.66	27.09	5.74	18.39	5.41	18.93	5.09	20.46	5.41
Support	63.43***	12.57	67.05***	12.36	69.54**	12.53	24.09***	5.45	26.06*	5.14	27.37	5.67	20.66***	5.46	21.31***	5.22	22.41***	5.44
Country of Origin																		
Not US	60.70*	11.58	65.43*	12.22	71.07***	12.44	23.80	5.58	25.95*	5.61	28.12***	5.78	19.47*	5.22	19.83	5.16	22.64***	5.08
US Born	58.87	11.99	63.76	11.99	66.41	12.63	23.15	5.63	25.15	5.53	26.92	5.68	18.71	5.57	19.40	5.23	20.56	5.51

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

10.3389/fpsyg.2023.977680

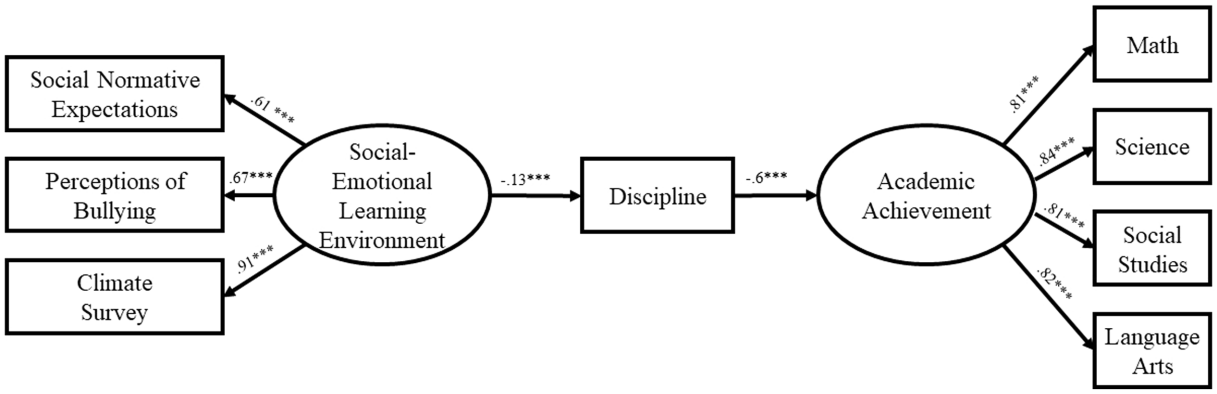


FIGURE 1
Model at year 1.

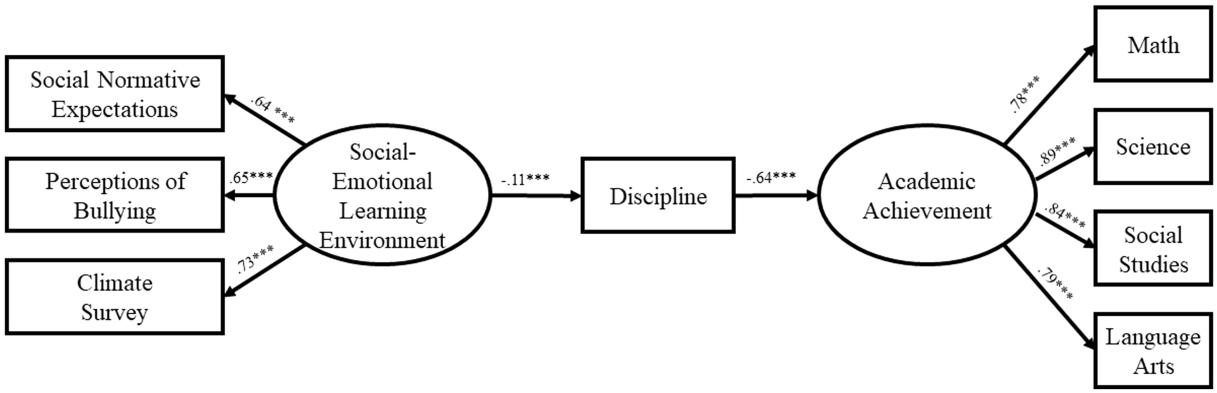


FIGURE 2
Model at year 2.

variability of some factors (e.g., overrepresentation of 70% or more), exploration into the role of demographic factors was limited. Finally, the current study does not address the intervention itself, only the theory of the proposed change model. As a result, the current study cannot provide evidence for a causal link between the social–emotional learning environment, discipline and academic achievement, nor can we validate that changing student perceptions of the social–emotional learning environment will invariably result in a change in academic achievement. Our findings use the available data to identify the relationships between the constructs targeted in an intervention, but do not provide longitudinal or causal evidence for the efficacy of the mechanism of change itself. In light of these limitations, the results here can be considered a first step towards further research to support the intervention implications, particularly individual level analysis to test the efficacy of hypothesized mechanism for change across a range of settings.

4.2. Future research

The current study tested a hypothesized conceptual model that theorized a specific mechanism for change in an urban middle school: that improved academic achievement can occur as a function of perceptions of social–emotional environment and disciplinary experiences. The results are promising, as the logic model was found to be supported, with constructs relationally linked in a valid path model. The context in which this conceptual model for change was evaluated reflected a "high needs" population, thus, any factors that influence students' achievement outcomes may present a valuable next step in resilience research. The results of this study also suggest that future research would benefit from expanded exploration of interventions targeted at these factors. If perceptions of climate, bullying and social expectations impact behavior, and which then impacts academic achievement, it may be that this relationship represents an area of resilience that can be enhanced deliberately by intervention programing that is coordinated with the elements of the model and evaluated more explicitly in sequence. To fully test the efficacy of school-level intervention programs, further research must occur in a range of schools over a number of years to see if systems level change can be executed through the path mechanism identified here. In an era when both student mental health and academic achievement are in a state of distress due to global factors beyond an individual students' control, it is important to understand what can support our students in reaching their potential.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

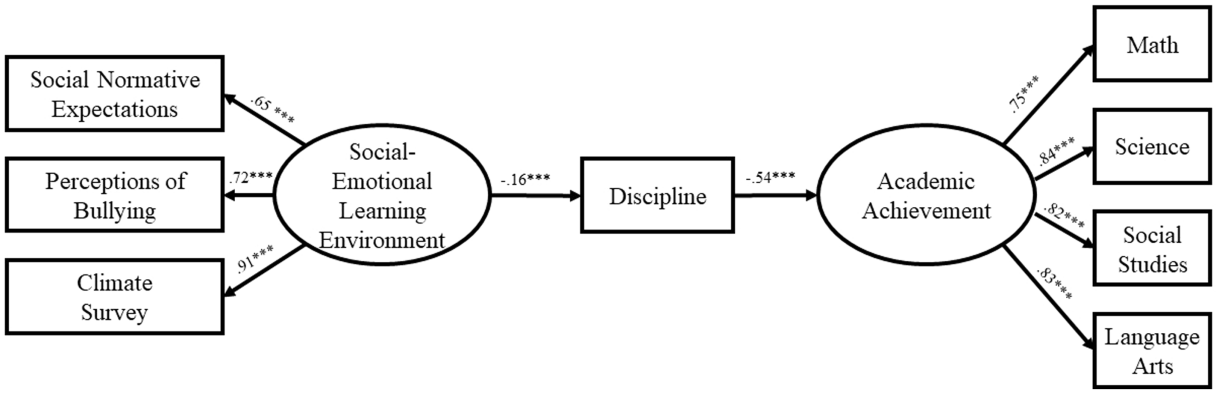


FIGURE 3
Model at year 3.

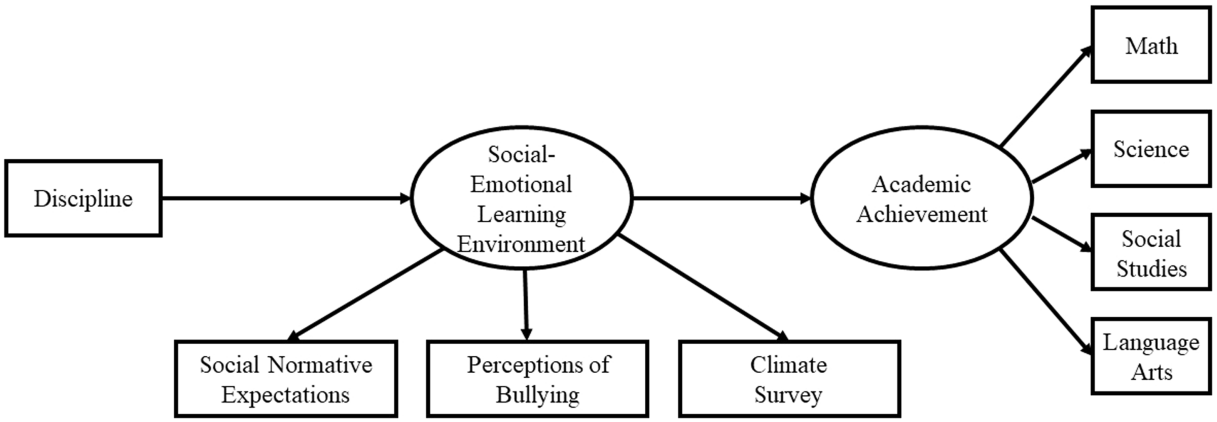


FIGURE 4
Proposed alternative model.

Ethics statement

The studies involving human participants were reviewed and approved by Rutgers University Institutional Review Board. Written informed consent from the participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

ME contributed to conception and design of the study. GW, DH, and EV contributed to the implementational and data collection, as well as the organization of the database. GW and CS performed the statistical analyses. GW wrote the first draft of the manuscript. DH, EV, and ME contributed significantly to its revisions. MY and AW contributed to the literature review and revisions. MY and CS both wrote sections of the manuscript. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.977680/full#supplementary-material

References

Anderson, K., Kochan, F., Kensler, L. A. W., and Reames, E. H. (2018). Academic optimism, enabling structures, and student achievement: delving into relationships. *J. Sch. Leadersh.* 28, 434–461. doi: 10.1177/105268461802800401

Anderson, K. P., Ritter, G. W., and Zamarro, G. (2019). Understanding a vicious cycle: the relationship between student discipline and student academic outcomes. *Educ. Res.* 48, 251–262. doi: 10.3102/0013189X19848720

Anyon, Y., Jenson, J. M., Altschul, I., Farrar, J., McQueen, J., Greer, E., et al. (2014). The persistent effect of race and the promise of alternatives to suspension in school discipline outcomes. *Child Youth Serv. Rev.* 44, 379–386. doi: 10.1016/j.childyouth.2014.06.025

Anyon, Y., Lechuga, C., Ortega, D., Downing, B., Greer, E., and Simmons, J. (2018). An exploration of the relationships between student racial background and the school sub-contexts of office discipline referrals: a critical race theory analysis. *Race Ethn. Educ.* 21, 390–406. doi: 10.1080/13613324.2017.1328594

Arbuckle, J. L., and Wothke, W. (2006). *AMOS user's guide 7.0*. Spring House, PA: AMOS Development Corporation.

Aygün, H. E., and Taşkın, Ç. Ş. (2022). The effect of social-emotional learning program on social-emotional skills, academic achievement and classroom climate. *Psycho-Educ. Res. Rev.* 11, 59–76. doi: 10.52963/PERR_Biruni_V11.N1.05

Barksdale, C., Peters, M. L., and Corrales, A. (2021). Middle school students' perceptions of classroom climate and its relationship to achievement. *Educ. Stud.* 47, 84–107. doi: 10.1080/03055698.2019.1664411

Bell, P. J., White, G. W., Hatchimonji, D. R., Stepney, C. T., Linsky, A. V., Vaid, E., et al. (2019). Social-normative expectations mediates school climate's association with academic achievement in latino middle school students. *Educ. Urban Soc.* 51, 374–394. doi: 10.1177/0013124517719972

Bouman, T., van der Meulen, M., Goossens, F. A., Olthof, T., Vermande, M. M., and Aleva, E. A. (2012). Peer and self-reports of victimization and bullying: their differential association with internalizing problems and social adjustment. *J. Sch. Psychol.* 50, 759–774. doi: 10.1016/j.jsp.2012.08.004

Bradshaw, C. P., Cohen, J., Espelage, D. L., and Nation, M. (2021). Addressing school safety through comprehensive school climate approaches. *Sch. Psychol. Rev.* 50, 221–236. doi: 10.1080/2372966X.2021.1926321

Brumley, L. D., Jaffee, S. R., and Brumley, B. P. (2017). Pathways from childhood adversity to problem behaviors in young adulthood: the mediating role of adolescents' future expectations. *J. Youth Adolesc.* 46, 1–14. doi: 10.1007/s10964-016-0597-9

Bryson, S. L., and Childs, K. K. (2018). Racial and ethnic differences in the relationship between school climate and disorder. *Sch. Psychol. Rev.* 47, 258–274. doi: 10.17105/SPR-2018-0016.V47-3

Buzzai, C., Sorrenti, L., Costa, S., Toffle, M. E., and Filippello, P. (2021). The relationship between school-basic psychological need satisfaction and frustration, academic engagement and academic achievement. *Sch. Psychol. Int.* 42, 497–519. doi: 10.1177/01430343211017170

Capp, G., Astor, R. A., and Gilreath, T. (2020). School staff members in California: how perceptions of school climate are related to perceptions of student risk and well-being. *J. Soc. Soc. Work Res.* 11, 415–442. doi: 10.1086/710974

Cohen, P. R., Carlsson, A., Ballesteros, L., and Amant, R. S. (1993). Automating path analysis for building causal models from data. *Proceedings of the International Workshop on Machine Learning*, 57–64.

Cook, C. R., Coco, S., Zhang, Y., Fiat, A. E., Duong, M. T., Renshaw, T. L., et al. (2018). Cultivating positive teacher–student relationships: preliminary evaluation of the establish–maintain–restore (EMR) method. *Sch. Psychol. Rev.* 47, 226–243. doi: 10.17105/SPR-2017-0025.V47-3

Corcoran, R. P., Cheung, A. C., Kim, E., and Xie, C. (2018). Effective universal school-based social and emotional learning programs for improving academic achievement: a systematic review and meta-analysis of 50 years of research. *Educ. Res. Rev.* 25, 56–72. doi: 10.1016/j.edurev.2017.12.001

Credé, M., Tynan, M. C., and Harms, P. D. (2017). Much ado about grit: a meta-analytic synthesis of the grit literature. *J. Pers. Soc. Psychol.* 113, 492–511. doi: 10.1037/pspp0000102

Cruz, R. A., Firestone, A. R., and Rodl, J. E. (2021). Disproportionality reduction in exclusionary school discipline: a best-evidence synthesis. *Rev. Educ. Res.* 91, 397–431. doi: 10.3102/0034654321995255

Dahl, R. E., Allen, N. B., Wilbrecht, L., and Suleiman, A. B. (2018). Importance of investing in adolescence from a developmental science perspective. *Nature* 554, 441–450. doi: 10.1038/nature25770

Daily, S. M., Mann, M. J., Kristjansson, A. L., Smith, M. L., and Zullig, K. J. (2019). School climate and academic achievement in middle and high school students. *J. Sch. Health* 89, 173–180. doi: 10.1111/josh.12726

Damasio, A. R. (1999). *The feeling of what happens: Body and emotion in the making of consciousness*. New York, NY: Harcourt Brace and Co. doi. 10.1353/jsp.2001.0038

Damasio, A., and Carvalho, G. B. (2013). The nature of feelings: evolutionary and neurobiological origins. *Nat. Rev. Neurosci.* 14, 143–152. doi: 10.1038/nrn3403

Darling-Hammond, L., and Cook-Harvey, C. M. (2018). *Educating the whole child: Improving school climate to support student success*. Palo Alto, CA: Learning Policy Institute. doi:10.54300/145.655.

De Witte, K., Cabus, S., Thyssen, G., Groot, W., and van den Brink, H. M. (2013). A critical review of the literature on school dropout. *Educ. Res. Rev.* 10, 13–28. doi: 10.1016/j.edurev.2013.05.002

Del Toro, J., and Wang, M.-T. (2021). Longitudinal inter-relations between school cultural socialization and school engagement among urban early adolescents. *J. Youth Adolesc.* 50, 978–991. doi: 10.1007/s10964-020-01377-w

Dermody, C., Dusenbury, L., Greenberg, M., Godek, D., Connor, P., Cross, R., et al. (2022). *A developmental framework for the integration of social and emotional learning and career and workforce development*. Chicago, IL: CASEL.

Dijkstra, J. K., and Gest, S. D. (2015). Peer norm salience for academic achievement, prosocial behavior, and bullying: implications for adolescent school experiences. *J. Early Adolesc.* 35, 79–96. doi: 10.1177/0272431614524303

Elias, M. J. (2014). The future of character education and social-emotional learning: "the need for whole school and community-linked approaches". *J. Character Educ.* 10, 37–42.

Elias, M. J., Zins, J. E., Weissberg, R. P., Frey, K. S., Greenberg, M. T., Haynes, N. M., Kessler, R., Schwab-Stone, M.E., and Shriver, T. P. (1997). *Promoting social and emotional learning: Guidelines for educators*. Association for Supervision and Curriculum Development, Virginia, USA

Espelage, D. L., and Hong, J. S. (2019). "School climate, bullying, and school violence" in *School safety and violence prevention: Science, practice, policy*. eds. M. J. Mayer and S. R. Jimerson (Washington, DC: American Psychological Association), 45–69. doi: 10.1037/0000106-003

Eugene, D. R. (2020). A multilevel model for examining perceptions of school climate, socioeconomic status, and academic achievement for secondary school students. *JESPAR* 25, 79–99. doi: 10.1080/10824669.2019.1670067

Farrington, C. A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T. S., Johnson, D. W., and Beechum, N. O. (2012). Teaching adolescents to become learners: the role of noncognitive factors in shaping school performance--A critical literature review. In Consortium on Chicago School Research. Consortium on Chicago School Research. Available at: https://eric.ed.gov/?id=ED542543

Fauth, B., Decristan, J., Decker, A. T., Büttner, G., Hardy, I., Klieme, E., et al. (2019). The effects of teacher competence on student outcomes in elementary science education: the mediating role of teaching quality. *Teach. Teach. Educ.* 86:102882. doi: 10.1016/j.tate.2019.102882

Felmlee, D., McMillan, C., Inara Rodis, P., and Osgood, D. W. (2018). Falling behind: lingering costs of the high school transition for youth friendships and grades. *Sociol. Educ.* 91, 159–182. doi: 10.1177/0038040718762136

Fitzpatrick, C., Côté-Lussier, C., Pagani, L. S., and Blair, C. (2015). I Don't think you like me very much: child minority status and disadvantage predict relationship quality with teachers. *Youth Soc.* 47, 727–743. doi: 10.1177/0044118X13508962

Gage, N. A., Larson, A., Sugai, G., and Chafouleas, S. M. (2016). Student perceptions of school climate as predictors of office discipline referrals. *Am. Educ. Res. J.* 53, 492–515. doi: 10.3102/0002831216637349

Gašević, D., Dawson, S., Rogers, T., and Gasevic, D. (2016). Learning analytics should not promote one size fits all: the effects of instructional conditions in predicting academic success. *Internet High. Educ.* 28, 68–84. doi: 10.1016/j.iheduc.2015.10.002

Geiser, S., and Santelices, M. V. (2007). Validity of high-school grades in predicting student success beyond the freshman year: High-school record vs. standardized tests as indicators of four-year college outcomes. *Research and Occasional Paper Series: CSHE. 6.07*. Berkeley, CA: Center for studies in higher education, University of California.

Gerard, J. M., and Booth, M. Z. (2015). Family and school influences on adolescents' adjustment: the moderating role of youth hopefulness and aspirations for the future. *J. Adolesc.* 44, 1–16. doi: 10.1016/j.adolescence.2015.06.003

Gerlinger, J., Viano, S., Gardella, J. H., Fisher, B. W., Chris Curran, F., and Higgins, E. M. (2021). Exclusionary school discipline and delinquent outcomes: A meta-analysis. *J. Youth Adolesc.* 50, 1493–1509. doi: 10.1007/s10964-021-01459-3

Goleman, D. (1995). *Emotional intelligence: Why it can matter more than IQ*. Bantam Books, Penguin Random House LLC, New York, NY

Grazia, V., and Molinari, L. (2021). School climate multidimensionality and measurement: a systematic literature review. *Res. Pap. Educ.* 36, 561–587. doi: 10.1080/02671522.2019.1697735

Gregory, A., Osher, D., Bear, G. G., Jagers, R. J., and Sprague, J. R. (2021). Good intentions are not enough: centering equity in school discipline reform. *Sch. Psychol. Rev.* 50, 206–220. doi: 10.1080/2372966X.2020.1861911

Gremmen, M. C., Dijkstra, J. K., Steglich, C., and Veenstra, R. (2017). First selection, then influence: developmental differences in friendship dynamics regarding academic achievement. *Dev. Psychol.* 53, 1356–1370. doi: 10.1037/dev0000314

Halliday, S., Gregory, T., Taylor, A., Digenis, C., and Turnbull, D. (2021). The impact of bullying victimization in early adolescence on subsequent psychosocial and academic outcomes across the adolescent period: a systematic review. *J. Sch. Violence* 20, 351–373. doi: 10.1080/15388220.2021.1913598

Hamlin, D. (2021). Can a positive school climate promote student attendance? Evidence from New York City. *Am. Educ. Res. J.* 58, 315–342. doi: 10.3102/0002831220924037

Herman, R. (2012). Scaling school turnaround. *JESPAR* 17, 25–33. doi: 10.1080/10824669.2012.637166

Herrenkohl, T. I., Lee, J., and Hawkins, J. D. (2012). Risk versus direct protective factors and youth violence: Seattle social development project. *Am. J. Prev. Med.* 43, S41–S56. doi: 10.1016/j.amepre.2012.04.030

Huang, L. (2022). Exploring the relationship between school bullying and academic performance: the mediating role of students' sense of belonging at school. *Educ. Stud.* 48, 216–232. doi: 10.1080/03055698.2020.1749032

Hubbard, L. (2005). The role of gender in academic achievement. *Int. J. Qual. Stud. Educ.* 18, 605–623. doi: 10.1080/09518390500224887

IBM Corp. Released 2021. *IBM SPSS Statistics for Windows, Version 28.0*. Armonk, NY: IBM Corp

Jia, Y., Konold, T. R., and Cornell, D. (2016). Authoritative school climate and high school dropout rates. *Sch. Psychol. Q.* 31, 289–303. doi: 10.1037/spq0000139

Jussim, L., Robustelli, S. L., and Cain, T. R. (2009). "Teacher Expectations and Self-Fulfilling Prophecies" in *Handbook of Motivation at School*. eds. K. R. Wentzel and D. B. Miele (New York, NY: Routledge), 1–32.

Juvonen, J., Wang, Y., and Espinoza, G. (2011). Bullying experiences and compromised academic performance across middle school grades. *J. Early Adolesc.* 31, 152–173. doi: 10.1177/0272431610379415

Kanesan, P., and Fauzan, N. (2019). Models of emotional intelligence: A review. *e-Bangi* 7, 1–9.

Kim, Y. K., Sanders, J. E., Makubuya, T., and Yu, M. (2020). Risk factors of academic performance: experiences of school violence, school safety concerns, and depression by gender. *Child Youth Care Forum* 49, 725–742. doi: 10.1007/s10566-020-09552-7

Kline, R. B. (2011). "Convergence of structural equation modeling and multilevel modeling" in *Handbook of methodological innovation in social research methods*. eds. E. M. Williams and W. P. Vogt (London: Sage), 562–589.

Kyriakides, L., Creemers, B. P., Panayiotou, A., and Charalambous, E. (2020). *Quality and equity in education: Revisiting theory and research on educational effectiveness and improvement*. London: Routledge. doi:10.4324/9780203732250

La Salle, T. P., Roach, A. T., and McGrath, D. (2013). The relationship of IEP quality to curricular access and academic achievement for students with disabilities. *Int. J. Spec. Educ.* 28, 135–144.

Lacoe, J., and Steinberg, M. P. (2018). Rolling Back zero tolerance: the effect of discipline policy reform on suspension usage and student outcomes. *Peabody J. Educ.* 93, 207–227. doi: 10.1080/0161956X.2018.1435047

Lester, L., and Cross, D. (2015). The relationship between school climate and mental and emotional wellbeing over the transition from primary to secondary school. *Psychol. Well-Being* 5:9. doi: 10.1186/s13612-015-0037-8

Liang, X., and Zhang, G. (2009). Indicators to evaluate pre-service teachers' cultural competence. *Eval. Res. Educ.* 22, 17–31. doi: 10.1080/09500790903082354

Lickona, T., and Davidson, M. (2003). *School as a caring community profile-II: A survey of students, staff, and parents*. Southern University of New York. Available at: https://oira.cortland.edu/msche/instschoolcharacter.pdf

Linsky, A. C., Hatchimonji, D., Kruzik, C., Kifer, S., Franza, N., McClain, K., et al. (2018). Students taking action together: social action in urban middle schools. *Middle Sch. J.* 49, 4–14. doi: 10.1080/00940771.2018.1488469

Maxwell, S., Reynolds, K. J., Lee, E., Subasic, E., and Bromhead, D. (2017). The impact of school climate and school identification on academic achievement: multilevel modeling with student and teacher data. *Front. Psychol.* 8, 1–21. doi: 10.3389/fpsyg.2017.02069

McCoach, D. B., Yu, H., Gottfried, A. W., and Gottfried, A. E. (2017). Developing talents: A longitudinal examination of intellectual ability and academic achievement. *High Abil. Stud.* 28, 7–28. doi: 10.1080/13598139.2017.1298996

McKinley Yoder, C., Cantrell, M. A., and Hinkle, J. L. (2022). Disparities in high school graduation by identity and disability using intermediate and long-term educational outcomes. *J. Sch. Nurs.* 10598405221078988:105984052210789. doi: 10.1177/10598405221078989

McKown, C., and Weinstein, R. S. (2008). Teacher expectations, classroom context, and the achievement gap. *J. Sch. Psychol.* 46, 235–261. doi: 10.1016/j.jsp.2007.05.001

Miller, D. T., and Ross, M. (1975). Self-serving biases in the attribution of causality: fact or fiction? *Psychol. Bull.* 82, 213–225. doi: 10.1037/h0076486

Moreno, G., and Segura-Herrera, T. (2013). Special education referrals and disciplinary actions for Latino students in the United States. *Multicult. Learn.Teach.* 9, 33–51. doi: 10.1515/mlt-2013-0022

Morris, E. W., and Perry, B. L. (2017). Girls behaving badly? Race, gender, and subjective evaluation in the discipline of African American girls. *Sociol. Educ.* 90, 127–148. doi: 10.1177/0038040717694876

Murano, D., Sawyer, J. E., and Lipnevich, A. A. (2020). A meta-analytic review of preschool social and emotional learning interventions. *Rev. Educ. Res.* 90, 227–263. doi: 10.3102/0034654320914743

O'Connor, K. E., Hitti, S. A., Thompson, E. L., Farrell, A. D., and Sullivan, T. N. (2020). Perceptions of school climate among subgroups of aggressive and victimized youth. *Sch. Ment. Heal.* 12, 169–181. doi: 10.1007/s12310-019-09343-z

Olivier, E., Archambault, I., and Dupéré, V. (2020). Do needs for competence and relatedness mediate the risk of low engagement of students with behavior and social problem profiles? *Learn. Individ. Differ.* 78:101842. doi: 10.1016/j.lindif.2020.101842

Ou, S.-R., and Reynolds, A. J. (2008). Predictors of educational attainment in the Chicago Longitudinal Study. *School Psychology Quarterly* 23, 199–229.

Osher, D., Kidron, Y., Brackett, M., Dymnicki, A., Jones, S., and Weissberg, R. (2016). Advancing the science and practice of social and emotional learning: looking back and moving forward. *Rev. Res. Educ.* 40, 644–681. doi: 10.3102/0091732X16673595

Porter, R. P. (2000). Accountability is overdue: testing the academic achievement of limited English proficient students. *Appl. Meas. Educ.* 13, 403–410. doi: 10.1207/S15324818AME1304_06

Reaves, S., McMahon, S. D., Duffy, S. N., and Ruiz, L. (2018). The test of time: a meta-analytic review of the relation between school climate and problem behavior. *Aggress. Violent Behav.* 39, 100–108. doi: 10.1016/j.avb.2018.01.006

Rocque, M., Jennings, W. G., Piquero, A. R., Ozkan, T., and Farrington, D. P. (2017). The importance of school attendance: findings from the Cambridge study in delinquent development on the life-course effects of truancy. *Crime Delinq.* 63, 592–612. https://doi:10.1177001112871666052O. doi: 10.1177/0011128716660520

Roorda, D. L., Koomen, H. M., Spilt, J. L., and Oort, F. J. (2011). The influence of affective teacher–student relationships on students' school engagement and achievement: A meta-analytic approach. *Rev. Educ. Res.* 81, 493–529. doi: 10.3102/0034654311421793

Saadat, S., Kord, N., and Jalali, M. (2019). The role of psychological capital (self-efficacy, Hope, resilience, and optimism) in student achievement motivation. *Rooyesh-e-Ravanshenasi J.* 8, 167–174.

Salovey, P., and Mayer, J. D. (1990). Emotional intelligence. *Imagin. Cogn. Pers.* 9, 185–211. doi: 10.2190/DUGG-P24E-52WK-6CDG

Samuels, W. E., Tournaki, N., Blackman, S., and Zilinski, C. (2016). Executive functioning predicts academic achievement in middle school: a four-year longitudinal study. *J. Educ. Res.* 109, 478–490. doi: 10.1080/00220671.2014.979913

Santiago, C. D., Gudiño, O. G., Baweja, S., and Nadeem, E. (2014). Academic achievement among immigrant and US-born Latino adolescents: associations with cultural, family, and acculturation factors. *J. Commun. Psychol.* 42, 735–747. doi: 10.1002/jcop.21649

Seeskin, A., Nagaoka, J., and Mahaffie, S. (2018). Hidden risk: changes in GPA across the transition to high school. Research Report. University of Chicago Consortium on School Research.

Shepperd, J., Malone, W., and Sweeny, K. (2008). Exploring causes of the self-serving bias. *Soc. Personal. Psychol. Compass* 2, 895–908. doi: 10.1111/j.1751-9004.2008.00078.x

Skiba, R. J., Horner, R. H., Chung, C.-G., Rausch, M. K., May, S. L., and Tobin, T. (2011). Race is not neutral: a National Investigation of African American and Latino disproportionality in school discipline. *Sch. Psychol. Rev.* 40, 85–107. doi: 10.1080/02796015.2011.12087730

Sommerfeld, A. K. (2016). Education as a collective accomplishment: how personal, peer, and parent expectations interact to promote degree attainment. *Soc. Psychol. Educ.* 19, 345–365. doi: 10.1007/s11218-015-9325-7

Sorensen, L. C., Bushway, S. D., and Gifford, E. J. (2021). Getting tough? The Effects of Discretionary Principal Discipline on Student Outcomes. In EdWorkingPapers.com. Annenberg Institute at Brown University. Available at: https://www.edworkingpapers.com/ai20-216

Spilt, J. L., Koomen, H. M., and Thijs, J. T. (2011). Teacher wellbeing: the importance of teacher–student relationships. *Educ. Psychol. Rev.* 23, 457–477. doi: 10.1007/s10648-011-9170-y

Steinmayr, R., Heyder, A., Naumburg, C., Michels, J., and Wirthwein, L. (2018). School-related and individual predictors of subjective well-being and academic achievement. *Front. Psychol.* 9:2631. doi: 10.3389/fpsyg.2018.02631

Stoddard, S. A., and Pierce, J. (2015). Promoting positive future expectations during adolescence: the role of assets. *Am. J. Commun. Psychol.* 56, 332–341. doi: 10.1007/s10464-015-9754-7

Strunk, K. O., Marsh, J. A., Hashim, A. K., Bush-Mecenas, S., and Weinstein, T. (2016). The impact of turnaround reform on student outcomes: evidence and insights from the Los Angeles Unified School District. *Educ. Finance Policy* 11, 251–282. doi: 10.1162/EDFP_a_00188

Swearer, S. M., Espelage, D. L., Vaillancourt, T., and Hymel, S. (2010). What can be done about school bullying? Linking research to educational practice. *Educational researcher* 39, 38–47.

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

Thapa, A., Cohen, J., Guffey, S., and Higgins-D'Alessandro, A. (2013). A review of school climate research. *Rev. Educ. Res.* 83, 357–385. doi: 10.3102/0034654313483907

Thompson, E. (2019). Understanding bullying and the necessity for prevention and intervention in schools. Senior Honors Theses. Available at: https://digitalcommons.liberty.edu/honors/869

U.S. Bureau of Labor Statistics (2019). Available at: https://www.bls.gov/home.htm [Accessed November 21, 2021]

Vaid, E., Jofen, T., Hatchimonji, D. R., DeMarchena, S. L., White, G., and Elias, M. J. (2023). The impact of school climate in social-normative expectations in low and high SES schools. *Journal of Community Psychology* 51, 219–233. doi: 10.1002/jcop.22898

Valente, S., Monteiro, A. P., and Lourenço, A. A. (2019). The relationship between teachers' emotional intelligence and classroom discipline management. *Psychol. Sch.* 56, 741–750. doi: 10.1002/pits.22218

Wang, M.-T., and Degol, J. L. (2016). School climate: a review of the construct, measurement, and impact on student outcomes. *Educ. Psychol. Rev.* 28, 315–352. doi: 10.1007/s10648-015-9319-1

Wang, M.-T., Degol, L., Amemiya, J., Parr, A., and Guo, J. (2020). Classroom climate and children's academic and psychological wellbeing: A systematic review and meta-analysis. *Dev. Rev.* 57:100912. doi: 10.1016/j.dr.2020.100912

Weissberg, R. P., Durlak, J. A., Domitrovich, C. E., and Gullotta, T. P. (2015). "Social and Emotional Learning: Past, Present and Future" in *Handbook for Social and Emotional Learning*. (New York, NY: Guilford), 3–19.

White, G. W., Stepney, C. T., Hatchimonji, D. R., Moceri, D. C., Linsky, A. V., Reyes-Portillo, J. A., et al. (2016). The increasing impact of socioeconomics and race on standardized academic test scores across elementary, middle, and high school. *Am. J. Orthop.* 86, 10–23. doi: 10.1037/ort0000122

Williams, S., Schneider, M., Wornell, C., and Langhinrichsen-Rohling, J. (2018). Student's perceptions of school safety: it is not just about being bullied. *J. Sch. Nurs.* 34, 319–330. doi: 10.1177/1059840518761792

Williams, K. R., and Guerra, N. G. (2007). Prevalence and predictors of internet bullying. *J. Adoles. Health.* 41, S14–S21.

Wortham, S., Love-Jones, R., Peters, W., Morris, S., and García-Huidobro, J. C. (2020). Educating for comprehensive well-being. *ECNU Rev. Educ.* 3, 406–436. doi: 10.1177/2096531120928448

Yeager, D. S., and Dweck, C. S. (2012). Mindsets that promote resilience: when students believe that personal characteristics can be developed. *Educ. Psychol.* 47, 302–314. doi: 10.1080/00461520.2012.722805

Zajda, J. (2021). "Discrimination and self-fulfilling prophecy in schools globally" in *Globalisation and education reforms: Creating effective learning environments*. ed. J. Zajda (Switzerland: Springer International Publishing), 51–70. doi: 10.1007/978-3-030-71575-5

Zullig, K. J., Ward, R. M., Huebner, E. S., and Daily, S. M. (2018). Association between adolescent school climate and perceived quality of life. *Child Indic. Res.* 11, 1737–1753. doi: 10.1007/s12187-017-9521-4

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 16 February 2023
DOI 10.3389/feduc.2023.1070591

OPEN ACCESS

EDITED BY
Celene Domitrovich,
Georgetown University,
United States

REVIEWED BY
Deborah Perry,
Georgetown University,
United States
Karyn Hartz,
University of Rochester,
United States

*CORRESPONDENCE
Ann M. Partee
✉ amp9as@virginia.edu

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Educational Psychology, a section of the journal Frontiers in Education

RECEIVED 15 October 2022
ACCEPTED 23 January 2023
PUBLISHED 16 February 2023

CITATION
Partee AM, Sachdeva S, Bivona MA, Clayback KA, Miller-Marshall S, Parker K, Alamos P, Frank C, Downer JT and Williford AP (2023) Implementation of an early childhood mental health consultation pilot in Virginia: Critical tensions and implications for scale-up.
Front. Educ. 8:1070591.
doi: 10.3389/feduc.2023.1070591

Implementation of an early childhood mental health consultation pilot in Virginia: Critical tensions and implications for scale-up

Ann M. Partee*, Shubhi Sachdeva, Marissa A. Bivona†, Kelsey A. Clayback†, Sasha Miller-Marshall†, Kyra Parker†, Pilar Alamos, Cloe Frank, Jason T. Downer and Amanda P. Williford

Center for Advanced Study of Teaching and Learning, University of Virginia, Charlottesville, VA, United States

Early childhood mental health consultation (ECMHC) is a targeted prevention service that aims to build the capacity of early care and education (ECE) professionals and foster supportive environments that promote children's social–emotional competence and improve mental health and well-being. A key challenge to delivering ECMHC at scale is navigating complex multi-level factors to maximize successful implementation and program benefits at scale. The current study describes the implementation tensions arising during the first year of a pilot ECMHC program conducted in partnership across multiple agencies and a state's department of education. In the 2021–2022 pilot year, ECMHC was offered as a free service to ECE programs in one large region of Virginia, with the goal of examining feasibility to scale statewide in future years. Consultation was implemented in 45 preschool classrooms across 30 programs. Implementation data were collected using consultation logs and participant surveys, and 20 participants (educators, families, program directors) participated in focus groups. Three implementation tensions are highlighted in this paper: (1) ideal plans versus reality of a new ECMHC roll-out; (2) how to support ECE professionals' practice as it relates to children's behavior, without contributing to a deficit view that children need to be "fixed;" and (3) systemic factors in the early childhood field that undermine the implementation and effectiveness of ECMHC. For each tension, we provide context from the larger literature on ECMHC, describe relevant decision points from Virginia's pilot ECMHC program, and present implementation data to illustrate these tensions in practice. We conclude with reflections on lessons learned that have implications for other ECMHC and SEL intervention scale-up efforts.

KEYWORDS

social–emotional learning, implementation, scale–up, early childhood mental health consultation, early childhood education

1. Introduction

During early childhood, young children have countless experiences with early care and education (ECE) providers and families that contribute to children's emerging social–emotional development, mental health, and well-being. Early childhood mental health is defined as "the developing capacity of the child from birth to 5 years old to form close and secure adult and peer relationships; experience, manage, and express a full range of emotions; and explore the environment and learn – all in the context of family, community, and culture" (Zero to Three, 2017). Children's mental health, synonymous with social–emotional competence in early childhood, develops through

their relationships with adults (Zeanah and Zeanah, 2019). Indeed, an emphasis on the importance of relationships between children and their adult caregivers is a distinguishing feature of mental health in early childhood from mental health in adolescence or adulthood (Zeanah and Zeanah, 2019). Social–emotional learning (SEL) programs and interventions seek to strengthen children's abilities to form secure relationships, manage emotions, and engage in their learning environment (McClelland et al., 2017). One type of SEL intervention is early childhood mental health consultation (ECMHC). ECMHC is a targeted prevention service that builds the capacity of ECE professionals to interact with young children in ways that promote their social–emotional competence and improve their mental health and well-being.

Critically, Virginia, along with the entire United States and rest of the world, is facing an unprecedented health, social, and economic crisis with the COVID-19 pandemic. Even before COVID-19, many young children, particularly those from low-income families and communities of color, faced traumatic experiences that have had a significant impact on their social–emotional competence and mental health (National Scientific Council on the Developing Child, 2012; Bartlett and Smith, 2019; Shonkoff et al., 2021). Further amplifying these existing inequities, there are well-documented disparities in the rates of COVID-19 infection, illness, and death among communities of color due to inequities in social determinants of health such as discrimination, crowded housing, and access to health care (Center for Disease Control and Prevention, 2022). Further, low-wage workers are more likely to experience job loss due to the pandemic, resulting in greater economic hardship, food insecurity, and evictions (Center on Budget and Policy Priorities, 2020). Children are impacted even if they do not directly experience these traumatic events. For instance, the pandemic adds stress to families who may be trying to simultaneously parent and work from home. Young children's daily routines have been upended and many had to adapt to remote learning for an extended period of time, which made it more difficult to form meaningful connections with their ECE providers[1] and peers. In this current context, there is a great need for high-quality and effective services that promote young children's social–emotional competence and mental health.

Early childhood education[2] is a key setting in which to situate efforts to enhance children's social–emotional competence and mental health as well as address and prevent any concerns in these areas from further progressing (Trigg and Keyes, 2019). Children spend a significant amount of time in ECE settings, and there is a long history of ECE programs supporting young children's social–emotional competence and mental health through relationship building, classroom curricula, and programming (McClelland et al., 2017). For over a decade, the Center on the Social and Emotional Foundations of Early Learning (CSEFEL) has provided resources nationally to ECE programs that support providers' implementation of the Pyramid Model for Promoting Social–Emotional Competence in Infants and Young Children framework (Hemmeter et al., 2006). However, addressing challenging behaviors in the classroom is still an area of stress for ECE providers (Friedman-Krauss et al., 2014; Clayback and Williford, 2022), and programs resort to exclusionary discipline such as suspensions and expulsions from early childhood programs and at inequitably higher rates for young Black boys (Albritton et al., 2019; Garro et al., 2021). Disproportionate use of exclusionary discipline with Black children is not explained by the level of disruptive behavior (Bradshaw et al., 2010), suggesting that racial bias may play a role (Gregory et al., 2017). Concerns about exclusionary discipline in early childhood and programs' disproportionate use with Black children led the U.S. Department of Health and Human Services and the U.S. Department of Education to release a policy statement with recommendations for early childhood programs to promote children's social–emotional competence and mental health and reduce exclusionary discipline (U.S. Department of Health and Human Services and Education, 2014). ECMHC was included as a recommended strategy in this policy statement and is increasingly being provided to early childhood teachers to support young children's social–emotional competence and mental health.

ECMHC in ECE programs is a prevention-oriented service that aims to build the capacity of ECE providers and families to foster supportive environments that promote children's social-emotional competence and mental health. Through a collaborative relationship, consultants with early childhood mental health expertise support ECE program staff, teachers, and families to prevent and address concerns related to children's mental health and behavior (Cohen and Kaufmann, 2005). Some ECMHC programs have an explicit goal to prevent and reduce the practice of suspensions and expulsions from ECE programs (Conners Edge et al., 2021). Consultation activities that are implemented with ECE providers and families vary depending on the specific ECMHC model and whether the consultation is focused on addressing issues at the program- (supports the overall quality and climate of the program), classroom- (supports classroom systems and processes that shape children's social–emotional development such as warm interactions, positive teacher-child relationships, and consistent routines), or child/family- (supports a child or family's mental health and/or behavioral needs) level (Hunter et al., 2016). Despite this variation, most ECMHC approaches incorporate common components, such as a referral system, needs assessment, feedback to key stakeholders, and strategy implementation; however, there is still much to be learned about how best to roll out these ECMHC features, particularly when trying to scale up availability of these resources beyond a single classroom or program.

In fall 2020, state legislation[3] in Virginia required that a workgroup composed of stakeholders in infant and early childhood mental health study the feasibility of adopting an ECMHC program to prevent suspensions and expulsions of young children attending ECE programs. Based on recommendations made by this workgroup, the Virginia Department of Education (VDOE) funded a university research center and a statewide early childhood service provider to develop, implement, and evaluate a birth-to-five pilot model of ECMHC in 2021–2022. Over the 2021–2022 pilot year, ECMHC services were delivered primarily in one large region of the state, but a key aim of the pilot was to understand and learn from implementation successes and barriers as the state

1 In this article, we use the terms providers, educators, teachers interchangeably to describe the adults who work as early childhood professionals and provide care and education to young children from birth through preschool in private, faith-based, public, and family day home settings.

2 Early childhood education (ECE), early childhood care and education (ECCE), and early childhood education and care (ECEC) are terms that are often used synonymously. In this paper, we define early childhood education (ECE) inclusive of early childhood programs that provide care and education to young children from birth through preschool in private, faith-based, public, and family day home settings.

3 House Joint Resolution No. 51 requested that the Virginia Departments of Education, Social Services, and Behavioral Health and Developmental Services convene a workgroup to study and provide recommendations on an ECMHC model and submit a report to the Virginia Governor and General Assembly. The workgroup report is available here.

considers a potential expansion of services state-wide. The ECMHC model was designed to serve providers of children birth-to-five in ECE settings, including child care centers, family day homes, Early Head Start and Head Start, and school-based ECE programs. Priority was given to programs that received public funds, but any program was eligible to receive services. Services were delivered to programs at no charge.

The goal of the current paper is to describe Virginia's pilot ECMHC program during its first year of implementation, with a particular focus on grappling with tensions that arose in the implementation process. We highlight three tensions that illustrate competing needs and values that arose when designing and rolling out this new ECMHC program to serve ECE programs: (1) the tension between ideal and pragmatic roll-out; (2) the tension between supporting teachers' practice as it relates to an individual child and contributing to a deficit view that children need to be "fixed"; and (3) the tension with addressing systemic factors impacting the ECE field that can undermine the implementation and effectiveness of ECMHC. For each tension, we provide context from the broader literature on ECMHC and describe the decision points that were made for Virginia's pilot ECMHC program. We present relevant implementation data to illustrate these tensions in practice and then reflect on lessons learned that have implications for other ECMHC scale-up efforts.

2. Conceptual model of implementation

Implementation of ECMHC involves the amount or dosage of consultation that is delivered to ECE professionals, the quality of consultation, the extent to which consultation matches participants' needs, and participants' responsiveness to services (Durlak and DuPre, 2008). Implementation frameworks emphasize myriad factors that influence dosage, quality, alignment, and responsiveness. For example, Domitrovich et al. (2008) propose a conceptual framework for understanding implementation fidelity of school-based interventions. The authors define an intervention as a set of features or practices (referred to as core elements) linked to an intended outcome. In this model, effective implementation is bolstered by a "support system," which may include pre-intervention training and ongoing coaching or consultation. This conceptual framework highlights multi-level factors that may influence implementation: (1) macro-level (e.g., federal, state, and local policies that impact schools; university partnerships; funding; leadership), (2) school-level (e.g., school policies; school and classroom climate; size), and (3) individual-level (e.g., educator professional and psychological characteristics; perception of the intervention). At all three levels of the model, these contextual factors are interdependent and influence quality of implementation and ultimately children's outcomes.

Though this framework has predominantly been applied to K-12 settings, we use the model to understand implementation of the Virginia ECMHC model in ECE contexts. Early childhood contexts differ in important ways from older grades (Hindman and Bustamante, 2019), and thus the framework by Domitrovich et al. (2008) must be modified to apply. At the macro-level, factors influencing implementation of ECMHC include the broader system of early childhood, which is characterized by a multitude of structural and policy challenges. For example, compared to the K-12 system, the early education system is underfunded, and educators experience low wages (Whitebook et al., 2014), high stress and turnover (Schaack et al., 2020; Doromal et al., 2022), and report lacking adequate professional development (Gomez et al., 2015; Schaack et al., 2022). We refer to the school-level of the model as "program-level," since many early childhood settings operate outside of a typical school system and include a variety of auspices (e.g., Head Start/Early Head Start, state-funded, private), each of which operates differently. At the program-level, many early childhood programs do not have protected planning time, in contrast to the K-12 system. As a result, ECE providers are forced to engage in professional development activities during other times such as nap time and before or after the school day (Fettig and Artman-Meeker, 2016). Education requirements and training opportunities are also more variable across programs in the ECE system compared to K-12. At the individual level, ECE providers need to be open to adopting new strategies or making shifts to their classroom practice to support children (Domitrovich et al., 2009; Cook et al., 2015; Domitrovich et al., 2019). Providers' own mental health and beliefs may act as a barrier or facilitator to implementation. For example, teachers report that stress is a major barrier to implementation (McGoey et al., 2014), and teacher burnout is associated with lower implementation fidelity (Domitrovich et al., 2009). In ECE settings, provider mental health may be especially relevant for implementation, given the macro-level factors mentioned above. Understanding how implementation frameworks developed for the K-12 context apply to ECE settings is important when bringing any social–emotional learning program into early learning settings for young children. In the context of ECMHC, beliefs and biases may also influence engagement in consultation if providers attribute behavioral difficulties to the child or family rather than factors that the provider has some control over (Nemer et al., 2019). Racial bias may be particularly salient, given our focus on using ECMHC to reduce exclusionary discipline and specifically eliminate racial disparities in these practices (Davis et al., 2020).

A key challenge to delivering ECMHC at scale is navigating these interdependent multi-level factors to maximize successful implementation and program benefits at scale. The current paper describes implementation of the Virginia ECMHC pilot and raises critical tensions that relate to factors at the macro-, program-, and individual-levels.

3. Implementation of early childhood mental health consultation in Virginia

Team members from a university research center and a statewide early childhood service provider designed a birth-5 ECMHC model, in partnership with the VDOE, that aligned with recommendations from the state's HJ51 workgroup report. Key components of the model include aligned infant/toddler and preschool services, an open referral system, and multi-tiered services based on identified needs. Though the Virginia ECMHC model is coordinated to ensure similar services are delivered across the entire birth-to-five continuum, this paper focuses explicitly on data stemming from services provided to preschool-aged classrooms and children.

3.1. Early childhood mental health consultation program components and procedures

Figure 1 displays Virginia's birth-to-five ECMHC model. The first step to implement ECMHC services was to seek referrals for children and teachers who needed support. The Virginia ECMHC team developed flyers, videos, and other recruitment materials which were sent to program

leaders, teachers, families, and communities through various channels. Those who were interested in requesting ECMHC services were asked to fill out an online referral form that included questions about the referral/request for services (e.g., name of program, locality, program type); contact information for follow-up (e.g., phone number or email of the person submitting the referral); and the reason for requesting/referring to services. Respondents were able to request consultation at the classroom level or more targeted support around one or a few individual children in a classroom. After a referral was received, teachers were asked to complete a teacher intake survey, and for child-specific referrals, families were asked to complete a family intake survey and family permission form. ECMHC services were not initiated until family permission was obtained. Families and teachers could elect to participate in services but decline that their data collected as part of the pilot be used for research purposes.

The service tier (i.e., higher intensity services versus lower intensity services) was determined based on identified needs. All classroom-wide referrals were assigned to the lower intensity service tier. For child-specific referrals, needs were assessed using two teacher-reported rating scales: the Preschool Expulsion Risk Measure (PERM; Gilliam and Reyes, 2018) and the Child Behavior Rating Scale (CBRS; Bronson et al., 1990). The PERM was the primary measure used to guide assignment to the service tier. In cases where PERM scores indicated substantial teacher frustration or risk of suspension/expulsion, the case was assigned to receive higher intensity services. Cases with low frustration or risk of suspension/expulsion were assigned to receive lower intensity services. The CBRS was used as a secondary measure in cases where the PERM score did not reach the research team's high or low benchmark score.

The lower intensity tier begins with a kick-off session in which teachers complete a self-assessment to identify two social–emotional teaching topics on which they would like to enhance their practice. Next, teachers engage in consultation cycles around the two selected topics. Consultation cycles involve the creation of an action plan with the consultant, teacher implementation of the action plan (videotaped when possible), and the consultant providing feedback to the teacher after observing their implementation of the action plan. The intended dosage for the lower intensive tier is the kick-off session and four consultation cycles, with each cycle consisting of a session and observation. The higher intensity tier also begins with a kick-off session. After the kick-off session, the consultant conducts a baseline observation of the classroom environment using the Teaching Pyramid Observation Tool-Short Form (TPOT-S) based on the research edition in Hemmeter et al. (2014) and the referred child's engagement in the classroom using the Individualized Classroom Assessment Scoring System (inCLASS; Downer et al., 2010), to gather key information to inform consultation. Consultation cycles follow the same format as the lower intensity tier, however, in the higher intensity tier, the strategies that are included in the action plan emerge from the consultant's observation and discussion with the teacher. In the higher intensity tier, consultants also support teacher-family collaboration and facilitate referrals to community services, when needed. The intended dosage for the higher intensity tier is the kick-off session and seven consultation cycles, with each cycle consisting of a session and observation. ECMHC services were offered in-person and virtually; however, in some areas, in-person services were not possible due to the distance between consultants' location and the program. Two full-time consultants and one part-time consultant were hired and trained by the university research center to serve preschool referrals.

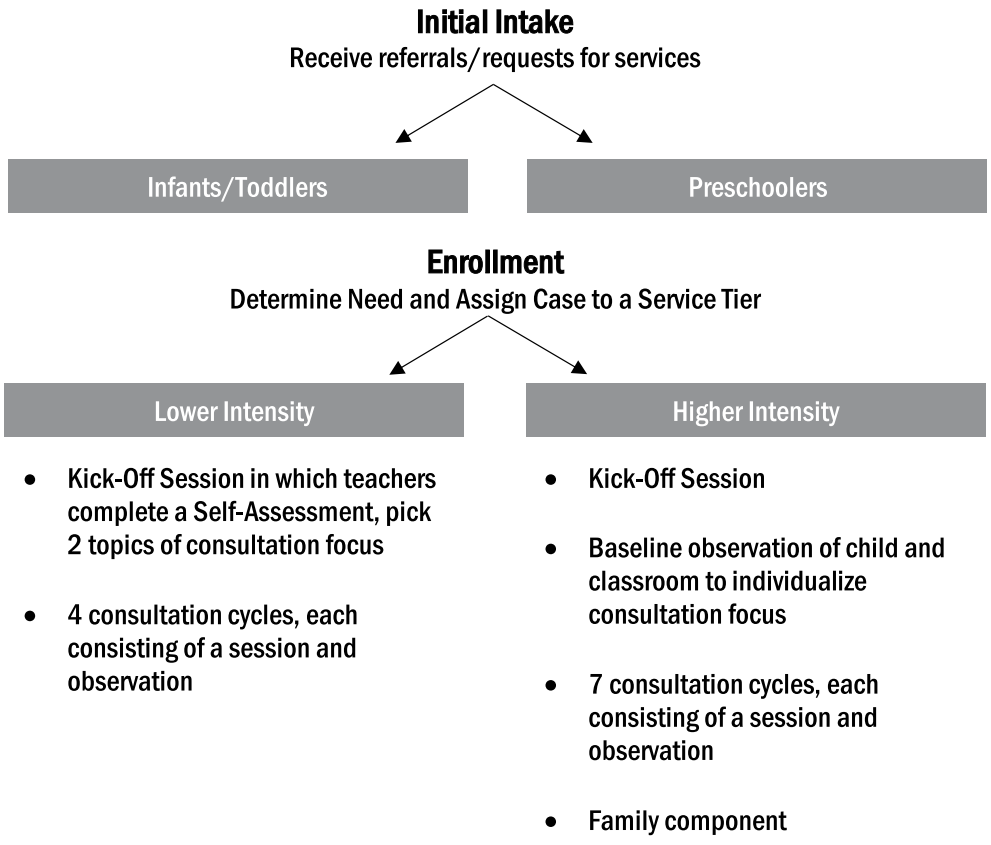


FIGURE 1
Birth-5 ECMHC model.

3.2. Theory of change

The ECMHC model implemented with teachers and families of preschool-aged children (defined here as children 36–60 months of age) draws heavily from a previously developed ECMHC model called Learning to Objectively Observe Kids (LOOK). Downer et al. (2018) describes the LOOK model's theory of change, core components, and initial evidence of its impact on teacher practices and preschool children's outcomes. Central to LOOK's theory of change is using guided video review to target ECE providers' beliefs and classroom practice to enhance children's social–emotional competence and mental health. Providers are asked to film themselves implementing evidence-based strategies with specific children whom the teacher perceives to display challenging behaviors. Consultants then select short video clips and write prompts that encourage providers to observe the child's engagement in the classroom and analyze their role in creating a supportive environment for the child. The guided video review is intended to promote teachers' understanding of the role of the classroom context for children's ability to effectively engage with ECE providers, peers, and tasks, to move away from a perspective that the child is the problem. Additionally, the guided video review helps providers link their use of strategies to improvements in children's engagement and behavior, thereby increasing providers' perceived self-efficacy to successfully respond to instances of challenging behavior. The Virginia ECMHC model incorporated LOOK's theory of change and use of guided video review to facilitate providers' implementation of evidence-based strategies and reflection on their practice.

3.3. Data collection and methods

To understand implementation of the Virginia ECMHC pilot, data were collected from consultants, teachers, program directors, and families using qualitative and quantitative methods. Regarding quantitative methods, consultants entered their consultation data into a consultant log. The consultant log collected information on the dosage of meetings between consultants and providers and families, the topics of those meetings, the dosage of observations conducted, the format of meetings and observations, and whether the consultant made referrals to external agencies to supplement consultation. Providers completed surveys at the onset and conclusion of consultation. Before and after consultation, teachers reported on their self-efficacy using a modified self-efficacy scale from Bandura (1997), emotional exhaustion using two items based on Jeon et al. (2018), and knowledge of early childhood social–emotional development and effective teacher practice (practitioner-developed items used in the pilot). Teachers also reported on their use of SEL-related resources in their classroom at the beginning of consultation. Additionally, for child-specific referrals, teachers reported pre- and post-consultation on children's self-regulation and social skills using the Child Behavior Rating Scale (Bronson et al., 1990), challenging behaviors using the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997), and expulsion risk using the Preschool Expulsion Risk Measure (PERM; Gilliam and Reyes, 2018). At the end of consultation, providers, families, and program directors were asked to respond to items asking about their experiences and satisfaction with the ECMHC pilot. Teacher and child sociodemographic data were collected *via* teacher and family surveys. Quantitative data that are presented in this paper include: sociodemographic characteristics; dosage of consultation; teacher-reported use of SEL-related resources; and teacher emotional exhaustion due to children's behaviors pre-consultation ("I am emotionally exhausted by children's behaviors," rated on a 5-point scale from Strongly Disagree to Strongly Agree).

To gain a better understanding of the experiences of teachers, families, and program directors who engaged with the ECMHC program, we also employed qualitative research methods to hear the stories and lived experiences of our participants. We did this in two ways. First, we conducted one-on-one interviews with participants who chose *not to* participate in ECMHC or opted out after consultation began. Second, we held video-cued focus groups (Adair and Kurban, 2019; Tobin, 2019) with participants who had more sustained engagement with ECMHC. Participants were contacted *via* email or phone and invited to participate in the interview or focus group. A $50 e-gift card was offered to incentivize participation. Six program directors and one provider participated in an interview. Eight focus groups were held with a total of 10 teachers, six parents, and four program directors. Rates of participation were as follows: 7.5% for program director focus group, 10% for family focus group, 15% for teacher focus group, and 35% for interviews. At the beginning of each focus group, participants were shown 3 video clips of children in early childhood classrooms. Each video depicted scenes that would be considered relevant to discussions about children's mental health, classroom behavioral expectations, and teachers' projected roles within those contexts. We chose scenes depicting conflict between children, teachers teaching social–emotional skills, and children throwing objects, kicking, and hitting other children, since these were some of the common reasons children were referred for ECMHC services.

Interviews and video-cued focus groups were facilitated virtually over Zoom and were video recorded and transcribed with participants' permission. The transcripts were first checked for accuracy before they went through several rounds of qualitative analysis. For this analysis, we used an inductive approach (Creswell and Poth, 2016; Miles et al., 2018). In the first round, a faculty member who led and participated in all the qualitative interviews and focus groups read through the transcripts and did open coding (Glaser, 2016) and identified thematic codes. In this round, some of the codes that emerged were communication issues, challenges with modalities of how services were offered (online vs. in-person), and lack of access to available resources for families. The researcher then used axial coding (Scott and Medaugh, 2017) to find interconnected thematic codes. This helped to connect interconnected categories. For instance, "challenges in communication with consultants" was connected to "challenges to uptake of services," as well as "suggestions to improve services." Similarly, "lack of adequately trained teachers" and "teachers leaving ECE" was connected to "systemic challenges faced by programs and families." The data began to show three main interconnected themes: (1) challenges to uptake of ECMHC services, (2) systemic challenges faced by families and programs, and (3) suggestions to improve ECMHC services. Next, this researcher went through the data again and did selective coding (Williams and Moser, 2019) for the three emergent themes. Research team members who participated in the interviews and focus groups then worked together, in collaboration with the faculty member, to identify selective quotes that illustrated the emergent themes.

3.4. Overview of participants and implementation findings

During the 2021–2022 pilot year, the Virginia ECMHC pilot received referrals to provide mental health consultation to teachers in 89 preschool classrooms. These referrals were linked to 106 individual

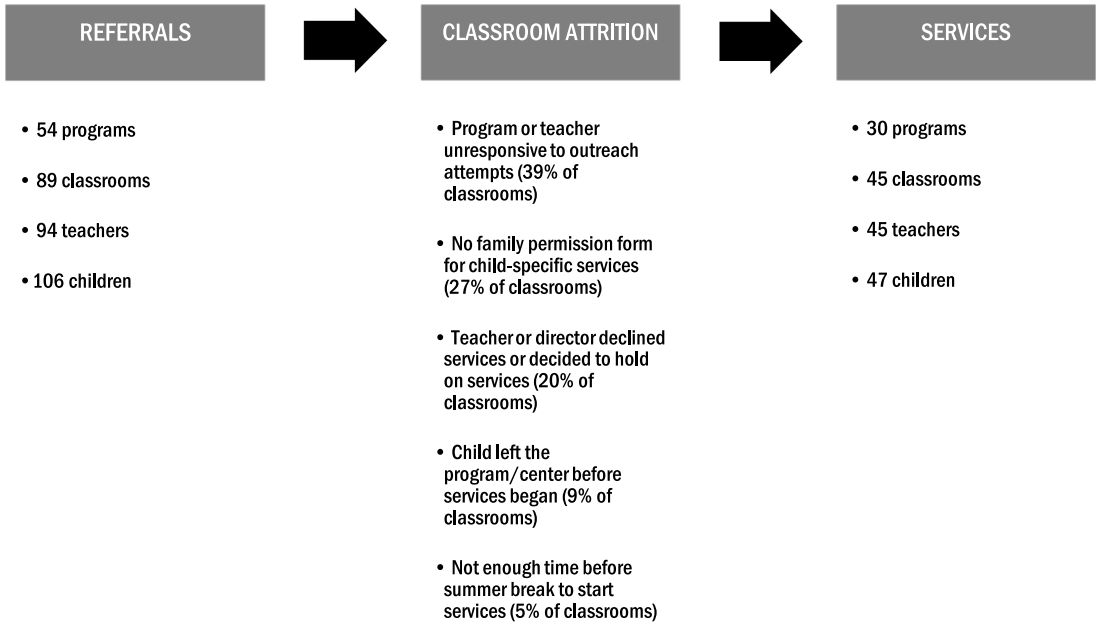


FIGURE 2
Flowchart from ECMHC referrals to services.

children, 94 teachers, 54 ECE programs, and 27 cities or counties across Virginia. Nearly half of the programs (48%) were child care centers.

Figure 2 displays a flowchart of ECMHC participants from the point of referral to services. Teachers in 51% of the 89 referred classrooms ($n = 45$) were served by the ECMHC pilot. Classroom attrition between receiving the referral and initiating services was primarily due to program or teacher unresponsiveness or not receiving family permission to provide child-specific services. Table 1 reports data on ECMHC implementation. Of the 89 referred classrooms, 88.8% requested consultation that supported the teacher to address a specific child's challenging behaviors, while 11.2% of classrooms requested classroom-wide support. Among the 45 classrooms that were served by the ECMHC pilot, 57.8% were assigned to receive the higher intensity services, while 35.6% received lower intensity services. A small number of classrooms did not have a service tier assigned because of incomplete rating scales. All 45 classrooms that were served had teachers who participated in a teacher kick-off meeting. However, only 14 teachers participated in at least one consultation session ($M = 2.14$ sessions, $SD = 0.95$, range = 1–4) and 6 teachers received at least one observation ($M = 1.67$, $SD = 1.63$, range = 1–5). Reasons for the drop-off in services from the kick-off to consultation sessions were complex, and we were not able to capture through systematic quantitative data collection; however, qualitative data described later in section 4 of this paper offer some insight into the low uptake of services.

Table 2 reports information about children who were served by the ECMHC pilot. A total of 47 preschool-age children were served by the ECMHC pilot, and demographic data were available for 32 of them (68%). Among these children, 63% were boys ($n = 20$) and 37% were girls ($n = 12$). The racial/ethnic composition of children was: 38% Black ($n = 12$), 47% White ($n = 15$), 9% Multiracial ($n = 3$), and 6% Latino ($n = 2$). Families reported that the majority of children spoke English at home (97%; $n = 31$) and some spoke Spanish (13%; $n = 4$). On average, children were 4 years old ($M = 4.11$, $SD = 0.64$). A small number of children had an Individualized Education Plan (9%; $n = 3$). Missing child data are due to not receiving a family intake survey or not having permission to report data for research purposes.

Table 3 reports information about the educators who were served by the ECMHC pilot. A total of 45 preschool teachers were served by the ECMHC pilot, and sociodemographic data were available for 22 (49%). Of these, all identified as female. The racial/ethnic composition was: 73% White ($n = 16$), 23% Black ($n = 5$), and 5% Other race/ethnicity ($n = 1$). On average, teachers had 16 years of experiences ($M = 15.77$, $SD = 10.06$), and 81% had at least a Bachelor's degree ($n = 17$). Missing teacher data are due to not receiving a teacher intake survey or not having permission to report data for research purposes.

4. Tensions that arose in context of the Virginia early childhood mental health consultation pilot's implementation

The Virginia ECMHC model launched in a single, large community in the state with the goal of closely monitoring implementation to understand how making this new service available to all publicly funded ECE providers would be received, during a time when COVID-19 disruptions continued to place stress on local ECE programs. This emphasis on feasibility and uptake was of particular importance given growing evidence that social–emotional supports in ECE programs are only helpful when implemented consistently and well (Hemmeter et al., 2022).

In seeking to understand the first year of the Virginia ECMHC pilot, from developing the model and delivering services, we identified critical tensions in implementing social–emotional and mental health services in early childhood settings at scale. The tensions we lay out are ones we grappled with throughout the pilot; our goal is not to provide answers, but to discuss these topics and their implications for ECMHC

TABLE 1 ECMHC pilot implementation.

Requested consultation focus	*N*= 89 referred classrooms	
	n	%
Classroom-wide	10	11.24
Child specific	79	88.76
Service tier	*n* =45 served classrooms	
	n	%
Lower intensity	16	35.56
Higher intensity	26	57.78
Unassigned[a]	3	6.67
Teachers' use of SEL resources before ECMHC pilot	*n* =30 served teachers reporting SEL resources[b]	
	n[c]	%
Coaching	15	50.00
SEL curricula	16	53.33
Webinars/trainings	11	36.67
Online materials	11	36.67
Emotionally exhausted by children's behaviors	*n* =25 served teachers reporting emotional exhaustion	
	n	%
Strongly disagree or disagree	7	28.00
Neither agree nor disagree	7	28.00
Agree or strongly agree	11	44.00

[a]Service tier was unassigned if the PERM and CBRS ratings scales were incomplete. [b]Sample sizes for SEL Resources and Emotional Exhaustion are lower due to missing data (e.g., teacher surveys were incomplete, or teachers did not give permission to report data for research purposes). [c]Teachers could select multiple types of SEL Resources, so the total does not add up to 30.

scale-up efforts (and more generally for SEL programming in ECE). Three key tensions are outlined in the following sub-sections: (1) the tension between ideal and pragmatic roll-out; (2) the tension between supporting teachers' practice as it relates to an individual child and contributing to a deficit view that children need to be "fixed"; and (3) the tension with addressing systemic factors impacting the ECE field that can undermine the implementation and effectiveness of ECMHC. For each tension, we describe relevant literature and present implementation data from the Virginia ECMHC pilot that illustrates how the tension arose in our pilot.

4.1. The tension between ideal and pragmatic roll-out

Tensions related to what is ideal and what is realistic and feasible at scale arose frequently during the implementation of the ECMHC pilot. Our pilot included various stakeholders including educators, families, policymakers, and researchers whom, at times, held different perspectives on what ideal ECMHC implementation looks like. We describe how we grappled with three specific issues around what is ideal versus pragmatic: (1) consultant workforce and qualifications; (2) access to consultation; and (3) systems of support. Decisions related to these tensions were guided by our ongoing partnership with the VDOE and their requests to prioritize consultation for the highest need cases, specifically where a child was at risk of being suspended or expelled.

4.1.1. Consultant workforce and qualifications

> *In some areas of the United States, finding a candidate who fits the standard job requirements of an IECMH consultant is difficult or almost impossible. The reality is that there are mental health professional deserts, and hiring teams may need to be flexible and discerning when hiring. – Center of Excellence Hiring Guidance.*[4]

The Center of Excellence for Infant and Early Mental Health Consultation recommends the following minimum qualifications for consultants: Master's degree in social work, psychology, or related field (preferably licensed); at least 2–3 years of experience working as a mental health professional; possess attributes and skills critical to this work (e.g., facilitates consultative stance, culturally sensitive, and empathetic); have specialized knowledge and deep understanding of early childhood development and social, emotional, and relational health. These highly qualified consultants would deliver promotion, prevention, and intervention services to programs, teachers, and children/families, depending on need.

However, these ideal qualifications bump up against reality that the current ECMHC workforce is small and requiring this amount of training is long and costly. At this time, Virginia does not have a large workforce from which to recruit. Further, using licensed mental health professionals as consultants for the full spectrum of services (from promotion and prevention to intervention) takes an already small workforce away from providing other services, such as direct mental health services to children and families. This unintended consequence may contribute to long wait times for those most in need of intensive, targeted support. In addition, given that ECMHC was being provided within classroom and family day home settings, we deemed prior experience within ECE as critical. We addressed this tension by requiring the following key qualifications for consultants: a Bachelor's degree in psychology, counseling, social work, education, or a related field; a minimum of 2 years of relevant experience, including work in an early childhood environment; and experience with consultation and/or coaching educators. Our qualifications favored certain experiences we deemed critical for consultants who would be working in ECE settings (e.g., experience in early childhood classrooms) and implemented a robust system of initial and ongoing training and support. We did not consider this to be a 'less than' approach in relation to Center for Excellence standards, but rather an intentional adaptation that worked for implementing ECMHC in this state and region at this moment in time.

4.1.2. Access to consultation

> *...I think we stopped utilizing the service because the only thing available to us was virtual…I need someone [an ECMHC consultant] physically to be able to come in and support these teachers. – Joanna*[5]*, Program Director.*

4 https://www.iecmhc.org/resources/hiring-guidance/

5 All names have been changed with permissions from participants to protect their identities.

TABLE 2 Served children's socio-demographic characteristics.

	N= 32			
	n	%	M	SD
Age			4.11	0.64
Male	20	63		
Female	12	37		
Race/ethnicity				
Black	12	38		
White	15	47		
Latino	2	6		
Multiracial	3	9		
Language spoken at home				
English	31	97		
Spanish	4	13		
Child has an IEP	3	9		

Data were missing for 15 children due to incomplete family surveys or because families did not give permission to report data for research purposes.

TABLE 3 Served educators' socio-demographic characteristics.

	N= 22			
	n	%	M	SD
Years of experience			15.77	10.06
Male	0	0		
Female	22	100		
Education				
Some college or 2-year degree	4	19		
Bachelor's degree	9	43		
Master's degree	8	38		
Race/ethnicity				
Black	5	23		
White	16	73		
Other race	1	5		

Data were missing for 23 teachers due to incomplete teacher surveys or because teachers did not give permission to report data for research purposes.

Another tension between the ideal and the real is related to access to consultation. Ideally, all programs and educators would have access to some type of mental health consultation. Consultants embedded within programs (i.e., internal consultants) may provide better services, since they are able to form strong relationships with educators and families, prior to challenges occurring. Embedded consultants have time to get to know the program climate and culture, including strengths and challenges, and can proactively support teachers, rather than reacting once a child is on the brink of expulsion or a teacher is burnt out. However, as described above, funding and workforce challenges result in difficulty providing highly trained consultants to work intensively in every program. This implementation challenge meant that we needed to allocate resources strategically.

Under realistic conditions of limited resources, we used external consultants (rather than internal consultants) and assigned consultation based on level of need in our ECMHC pilot. External consultants typically are only present in the classroom or program when providing services, compared to internal consultants who already work in a program (Giordano et al., 2020). Though external consultants are commonly used, external consultants may not be fully aware of the existing program, community, and cultural contexts, especially at the beginning of consultation, and this can lead to difficulty establishing trust and strong relationships (Giordano et al., 2020). Scheduling consultation activities, such as meetings and observations, is also more difficult for external consultants, compared to consultants who are internally embedded within a program.

We also experienced tensions related to access to different formats for consultation. For our pilot, ECMHC services were offered both in-person and virtually. We used exclusively virtual consultation in some areas to allow consultants to provide services within a wider geographical range. Though we anticipated that virtual consultation would help us reach more programs, our data showed that the teachers and directors preferred services to be delivered in-person as opposed to virtually. Program directors who chose not to take up the ECMHC services told us that one main reason was because only the virtual option was available to them, which they did not see as beneficial. Program directors and teachers shared that teachers were already spending significant amounts of time on screens and another virtual service would have added to "screen fatigue." In the focus groups, teachers who had participated in the ECMHC program shared that they would have liked the consultation and observations to be in-person.

4.1.3. Systems of support

> *I really wish I had mental health services in my school more than just the social worker…who could…step in and get a kid serious services without waiting…There's so many barriers…my two kids that I had last year…never got the services that they needed. – Sarah, Teacher.*

The most effective implementation happens within a system of support (Domitrovich et al., 2008). In the case of early childhood, this system of support also includes coordination across providers of different early childhood services, such as primary care, speech and occupational therapy, mental and behavioral health, and early intervention. Unfortunately, our early childhood system is fragmented, resulting in families navigating separate systems for each necessary support. This theme came up consistently during focus groups with families and teachers who participated in the pilot.

Prior to reaching the level of intensive intervention like ECMHC, systems should ideally be set up to universally promote social and emotional development and prevent challenges in the classroom. These universal supports include access to high quality comprehensive curricula, as well as training and classroom resources to support children's social and emotional development classroom wide. Fifty percent of teachers who were served by the ECMHC pilot reported having had some prior experience being coached on practices that support children's social-emotional learning, 53% of teachers reported using a social-emotional curricula, 37% reported accessing webinars and trainings, and 37% reported using online materials (see Table 1). While it is encouraging that half of teachers had received coaching prior to ECMHC, these universal supports for teachers are still not widely implemented in a coordinated and aligned system. The result can be overly relying on targeted and intensive support once challenges occur. Yet, these targeted and specialized interventions are more costly

and time intensive and contribute to the next tension—perceptions that a child's struggles in the classroom represent a deficit in that child and their family, rather than a product of resources available to them in their environment.

4.2. The tension between supporting teachers' practice as it relates to an individual child and contributing to a deficit view that children need to be "fixed"

> *Yeah, I think, and the other piece of that, of his behavior you know, mom to a Black boy, and he has behaviors that there are negative associations with, particularly for Black boys. When you hear him being described as, "aggressive," or you know, "violent," and stuff like that, when it's, like, that's not how he is. The folks who are working with him, and when they talk about him. The lady who works with him now, like, I almost was brought to tears when she first started talking about her experience with Khalil, because she was talking about how smart he is. You know, how he likes to learn. I mean, just, like, all of the strengths that Khalil possesses, where usually what I hear is, he's hyper, impulsive, you know, those are the things that you first see with Khalil. But after you get to know him, and get to understand him, then you get to see all those things.* ***But the reality is that the world is going to immediately see him as a Black boy who may be acting in a way that's been labeled as aggressive. Particularly, he has now a diagnosis of ADHD****. And where the developmental pediatrician really talked about that from a neurodevelopmental perspective. But even in my line of work, ADHD is not something that... I mean, people kinda laugh at that diagnosis a lot of time. Like, it's an excuse for why kids behave the way that they do, but when you have a better understanding of it, I mean, you almost are like... People have more of an understanding or sympathy or understanding if he, if he had an autism... diagnosis, than the ADHD. But he has so many of the autism features, but that is not the primary diagnosis that he has. You know, I hate to say that, but I... that's just kind of the reality.* ***And I, I want to get him into the school system early so that he's not really known for his behaviors and that they really get a chance to really understand him better because I know how those behaviors are labeled. And they do that for all kids, but the reality is, it is a different experience for little Black boys in particular. Where he can be acting out, and somebody thinks he's being aggressive and feel threatened by him. And so I'm trying to help him with that because I know what that can lead to later on. – Kacia, Mother****.*

The excerpt above is taken from a focus group discussion that we had with mothers of children who had been referred for ECMHC services. Kacia is a Black woman who described her experience of raising a Black boy who seemingly had behavioral challenges and was diagnosed to have Attention-Deficit/Hyperactivity Disorder (ADHD). She shared her worries around labels that get attached to children, especially Black boys. For young Black boys, their behaviors are often tagged with negative associations, such as "violent" or "aggressive." Black parents are aware of teachers' perceptions of their child's behavior and the different experiences of Black youth in the classroom. Like many other parents, Kacia became an advocate for her child, an expert, who sees the many strengths of her child and not how the world views her Black child and automatically labels his externalizing behaviors as "aggressive." It is interesting to note how Kacia described that an autism diagnosis was preferable over a diagnosis of ADHD. Kacia was a social worker and told us that she "knew the system" and even in her "line of work" she had experienced people lacking trust in an ADHD diagnosis.

The goal of the Virginia ECMHC pilot was to support ECE teachers in responding to the social–emotional and mental health needs of the children in their classrooms in the wake of the pandemic. As part of this goal, VDOE was especially interested in reducing and preventing suspensions and expulsions of young children attending ECE programs in Virginia. As a team, we were driven by the belief that behavioral difficulties did not lie with the children but often with the systems around them. Our work was also grounded in the realities of how race plays a role in the way young children and their families are perceived and treated in ECE programs.

In this section, we detail the tension that we experienced as we implemented the ECMHC services and collected and analyzed data. Despite our best intentions to move away from the idea of "fixing children," and focusing on "fixing contexts," we felt that we still centered the child, and possibly cast them in deficit ways. This happened in two ways: (1) Centering deficit views of children by seeking child referrals and (2) Referrals reinforcing systemic ways in which BIPOC children are over identified as having behavioral challenges.

4.2.1. Centering deficit views of children by seeking child referrals

ECMHC services were marketed to address teacher mental health concerns, classroom climate, teacher-child interactions, as well as child-specific behavior concerns. The referral form included a question about the reason for requesting services that was deliberately kept open ended, because we did not want to assign labels to the child or teacher. We wanted the family, teacher, or program leader who was requesting the referral to describe behaviors or challenges in their own words, rather than have them check boxes of pre-determined categories. Because consultation did not have to focus on a specific child, the intake form did not require that respondents list an individual child as the intended focus of consultation. However, nearly all classrooms requested support around addressing a specific child's challenging behavior (see Table 1). Although we deliberately tried to not take a deficit view of the child or their family, as we reflected on our work over the past year, we felt a tension with the mere act of seeking child-specific referrals. Even in the context of an open-ended prompt, the question around the reason for referral provided examples of ways in which a particular child was "behaving" or showing that they had "unmet needs." We ultimately received descriptors in the referrals such as "hitting others," "defying authority," "stealing toys," "disturbing others during naptime," and "refusing to follow classroom routines." As we can see from these examples, these responses described how the child was not fitting well into the context rather than how the context was making it difficult for the child. Further, some teachers that we interviewed—those who utilized the services and those who chose not to—mentioned that they were expecting the ECMHC services to "work with the child directly," instead of working with teachers and families to enhance the contexts that are possibly not working for the child.

Here we are struggling with the tension of wanting to provide services in cases where specific children were having a difficult time in the classroom and the teacher could use support in re-framing and

supporting the child's behavior while also contending with the possibility that the mere act of seeking referrals to implement a consultation program tends to reinforce deficit-based ideas about children and their families. However, to even begin a consultation process rooted in antiracist practices, where we try to help teachers to see the strengths of a child that they see as a problem or to recognize the funds of knowledge (González et al., 2006) that a family has, the starting point has to be the child and their behavior.

4.2.2. Referrals reinforced systemic ways in which BIPOC children are over identified as having behavioral challenges

A second tension we grappled with related to inadvertently contributing to deficit views of children is over identifying BIPOC children as exhibiting behavioral and mental health challenges. Young BIPOC children often attend early education spaces that are more tightly controlled than spaces White children attend (Adair et al., 2018), and their bodies and behaviors are heavily regulated (Hines-Datiri and Carter Andrews, 2020). BIPOC children's behaviors are then mislabeled and misinterpreted as problems, issues, or misbehaviors (Gregory et al., 2017). These labels can lead to disparities in the disciplinary experiences of BIPOC and White children (Gregory et al., 2010; Skiba et al., 2011; Losen and Gillespie, 2012; U.S. Department of Education Office for Civil Rights, 2016; Epstein et al., 2017; Boonstra, 2021).

In Virginia's ECMHC pilot, 41% of children who were referred for services identified as Black and 40% identified as White. Using the American Community Survey 5-year estimates from 2016 to 2020, we calculated an estimate of the racial/ethnic breakdown of children under 17 years for the top five cities and counties from which ECMHC referrals were made. On average, across these localities, the larger population of children was 50% White and 23% Black. As another point of comparison, the statewide sample of preschool children for whom the state had school readiness data was 35% White and 32% Black (Virginia Department of Education and the Center for Advanced Study of Teaching and Learning, University of Virginia, 2022). Comparing our referrals to these estimates, it is likely that Black children were over-identified for ECMHC services; however, we cannot definitively make this conclusion, since we do not have detailed information on race/ethnicity in classrooms served by ECMHC.

Previous studies have suggested that Black families are worried about the stigmas associated with ADHD diagnoses or other behavioral or learning disabilities due to their child being labeled as "crazy" or needing medication to make their child "behave" (Davison and Ford, 2001). Families of BIPOC children are left in limbo, and it is up to the ECE program to accurately accommodate their child and not subject them to the stigmas and stereotypes that come from labels that provide relief to White families as it explains their child's behavior (Davison and Ford, 2001). When Kacia expressed preferring an autism diagnosis over an ADHD diagnosis for her son, she was falling back on her experience and trying to ensure the best possible outcome for her son. Kacia knows that her son, due to him presenting as Black, would be viewed more negatively with an ADHD diagnosis than an autism diagnosis. She was trying to minimize the damage that these labels do by hoping for a diagnosis that minimizes the deficit view of her child.

These deficit views about BIPOC children get extended to their families too. The parents that we interviewed shared that they were working hard to support their children in school and at home. Like Kacia, we heard stories of parents advocating for their children while they confronted their own challenges concerning workload and income, especially in the aftermath of COVID. However, teachers sometimes expressed deficit views of families, saying that parents were "in denial" about their children's needs, were not parenting in the right way, or were unable to comprehend their children's needs. Parents felt they were labeled as uninvolved without consideration for the strides they take to support their children (Devlieghere et al., 2020).

Like families, ECE providers are also working to overcome incredible challenges to ensure the best outcomes for children in their care. What may be perceived by families as a provider pushing a diagnosis on their children might be explained by the provider as efforts to make resources and supports available for the child. Although ECE providers report being both devoted to and rewarded by their work, systemic issues exert real influence on teachers' stress and well-being which can undermine their responsiveness to children and families. Next, we discuss these systemic issues in the ECE field and how they relate to ECMHC implementation.

4.3. The tension with addressing systemic factors impacting the early care and education field that can undermine the implementation and effectiveness of early childhood mental health consultation

> *And I, I know, like as an outsider or a professional, and I feel like you, you know the long-term solutions, what needs to happen, but like just with like children and learning, if their basic needs aren't met, they cannot be there and be present to learn it. I think that's where teachers were at. They were in crisis. It was fight or flight in that stem of their brain. They were not in a place to learn or work towards a long-term solution. – Tammy, Program Director.*

ECMHC focuses on working with adults to better understand and respond to behavior in context, and thus requires the capacity to learn, engage with, and apply new information to make changes in practice. A model of ECMHC that targets the child's environment (versus a deficit view that sees the child as the point of intervention) requires additional responsibilities and time commitments from adults in the child's ecosystem. However, the same challenges that might lead a teacher or program to seek out ECMHC (e.g., educator stress, educator lack of self-efficacy, lack of resources, lack of work-time supports, students/families experiencing mental health challenges) may also act as barriers that interfere with access to and engagement with this model of consultation. This tension illustrates the need among ECE teachers for resources and support, but the foundation for receiving them is not always there, analogous to pouring water into a bucket with a hole in the bottom. Patching this hole (i.e., addressing systemic issues in the ECE field at the macro level) is necessary for supports and services like ECMHC, and other SEL programming in ECE settings, to be effective at scale. As previously noted, the dosage of ECMHC that providers received was below the intended dosage of the model. While we expected teachers to participate in 4–7 consultation cycles, depending on the service tier, the average number of sessions was 2.14 (range 1–4) and the average number of observations was 1.67 (range 1–5). We discuss three systemic issues that impact ECE providers' work and undermined teachers' engagement with and dosage of ECMHC: (1) turnover,

coverage, and workload; (2) provider stress, mental health, and well-being; and (3) compensation for ECE providers. Although these systemic issues were exacerbated by the COVID-19 pandemic, they were already impacting the ECE system; thus, it is important to understand and address these issues well beyond the pandemic when considering implementation of ECMHC and other SEL programs.

4.3.1. Turnover, coverage, and workload

> *Our teachers, as you know, everywhere with teaching, teachers are like burnt out. They feel like there's so much on their plate. And when they got down into this, it seemed like a lot more work on a teacher, and they just were not willing to take it on. It seems like a lot of training and things that they had to do and not the support that they were looking for. So that's why a lot of my teachers were like, "No, I'm not doing that [ECMHC] now." – Tammy, Program Director.*

Focus group and exit interview participants often expressed the sentiment that early childhood educators are at capacity. Turnover and coverage challenges resulted in high workloads and increased stress, which made taking on ECMHC services untenable for teachers in many programs. Teachers and leaders who remained at their programs have dealt with the fallout of high turnover rates during the pandemic (Quinones et al., 2021; Berger et al., 2022). Teachers who remained at their program were caught in a cycle in which turnover and lack of coverage led to a higher workload:

> *Um, and then, yeah, that was a huge part of it, because you have the teachers who do show up, who were dedicated and loyal to come into work every day, that are working overtime to make up for the lack of teachers that we have. – Liz, Program Director.*

Challenges associated with turnover including lack of coverage, overworked staff, inconsistencies in staffing within classrooms, and higher workloads for teachers lead to a lack of time and bandwidth for program staff to complete basic tasks and responsibilities. Early educators already complete many job demands without proper supports, such as paid planning time, adequate staffing, and training (Jeon et al., 2018; Jeon and Wells, 2018; Roberts et al., 2019; Kwon et al., 2022). Within the Virginia ECMHC pilot, programs were experiencing similar challenges that acted as barriers to engagement. Leaders mentioned that they were looking for support without placing additional demands on their teachers, who often need to complete work responsibilities outside of compensated hours. Teachers also declined ECMHC services due to a lack of capacity to engage:

> *I have no planning time [already]. I do not have time to get another resource [ECMHC services]. – Mandy, Teacher.*

Early childhood educators have shown resilience and creativity in their work to support children and families at the same time they have been experiencing high levels of stress, turnover, and financial insecurity (Beltman et al., 2020; Crawford et al., 2021; Eadie et al., 2021; Swigonski et al., 2021). However, the time, resources, and capacity for consistently engaging with consultation (e.g., attending meetings, uploading videos for review, implementing practices and reflecting on them) were not available to all educators and programs. Unfortunately, this likely leads to situations where programs with less resources are in greater need of consultation but have less bandwidth to engage. Additionally, if early educators had more time and resources at their disposal, they may have been more receptive to a model of ECMHC that targets changes to the classroom environment and their practice, rather than a frequently expressed sentiment that many educators wanted someone to pull out or work directly with the child.

4.3.2. Provider stress, mental health, and well-being

> *So I do know the importance of, um, supporting the teachers, not just with interventions and strategies to help children in the classroom, but dealing with, you know, their own baggage that makes, um, when you come in the classroom and you are already burnt out and tired, then it's hard to build a relationship with a student that is causing you more stress. And then that in itself causes teachers to make decisions that may not always be the best decision because you are doing it from frustration or because you are tired or you are already stressed and burnt out. – Laura, Program Director.*

Larger systems contribute to stress and mental health of adults, shaping the care they provide to children as well as their perceptions of and responses to children's behavior (Buettner et al., 2016; Jeon et al., 2019; Zinsser et al., 2019). ECE providers experience higher levels of stress, depressive symptoms, and burnout than the general population (Jeon et al., 2018; Jeon and Wells, 2018; Roberts et al., 2019; Kwon et al., 2022). ECE providers also report increased rates of stress, anxiety, anger, frustration, sleeping problems, and physical pain since the start of the pandemic (Berger et al., 2022; Farewell et al., 2022). Program directors, who play a critical role in supporting teacher engagement with ECMHC, including prioritization of consultation, scheduling coverage for consultation meetings and shaping policies and philosophies around interpreting and responding to behavior, also report high levels of stress that interfere with their ability to focus on educators' well-being and professional development (Kristiansen et al., 2021).

A desired outcome of ECMHC is to enhance teachers' ability to respond effectively to behaviors they perceive as challenging and to meet young children's social and emotional needs, which is one of the most commonly reported stressors for early educators (Hoover et al., 2012; Reyes et al., 2021). Nearly half (44%) of teachers served in the Virginia ECMHC pilot reported that they were emotionally exhausted by children's behaviors (see Table 1). Though the intent is to alleviate a source of stress for educators, ECMHC can add more burden or stress, particularly for educators who might be struggling to manage their own mental health needs. The Virginia ECMHC model worked from a framework of understanding behavior in context, so some educators were learning a new way of understanding and interpreting children's behavior while unlearning implicit biases and child-focused behavioral attributions. This learning and reflection requires personal work that can bring up difficult emotions at a time when teachers already feel targeted:

> *I feel like teachers are being targeted a lot right now and they feel like everyone's telling them how to do their job better- just feeling very criticized about a lot. So I think- if somebody was meeting them where they are and, you know, kind of modeling the, that support- would help them to say like, "Oh, okay. Yeah, that could really help." They just want to be heard about what's going on in their classroom. And that's a tough thing for a teacher. Like those are their four, four walls. That's the one area in the school they can control, so letting somebody*

> *into that is difficult, and especially via video camera. – Tammy, Program Director.*

4.3.3. Compensation for early care and education providers

> *We have staff that's like, in order for us to come to work, we need you to pay our gas. We need for you to give us the gas money because we just cannot afford it. Uh, so we, coming up with how we are gonna help staff get back and forth to work because gas prices went up. Um, I cannot increase your salary, but I can maybe do a gas voucher. Um, but the ins and outs of doing that is the paperwork for it is just ridiculous. And, when you have over almost 100 staff and maybe 75% of them need assistance with traveling to work, that paperwork is intense. So, it's like, do we help? – Whitney, Program Director.*

In addition to the heightened workload and stress that ECE providers are facing in the aftermath of the pandemic, many providers are struggling financially to meet their own and their family's basic needs due to disgracefully low compensation. In Virginia, median pay for ECE providers is $10.96 per hour, and 16.4% of providers live below the poverty line, twice the rate of workers in the state overall (McLean et al., 2021). Working during the pandemic, in many cases without health care or benefits such as paid sick leave, has placed additional financial and emotional stress on providers (Markowitz and Bassok, 2022). Low compensation relates to the other two systemic topics previously discussed. Financial insecurity leads to higher stress levels, turnover, and higher workload among remaining teachers.

Despite ECMHC being implemented at the child/family, classroom, or program level, this work cannot be separated from larger systemic forces that impact day-to-day functioning of educators, children, and families. Understanding and addressing systemic issues in the early childhood field (e.g., turnover, workload, stress, mental health, and compensation) will resolve some of the mental health and social–emotional challenges ECMHC targets, while also creating a stronger infrastructure for providing more effective support and implementing SEL interventions when needed. Therefore, we see part of our role as researchers being connected to social justice and advocating for systemic changes that would improve the lives of children, families, and early educators, while also helping us be effective in our role of developing, studying, and understanding supports. Understanding the larger context and the interconnected factors that shape educator, family, and child experiences is essential in approaching ECMHC, and other SEL programming in ECE, from a strengths-based, ecological systems orientation and promoting sustained change at scale.

5. Conclusion

ECMHC is a targeted prevention SEL service intended to build the capacity of ECE professionals to promote children's social-emotional competence and improve their mental health and well-being. In this paper, we described tensions that illustrate competing needs and values that arose when designing and rolling out Virginia's pilot ECMHC program during its first year of implementation. We categorized our tensions into three areas: (1) ideal versus pragmatic roll-out; (2) the potential to contribute to a deficit model where children are viewed as needing to be "fixed"; and (3) the systemic factors impacting the ECE field that undermine the implementation and effectiveness of ECMHC. For each tension, we presented a combination of qualitative and quantitative data to illustrate how these tensions played out for the children and providers we intended to serve through ECMHC.

We described these major tensions within the context of implementing a particular social–emotional learning program – ECMHC. However, the tensions arising from the practicality of implementing at scale, programming that may contribute to a deficit model of children who need to be "fixed," and systemic barriers that prevent or reduce successful uptake are not unique to our ECMHC pilot and are relevant to other SEL services, curricula, and interventions. For example, coaching and consultation are a core component of many SEL interventions and are used to help ensure high quality implementation of the particular SEL program (Pas et al., 2014). However, the cost of coaching and consultation is often prohibitive at scale, especially at the levels of frequency (i.e., weekly or bi-weekly for a full school year) and universality (i.e., to every educator who is engaging in the SEL intervention) delivered in the context of experimental evaluations. In reality, coaching and consultation will not be delivered at this intensity at scale and researchers might consider testing coaching and consultation delivery that could be practically delivered community-wide (e.g., coaching only teachers who are struggling to implement, during initial uptake only, and/or providing more scalable "nudge" supports). Specific to ECMHC programs, the field would benefit from a better understanding of the specific service components, and the dosage of those components, that lead to positive impacts for educators, families, and children, so that limited resources can be allocated most effectively at scale.

In addition, our work highlights the need for SEL interventionists, especially those operating at tier 2 or 3 in a multi-systemic system of support, to be reflective about whether their service may be inadvertently contributing to the idea that children need to be "fixed." For example, in what ways does pulling a child or group of children out of the classroom for a social skills group communicate to children, their teachers, and their parents that their child needs to be "fixed"? In the context of ECMHC programs specifically, services typically begin with some kind of referral system. Although the referral process is typically not considered to be a program component, we argue that it should be, because systematic investigation of the referral system can lead to helpful insights about how to best create a system that reaches potential participants but does not communicate unintended messages about children in the process. Finally, the systemic barriers our ECE workforce encounters create a system where most any SEL programming may not be able to achieve the intended positive impact. For example, it is near impossible to implement any service, curriculum, or program with good fidelity in partnership with a workforce that does not earn enough money to stay in their profession. Low wages of ECE providers result in constant educator turnover, which makes it very difficult to train and support providers over a longer time span. For example, time and money put into training providers on any SEL program is lost once the teacher has left the profession. Again, in this paper, we do not suggest there is a preferred or correct answer to solve these tensions, but we found that considering these tensions explicitly resulted in modifications that we hope will improve our implementation of ECMHC for children, educators, and families.

This paper also extended an implementation framework that was developed with K-12 school settings in mind to be relevant for the ECE context. We noted unique considerations for implementation within ECE contexts at the individual-, program-, and macro-levels. However, differences at the macro-level are perhaps the most stark when comparing factors that influence implementation of SEL programs in ECE versus K-12 settings. The ECE system is fragmented and

underfunded, has chronic workforce instability, and available resources for both children and educators are highly variable. The impacts of these systemic factors are felt everyday by providers and families across the United States. Without policy changes, wide scale implementation of SEL programs is insufficient to support children and families.

We focused on qualitative and survey data where we asked providers to describe their experience of and satisfaction with the ECMHC model to try to center the experiences of the children, providers, and families that we were intending to serve through this pilot. The voices of providers and families who were willing to share their experiences with us, especially through interviews and focus groups, helped us identify the tensions discussed in this paper. Our examination of the ECMHC pilot through these data sources highlight the value of embedding qualitative methodology when examining the scaling of SEL programming.

We recognize several limitations of this work and areas for future directions. First, we piloted the first year of Virginia's ECMHC model during the world-wide COVID-19 pandemic. As such, some of the tensions described in this paper were almost certainly exacerbated due to the negative repercussions of the pandemic. While we believe that the discussion of the tensions described in this paper will be applicable to the scaling of ECMHC in Virginia and beyond in a post-pandemic context, the fact that our pilot was implemented, and our data collected, during the COVID-19 pandemic limits our ability to compare our findings with prior research and may limit future researchers to compare their findings with our work. Second, the Virginia ECMHC model was designed and implemented birth-to-5, but only preschool data were accessible from this pilot year. In future years, we plan to incorporate data from infant/toddler programs, to better understand implementation of the entire Virginia ECMHC model. Additionally, data from a larger sample of participants will allow us to further examine the implementation tensions outlined in this paper. For example, it would be interesting to explore whether some tensions are more salient for certain ECMHC components versus others (e.g., the service intensity or consultation focus). Third, compared to the overall number of referrals that were received, a small percentage of providers engaged in ECMHC services. We also have much missing quantitative data due to challenges with reaching providers, which limits our understanding of who made a referral but did not progress through services and why. In future years, we will continue to center the voices of providers and families to understand their experiences engaging in the Virginia ECMHC program. Our goal is to continually apply lessons learned to improve our model, work to alleviate systemic barriers faced by providers and families when possible, and develop enhanced supports that will lead to better ECMHC implementation and social–emotional outcomes for young children in Virginia.

Data availability statement

The datasets presented in this article are not readily available because the data are restricted due to confidentiality agreements. Requests to access the datasets should be directed to corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by University of Virginia Institutional Review Board. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

AP wrote the introduction, implementation of ECMHC in Virginia, and conclusion, edited all other sections of the paper, and led the quantitative data collection and analysis. SS led the qualitative data collection and analysis and the second tension on contributing to deficit views of children. MB led the third tension on systemic issues. KC led the conceptual framework and first tension on ideal versus pragmatic roll-out. SM-M and KP contributed to the second tension on contributing to deficit views of children. PA contributed to analysis of implementation data, and CF contributed to qualitative data collection and identified quotes that best represented themes. AW and JD reviewed, edited, and provided feedback on the manuscript and contributed to the conclusion. All authors contributed to formulating the implementation tensions described in the paper.

Funding

This research was supported by the Virginia Department of Education with funding provided through a federal Governor's Emergency Education Relief (GEER) Fund. The opinions expressed in this paper are those of the authors and do not represent the views of the funding agencies.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Adair, J. K., Colegrove, K. S., and McManus, M. (2018). Troubling messages: agency and learning in the early school experiences of children of Latinx immigrants. *Teach. Coll. Rec.* 120, 1–40. doi: 10.1177/0161468118120006

Adair, J. K., and Kurban, F. (2019). Video-cued ethnography in educational settings. *Anthropol. Educ. Q.* 50, 245–254. doi: 10.1111/aeq.12311

Albritton, K., Mathews, R. E., and Anhalt, K. (2019). Systematic review of early childhood mental health consultation: implications for improving preschool discipline disproportionality. *J. Educ. Psychol. Consult.* 29, 444–472. doi: 10.1080/10474412.2018.1541413

Bandura, A. (1997). *Self-efficacy: The exercise of control* W.H. Freeman and Company.

Bartlett, J. D., and Smith, S. (2019). The role of early care and education in addressing early childhood trauma. *Am. J. Community Psychol.* 64, 359–372. doi: 10.1002/ajcp.12380

Beltman, S., Dobson, M. R., Mansfield, C. F., and Jay, J. (2020). "The thing that keeps me going": educator resilience in early learning settings. *Int. J. Early Years Educ.* 28, 303–318. doi: 10.1080/09669760.2019.1605885

Berger, E., Quinones, G., Barnes, M., and Reupert, A. (2022). Early childhood educators' psychological distress and wellbeing during the COVID-19 pandemic. *Early Child. Res. Q.* 60, 298–306. doi: 10.1016/j.ecresq.2022.03.005

Boonstra, K. E. (2021). Constructing "behavior problems": race, disability, and everyday discipline practices in the figured world of kindergarten. *Anthropol. Educ. Q.* 52, 373–390. doi: 10.1111/aeq.12374

Bradshaw, C. P., Mitchell, M. M., O'brennan, L. M., and Leaf, P. J. (2010). Multilevel exploration of factors contributing to the overrepresentation of black students in office disciplinary referrals. *J. Educ. Psychol.* 102, 508–520. doi: 10.1037/a0018450

Bronson, M. B., Goodson, B. D., Layzer, J. I., and Love, J. M. (1990). *Child behavior rating scale* Abt Associates.

Buettner, C. K., Jeon, L., Hur, E., and Garcia, R. E. (2016). Teachers' social–emotional capacity: factors associated with teachers' responsiveness and professional commitment. *Early Educ. Dev.* 27, 1018–1039. doi: 10.1080/10409289.2016.1168227

Center for Disease Control and Prevention. (2022). *Risk for COVID-19 infection, hospitalization, and death by race/ethnicity*. Available at: https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-race-ethnicity.html

Center on Budget and Policy Priorities (2020). *Tracking the COVID-19 recession's effects on food, housing, and employment hardships*. Available at: https://www.cbpp.org/sites/default/files/atoms/files/8-13-20pov.pdf

Clayback, K., and Williford, A. P. (2022). Teacher and classroom predictors of preschool teacher stress. *Early Educ. Dev.* 33, 1347–1363. doi: 10.1080/10409289.2021.1972902

Cohen, E., and Kaufmann, R. (2005). *Early childhood mental health consultation. DHHS pub. No. CMHS-SVP0151* Center for Mental Health Services, Substance Abuse and Mental Health Services Administration.

Conners Edge, N. A., Kyzer, A., Abney, A., Freshwater, A., Sutton, M., and Whitman, K. (2021). Evaluation of a statewide initiative to reduce expulsion of young children. *Infant Ment. Health J.* 42, 124–139. doi: 10.1002/imhj.21894

Cook, C. R., Lyon, A. R., Kubergovic, D., Browning Wright, D., and Zhang, Y. (2015). A supportive beliefs intervention to facilitate the implementation of evidence-based practices within a multi-tiered system of supports. *Sch. Ment. Heal.* 7, 49–60. doi: 10.1007/s12310-014-9139-3

Crawford, A., Vaughn, K. A., Guttentag, C. L., Varghese, C., Oh, Y., and Zucker, T. A. (2021). "Doing what I can, but I got no magic wand": A snapshot of early childhood educator experiences and efforts to ensure quality during the COVID-19 pandemic. *Early Childhood Educ. J.* 49, 829–840. doi: 10.1007/s10643-021-01215-z

Creswell, J. W., and Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches* Sage.

Davis, A. E., Perry, D. F., and Rabinovitz, L. (2020). Expulsion prevention: framework for the role of infant and early childhood mental health consultation in addressing implicit biases. *Infant Ment. Health J.* 41, 327–339. doi: 10.1002/imhj.21847

Davison, J. C., and Ford, D. Y. (2001). Perceptions of attention deficit hyperactivity disorder in one African American community. *J. Negro Educ.* 70, 264–274. doi: 10.2307/3211279

Devlieghere, J., Li, Y., and Vandenbroeck, M. (2020). Beyond the veil of parents: deconstructing the concept of parental involvement in early childhood education and care. *Early Years* 42, 587–598. doi: 10.1080/09575146.2020.1840526

Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., et al. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: a conceptual framework. *Adv. School Ment. Health Promot.* 1, 6–28. doi: 10.1080/1754730x.2008.9715730

Domitrovich, C. E., Gest, S. D., Gill, S., Jones, D., and DeRousie, R. S. (2009). Individual factors associated with professional development training outcomes of the head start REDI program. *Early Educ. Dev.* 20, 402–430. doi: 10.1080/10409280802680854

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

Doromal, J. B., Bassok, D., Bellows, L., and Markowitz, A. J. (2022). Hard-to-staff centers: exploring center-level variation in the persistence of child care teacher turnover. *Early Child. Res. Q.* 61, 170–178. doi: 10.1016/j.ecresq.2022.07.007

Downer, J. T., Booren, L. M., Lima, O. K., Luckner, A. E., and Pianta, R. C. (2010). The individualized classroom assessment scoring system (inCLASS): preliminary reliability and validity of a system for observing preschoolers' competence in classroom interactions. *Early Child. Res. Q.* 25, 1–16. doi: 10.1016/j.ecresq.2009.08.004

Downer, J. T., Williford, A. P., Bulotsky-Shearer, R. J., Vitiello, V. E., Bouza, J., Reilly, S., et al. (2018). Using data-driven, video-based early childhood consultation with teachers to reduce children's challenging behaviors and improve engagement in preschool classrooms. *Sch. Ment. Heal.* 10, 226–242. doi: 10.1007/s12310-017-9237-0

Durlak, J. A., and DuPre, E. P. (2008). Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am. J. Community Psychol.* 41, 327–350. doi: 10.1007/s10464-008-9165-0

Eadie, P., Levickis, P., Murray, L., Page, J., Elek, C., and Church, A. (2021). Early childhood educators' wellbeing during the COVID-19 pandemic. *Early Childhood Educ. J.* 49, 903–913. doi: 10.1007/s10643-021-01203-3

Epstein, R., Blake, J., and González, T. (2017). *Girlhood interrupted: the erasure of black girls' childhood* Georgetown Law Center on Poverty and Inequality doi: 10.2139/ssrn.3000695.

Farewell, C. V., Quinlan, J., Melnick, E., Power, J., and Puma, J. (2022). Job demands and resources experienced by the early childhood workforce serving high-needs populations. *Early Childhood Educ. J.* 50, 197–206. doi: 10.1007/s10643-020-01143-4

Fettig, A., and Artman-Meeker, K. (2016). Group coaching on pre-school teachers' implementation of pyramid model strategies. *Top. Early Child. Spec. Educ.* 36, 147–158. doi: 10.1177/0271121416650049

Friedman-Krauss, A. H., Raver, C. C., Morris, P. A., and Jones, S. M. (2014). The role of classroom-level child behavior problems in predicting preschool teacher stress and classroom emotional climate. *Early Educ. Dev.* 25, 530–552. doi: 10.1080/10409289.2013.817030

Garro, A., Giordano, K., Gubi, A., and Shortway., (2021). A consultation approach to target exclusionary discipline of students of color in early childhood education. *Contemp. Sch. Psychol.* 25, 124–135. doi: 10.1007/s40688-019-00258-9

Gilliam, W. S., and Reyes, C. R. (2018). Teacher decision factors that lead to preschool expulsion: scale development and preliminary validation of the preschool expulsion risk measure. *Infants Young Child.* 31, 93–108. doi: 10.1097/IYC.0000000000000113

Giordano, K., Eastin, S., Calcagno, B., Wilhelm, S., and Gil, A. (2020). Examining the effects of internal versus external coaching on preschool teachers' implementation of a framework of evidence-based social-emotional practices. *J. Early Child. Teach. Educ.* 42, 423–436. doi: 10.1080/10901027.2020.1782545

Glaser, B. G. (2016). Open coding descriptions. *Ground. Theor. Rev.* 15, 108–110.

Gomez, R. E., Kagan, S. L., and Fox, E. A. (2015). Professional development of the early childhood education teaching workforce in the United States: an overview. *Prof. Dev. Educ.* 41, 169–186. doi: 10.1080/19415257.2014.986820

González, N., Moll, L. C., and Amanti, C. (2006). *Funds of knowledge: Theorizing practices in households, communities, and classrooms* Routledge doi: 10.4324/9781410613462.

Goodman, R. (1997). The strengths and difficulties questionnaire: a research note. *J. Child Psychol. Psychiatry* 38, 581–586. doi: 10.1111/j.1469-7610.1997.tb01545.x

Gregory, A., Skiba, R. J., and Mediratta, K. (2017). Eliminating disparities in school discipline: a framework for intervention. *Rev. Res. Educ.* 41, 253–278. doi: 10.3102/0091732X17690499

Gregory, A., Skiba, R. J., and Noguera, P. A. (2010). The achievement gap and the discipline gap: two sides of the same coin? *Educ. Res.* 39, 59–68. doi: 10.3102/0013189X09357621

Hemmeter, M. L., Barton, E., Fox, L., Vatland, C., Henry, G., Phan, L., et al. (2022). Program-wide implementation of the pyramid model: supporting fidelity at the program and classroom levels. *Early Child. Res. Q.* 59, 56–73. doi: 10.1016/j.ecresq.2021.10.003

Hemmeter, M. L., Fox, L., and Snyder, P. (2014). *Teaching pyramid observation tool for preschool classrooms: Research edition*. Brookes.

Hemmeter, M. L., Ostrosky, M., and Fox, L. (2006). Social and emotional foundations for early learning: a conceptual model for intervention. *Sch. Psychol. Rev.* 35, 583–601. doi: 10.1080/02796015.2006.12087963

Hindman, A. H., and Bustamante, A. S. (2019). Understanding well-being among teachers in early childhood settings: challenges, supports, and implications for children's development. *J. Appl. Dev. Psychol.* 61, 1–3. doi: 10.1016/j.appdev.2019.03.005

Hines-Datiri, D., and Carter Andrews, D. J. (2020). The effects of zero tolerance policies on black girls: using critical race feminism and figured worlds to examine school discipline. *Urban Educ.* 55, 1419–1440. doi: 10.1177/0042085917690204

Hoover, S. D., Kubicek, L. F., Rosenberg, C. R., Zundel, C., and Rosenberg, S. A. (2012). Influence of behavioral concerns and early childhood expulsions on the development of early childhood mental health consultation in Colorado. *Infant Ment. Health J.* 33, 246–255. doi: 10.1002/imhj.21334

Hunter, A., Davis, A. E., Perry, D. F., and Jones, W. (2016). *The Georgetown manual for school-based early childhood mental health consultation services* Georgetown University Center for Child and Human Development https://www.iecmhc.org/documents/FCC_Manual_2020_Proof2.pdf.

Jeon, L., Buettner, C. K., and Grant, A. A. (2018). Early childhood teachers' psychological well-being: exploring potential predictors of depression, stress, and emotional exhaustion. *Early Educ. Dev.* 29, 53–69. doi: 10.1080/10409289.2017.1341806

Jeon, H.-J., Kwon, K.-A., Walsh, B., Burnham, M. M., and Choi, Y.-J. (2019). Relations of early childhood education teachers' depressive symptoms, job-related stress, and professional motivation to beliefs about children and teaching practices. *Early Educ. Dev.* 30, 131–144. doi: 10.1080/10409289.2018.1539822

Jeon, L., and Wells, M. B. (2018). An organizational-level analysis of early childhood teachers' job attitudes: workplace satisfaction affects early head start and head start teacher turnover. *Child Youth Care Forum* 47, 563–581. doi: 10.1007/s10566-018-9444-3

Kristiansen, E., Tholin, K. R., and B, M., (2021). Early childhood Centre directors coping with stress: firefighters and oracles. *Int. J. Educ. Manag.* 35, 909–921. doi: 10.1108/IJEM-12-2020-0584

Kwon, K.-A., Jeon, S., Castle, S., and Ford, T. G. (2022). Children's behavioral challenges in head start classrooms: links to teacher well-being and intent to leave. *Early Childhood Educ. J.* 50, 1221–1232. doi: 10.1007/s10643-021-01253-7

Losen, D. J., and Gillespie, J. (2012). *Opportunities suspended: The disparate impact of disciplinary exclusion from school* Civil Rights Project/Proyecto Derechos Civiles https://civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/federal-reports/upcoming-ccrr-research/losen-gillespie-opportunity-suspended-2012.pdf.

Markowitz, A. J., and Bassok, D. (2022). Understanding the wellbeing of early educators in the wake of the coronavirus pandemic: lessons from Louisiana. *Early Child. Res. Q.* 61, 1–14. doi: 10.1016/j.ecresq.2022.05.001

McClelland, M. M., Tominey, S. L., Schmitt, S. A., and Duncan, R. (2017). Social–emotional learning interventions in early childhood. *Future Child.* 27, 33–47. doi: 10.1353/foc.2017.0002

McGoey, K. E., Rispoli, K. M., Venesky, L. G., Schaffner, K. F., McGuirk, L., and Marshall, S. (2014). A preliminary investigation into teacher perceptions of the barriers to behavior intervention implementation. *J. Appl. Sch. Psychol.* 30, 375–390. doi: 10.1080/15377903.2014.950441

McLean, C., Austin, L. J. E., Whitebook, M., and Olson, K. L. (2021). *Early childhood workforce index–2020*. Berkeley: Center for the Study of Child Care Employment, University of California, Berkeley. Available at: https://cscce.berkeley.edu/workforce-index-2020/wp-content/uploads/sites/3/2021/02/Early-Childhood-Workforce-Index-2020.pdf

Miles, M. B., Huberman, A. M., and Saldaña, J. (2018). *Qualitative data analysis: A methods sourcebook* Sage.

National Scientific Council on the Developing Child (2012). *Establishing a level foundation for life: Mental health begins in early childhood: Working paper 6* Center on the Developing Child, Harvard University https://harvardcenter.wpenginepowered.com/wp-content/uploads/2008/05/Establishing-a-Level-Foundation-for-Life-Mental-Health-Begins-in-Early-Childhood.pdf.

Nemer, S. L., Sutherland, K. S., Chow, J. C., and Kunemund, R. L. (2019). A systematic literature review identifying dimensions of teacher attributions for challenging student behavior. *Educ. Treat. Child.* 42, 557–578. doi: 10.1353/etc.2019.0026

Pas, E., Bradshaw, C., and Cash, A. (2014). "Coaching classroom-based preventive interventions" in *Handbook of school mental health*. eds. M. Weist, N. Lever, C. Bradshaw and J. Owens. *2nd* ed (Springer), 255–267. doi: 10.1007/978-1-4614-7624-5_19

Quinones, G., Barnes, M., and Berger, E. (2021). Early childhood educators' solidarity and struggles for recognition. *Australas. J. Early Childhood* 46, 296–308. doi: 10.1177/18369391211050165

Reyes, C. R., Gilliam, W. S., Cicchetti, D., Luthar, S. S., and Burack, J. A. (2021). Addressing challenging behaviors in challenging environments: findings from Ohio's early childhood mental health consultation system. *Dev. Psychopathol.* 33, 634–646. doi: 10.1017/S0954579420001790

Roberts, A. M., Gallagher, K. C., Daro, A. M., Iruka, I. U., and Sarver, S. L. (2019). Workforce well-being: personal and workplace contributions to early educators' depression across settings. *J. Appl. Dev. Psychol.* 61, 4–12. doi: 10.1016/j.appdev.2017.09.007

Schaack, D. D., Donovan, C. V., Adejumo, T., and Ortega, M. (2022). To stay or to leave: factors shaping early childhood teachers' turnover and retention decisions. *J. Res. Child. Educ.* 36, 327–345. doi: 10.1080/02568543.2021.1955779

Schaack, D. D., Le, V. N., and Stedron, J. (2020). When fulfillment is not enough: early childhood teacher occupational burnout and turnover intentions from a job demands and resources perspective. *Early Educ. Dev.* 31, 1011–1030. doi: 10.1080/10409289.2020.1791648

Scott, C., and Medaugh, M. (2017). Axial coding. *Int. Encycl. Comm. Res. Methods* 10:12. doi: 10.1002/9781118901731.iecrm0012

Shonkoff, J. P., Slopen, N., and Williams, D. R. (2021). Early childhood adversity, toxic stress, and the impacts of racism on the foundations of health. *Annu. Rev. Public Health* 42, 115–134. doi: 10.1146/annurev-publhealth-090419-101940

Skiba, R. J., Horner, R. H., Chung, C. G., Rausch, M. K., May, S. L., and Tobin, T. (2011). Race is not neutral: a national investigation of African American and Latino disproportionality in school discipline. *Sch. Psychol. Rev.* 40, 85–107. doi: 10.1080/02796015.2011.12087730

Swigonski, N. L., James, B., Wynns, W., and Casavan, K. (2021). Physical, mental, and financial stress impacts of COVID-19 on early childhood educators. *Early Childhood Educ. J.* 49, 799–806. doi: 10.1007/s10643-021-01223-z

Tobin, J. (2019). The origins of the video-cued multivocal ethnographic method. *Anthropol. Educ. Q.* 50, 255–269. doi: 10.1111/aeq.12302

Trigg, A. B., and Keyes, A. W. (2019). "Child care and early education as contexts for infant mental health" in *Handbook of infant mental health*. ed. C. H. Zeanah. *4th* ed (Guilford Press), 599–609.

U.S. Department of Education Office for Civil Rights. (2016). *2013–2014 civil rights data collection: A first look*. United States: Department of Education Office for Civil Rights. Available at: https://ocrdata.ed.gov/assets/downloads/2013-14-first-look.pdf

U.S. Department of Health and Human Services and Education. (2014). *Policy statement on expulsion and suspension policies in early childhood settings* Available at: https://www2.ed.gov/policy/gen/guid/school-discipline/policy-statement-ece-expulsions-suspensions.pdf

Virginia Department of Education and the Center for Advanced Study of Teaching and Learning, University of Virginia (2022). *VKRP status report for the chairmen of house appropriations and senate finance committees*. Report prepared for the Virginia Chairmen of House Appropriations and Senate Finance Committees.

Whitebook, M., Phillips, D., and Howes, C. (2014). *Worthy work, STILL unlivable wages: The early childhood workforce 25 years after the National Child Care Staffing Study*. Center for the Study of Child Care Employment, University of California, Berkeley, Berkeley. Available at: https://cscce.berkeley.edu/worthy-work-still-unlivable-wages/

Williams, M., and Moser, T. (2019). The art of coding and thematic exploration in qualitative research. *Int. Manag. Rev.* 15, 45–55.

Zeanah, C. H., and Zeanah, P. D. (2019). "Infant mental health: the clinical science of early experience" in *Handbook of infant mental health*. ed. C. H. Zeanah. *4th* ed (Guilford Press), 5–24.

Zero to Three. (2017). *The basics of infant and early childhood mental health: a briefing paper*. Available at: https://www.zerotothree.org/wp-content/uploads/2017/08/The-Basics-of-Infant-and-Early-Childhood-Mental-Health_-A-Briefing-Paper.pdf

Zinsser, K. M., Zulauf, C. A., Das, V. N., and Silver, H. C. (2019). Utilizing social-emotional learning supports to address teacher stress and preschool expulsion. *J. Appl. Dev. Psychol.* 61, 33–42. doi: 10.1016/j.appdev.2017.11.006

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 24 February 2023
DOI 10.3389/feduc.2023.1100667

OPEN ACCESS

EDITED BY
Celene Domitrovich,
Georgetown University, United States

REVIEWED BY
Donna San Antonio,
Lesley University, United States
Janet Welsh,
The Pennsylvania State University (PSU), United States

*CORRESPONDENCE
Yanhua Shen
✉ yzt4rv@uvawise.edu

SPECIALTY SECTION
This article was submitted to Educational Psychology, a section of the journal Frontiers in Education

RECEIVED 17 November 2022
ACCEPTED 06 February 2023
PUBLISHED 24 February 2023

CITATION
Dyson B, Shen Y, Howley D and Baek S (2023) Social emotional learning matters: Interpreting educators' perspectives at a high-needs rural elementary school.
Front. Educ. 8:1100667.
doi: 10.3389/feduc.2023.1100667

Social emotional learning matters: Interpreting educators' perspectives at a high-needs rural elementary school

Ben Dyson[1], Yanhua Shen[2]*, Donal Howley[3] and Seunghyun Baek[1]

[1]Department of Kinesiology, The University of North Carolina at Greensboro, Greensboro, NC, United States, [2]Department of Education, The University of Virginia's College at Wise, Wise, VA, United States, [3]Department of Kinesiology, Towson University, Towson, MD, United States

Introduction: School educators' have a great influence on the adoption, sustainability, and development of school-based Social and Emotional Learning (SEL) programs. The purpose of this school-based research was to investigate educators' experiences and perspectives on implementing SEL in a high-needs rural elementary school setting.

Methods: Fifteen school educators (n = 15), including ten K-5 classroom teachers, one special education teacher, one social worker, and three school leaders, participated in this study. In addition, lessons were observed, and 17 sets of field notes were taken during 17 different days of visit (60–90 min) over the two semesters. A case study design drawing on qualitative research methods was utilized.

Results: The inductive analysis and constant comparison of the collected data generated six themes: *prerequisite for academic success, essential skills for everyday life, lack of time, lack of preparation and development, home-school disconnection,* and *pushback from students*.

Discussion: The study provided qualitative evidence to support the need for quality SEL implementation and revealed nested levels of constraints for school educators' implementing SEL from the "voices" of school educators. The study also calls for collaborative efforts and shared strategies to facilitate "legitimate" long-term partnerships between universities and schools, families, and communities, particularly in rural areas, in promoting a more holistic vision of the social and emotional development of our children.

KEYWORDS

social and emotional skills, life skills, high-needs, elementary school, school educators, buy-in, constraints

Introduction

As key stakeholders in the education process, teachers and school leaders have a great influence on the adoption, sustainability, and development of school-based Social and Emotional Learning (SEL) programs. SEL is defined as "the process through which individuals learn and apply a set of social, emotional, behavioral, and character skills required to succeed in schooling, the workplace, relationships, and citizenship"

(Jones et al., 2019, p. 19). Schools have a fundamental goal of teaching children to learn core curriculum subjects, such as mathematics and science. In addition to these fundamental academic skills, education studies, politics, and experience have become more aware that social and emotional competencies influence learning (Elias et al., 1997; Jones et al., 2013; Osher et al., 2016). The need to develop future citizens is now prioritized through the combination of academic and social-emotional development (Oberle et al., 2014). Typically, education policy has two goals that have been proposed for schools: (a) improve academic success; (b) strengthen social and emotional competencies (Ellis, 2003). Previous studies have suggested several potential SEL pedagogies to decrease behavioral issues, improve psychological well-being, and enhance academic achievement among students at elementary and middle school levels (Gordon et al., 2016; Hulvershorn and Mulholland, 2018; Dyson et al., 2021).

Teachers have been recognized as the most critical figures implementing school-based SEL programs (Humphrey et al., 2018; Dyson et al., 2019). However, previous research indicated that in-service teacher training in SEL was not adequately provided to address the needs of students (Walker, 2020). In addition, other studies argued that pre-service teachers should receive more training in SEL in their teacher education programs to be better prepared to deliver SEL-based programs (Fleming and Bay, 2004; Katz et al., 2020). Given that they are the ones who deliver SEL pedagogical practices to students in school, their perspectives toward SEL need to be investigated to promote the effectiveness of SEL-based programs (Avramidis and Norwich, 2002). In a longitudinal school-based SEL program, Humphrey et al. (2018) found a higher level of teachers' buy-in and enthusiasm for SEL would lead to a better quality of program implementation and more comprehensive student learning outcomes. However, when teachers' buy-in in SEL is low, the quality of the SEL-related practices and the students' SEL learning outcomes can be impeded (Ee and Cheng, 2013). Subsequently, teachers' buy-in and enthusiasm influenced their curricular and instructional decision-making toward SEL in classrooms (Ennis and Chen, 1995).

A successful school-based SEL program relies on joint efforts between school leaders and teachers. When school leaders' interests and supports toward SEL were weak, teacher's professional development in SEL and the program implementation at schools would highly unlikely to be encouraged and supported, which resulted in SEL programs being "insufficiently coordinated, monitored, evaluated, and improved" (Greenberg et al., 2003). However, when school leaders' interests and supports toward SEL were robust, sustainable support for teachers' professional development in SEL and the schoolwide implementation of SEL programs can be prioritized (Durlak and DuPre, 2008; Evans, 2014).

Research has studied constraints that educators encountered during the implementation of SEL programs. By investigating early childhood teachers' perspectives on SEL in urban classrooms, Humphrey et al. (2018) found limited time, lack of support, and insufficient resources as barriers for teachers to SEL implementation. Avramidis and Norwich (2002) identified several variables that influence teachers' attitudes toward SEL, including child-related variables, teacher-related variables, and educational environmental-related variables. Turnbull (2002) proposed several key factors influencing teachers' response to schoolwide reform programs (e.g., schoolwide SEL programs), including teachers' in-service training, school leaders' support, support from program developers and staff members, and control over classroom implementation. Jones and Cater (2020) pointed out that the lack of a clear understanding of SEL among school leaders is a significant constraint that impedes schoolwide SEL programs and reduces teachers' willingness to use SEL practices in classrooms. However, studies have shown that when those constraints are minimal, the implementation quality of SEL programs and students' SEL learning outcomes can be significantly improved (Zins and Elias, 2007; Collie et al., 2012; Anyon et al., 2016).

To date, only a few research have been conducted qualitatively to explore educators' perspectives on SEL programming, especially with teachers, school leaders, and other school staff in high-needs rural elementary school settings (Dyson et al., 2019; Jones and Cater, 2020). A qualitative study with seven teachers in a high-needs rural elementary school indicated that a new curriculum grounded in SEL empowered students to be responsible citizens by discussing issues over race, immigration, and gender discrimination (San Antonio, 2018). Another study, which was grounded in the promotion of a county-wide SEL program in rural schools, highlighted the importance of the interdisciplinary professionals' collaboration (e.g., teachers, social workers, school psychologists, university research experts, and parents) and the recursive organizational consultation process involving perspectives from multiple stakeholders (Meyers et al., 2015). Despite those encouraging findings, more research about the development and promotion of SEL programs in high-needs rural schools is needed.

Elementary schools are recognized as high-needs if there is a high percentage (>60%) of students from families with incomes below the poverty line, or a high teacher turnover (Higher Education Act, 2020). Students in high-needs schools often experienced "unhelpful schooling relationships and deleterious learning outcomes" (Palacios and Lemberger-Truelove, 2019), demanding a higher level of social and emotional support from school educators. Understanding school educators' perspectives of SEL in a high-needs elementary context can help us better conceptualize SEL and facilitate SEL programming for students (Humphrey et al., 2018). With this in mind, the purpose of this study was to explore educators' buy-in and constraints toward SEL in a high-needs elementary school setting, understanding what worked and what needs to improve. Recommendations that facilitate programs and practices grounded in SEL were discussed based on the findings.

Research on social and emotional learning

Social and Emotional Learning has been developed as a conceptual framework to promote children's cognitive, emotional, and academic competencies (Corcoran et al., 2018). SEL is a comprehensive concept and many researchers tried to define SEL in different ways with different terminologies. Therefore, a multi-level conceptualization of SEL is established to present the popular definitions of SEL and its nature of multiple layers (see **Figure 1**). At a macro-level, SEL can be defined as the process of acquiring "the

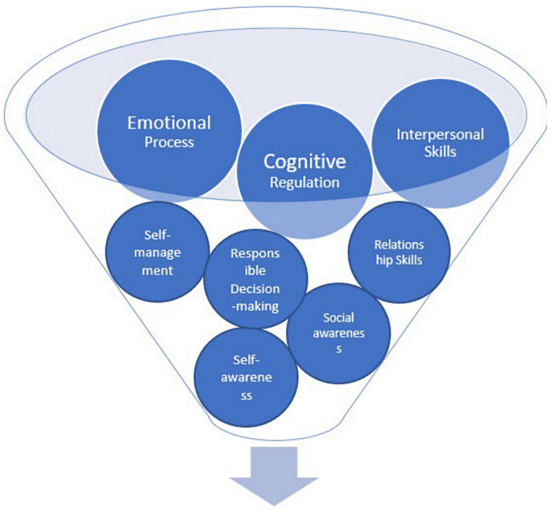


FIGURE 1
A multi-level conceptualization of social and emotional learning (SEL).

ability to understand, manage, and express the social and emotional aspects of one's life" (Elias et al., 1997, p. 2). At a meso-level, SEL includes competencies of cognitive regulation, emotional processes, and social/interpersonal skills (Jones and Bouffard, 2012). SEL can be further recognized at a micro-level as five interrelated skills: self-awareness, self-management, social awareness, relationship skills, and responsible decision-making (Collaborative for Academic, Social, and Emotional Learning [CASEL], 2023). Our research is based on the three SEL definitions and the study was grounded in the school context and the voice of the educators participating in the research.

An emerging body of literature has shown that a wide range of SEL competencies is connected with students' later success in multiple contexts, including school and workplace (Durlak et al., 2010; Jones and Doolittle, 2017). Some of the meta-analyses within the last ten years have added evidence to the benefits of SEL for students, advancing the case for further implementation within schools. Durlak et al.'s (2011) meta-analysis of 213 school-based SEL programs observed significant positive effects, including increased social-emotional competencies, enhanced behavioral adjustments, reduced mental stress, and improved academic performance. Sklad et al.'s (2012) meta-analysis also suggested that SEL programs significantly reduced antisocial behavior and substance abuse. By examining the core SEL competencies (e.g., attitudes toward self, pro-social behavior, conduct problems, emotional distress, etc.), Wigelsworth et al. (2016) confirmed that SEL programs effectively achieved the intended outcomes in these areas. Taylor et al.'s (2017) meta-analysis of 75 reports from 69 SEL programs found significantly improved academic and school performance. More recent findings from Corcoran et al.'s (2018) meta-analysis of academic achievement-oriented school SEL programs indicated that in comparison to traditionally teacher-centered classes, students in SEL-based classes achieved more significant improvement in reading and mathematics.

Despite the considerable amount of quantitative evidence in SEL (Durlak et al., 2011; Sklad et al., 2012; Taylor et al., 2017; Corcoran et al., 2018), there exists a need to investigate school educators' voices regarding their perspectives on SEL in school-based settings (Dyson et al., 2019). This study aimed to address this shortfall within a high-needs elementary school context (Elias and Haynes, 2008; Blair and Raver, 2015), particularly by examining school educators' voices through the theoretical lenses of human developmental perspectives (Jones et al., 2019).

Theoretical perspectives

Research on how interactions between individuals and contexts influence social and emotional development has been a significant area of focus in the study of human development (Smith and Thelen, 2003; Bornstein and Lamb, 2015). Two theories, the human developmental cascades theory (Masten and Cicchetti, 2010) and the social-ecological systems theory (Bronfenbrenner, 1979, 1992), were utilized to guide the interpretation of school educators' perspectives on the complex constructs and implementation of SEL in a high-needs elementary school setting.

Human developmental cascades theory refers to the cumulative development of a person's competencies in one domain, resulting in the far-reaching and non-obvious development of competencies in other domains (Masten and Cicchetti, 2010; Adolph and Robinson, 2013). Thus, the development of a person's competencies in one domain over a period of life becomes the cornerstone for competencies in other newly emerging domains so that "competence begets competence" (Masten et al., 2005, p. 492). Human developmental cascades have a spreading effect within and across domains of function in a developing system. As Described by Thelen (1989), "Changes in any one domain, therefore, may become amplified and have system-wide reverberations" (p. 349). Developmental cascades have a profound influence on human development, which may result in positive or negative adaptive behaviors (Masten and Cicchetti, 2010). Drawing on the human developmental cascades theory, we sought to understand why educators showed significant buy-in toward SEL in a high-needs elementary school setting.

Social-ecological systems theory (Bronfenbrenner, 1979, 1992) focuses on a continued state of human development with four interrelated vital factors: the process, the person, context, and time (Bronfenbrenner and Morris, 1998). Among those four factors, the process and the context are the two most important factors that have been addressed in Bronfenbrenner's social-ecological systems theory (Tudge et al., 2009). Bronfenbrenner and Morris (1998) described the process of human development as a "complex reciprocal interaction between an active, evolving biopsychological human organism and the persons, objects, and symbols in its immediate external environment" (p. 996). Bronfenbrenner also suggested that individuals develop within a multi-level system of environmental and social organizations, including microsystem, mesosystem, exosystem, and macrosystem. This framework represents SEL from a broader social-ecological model perspective where transactions among people within their social and physical settings, over time and across personal, cultural, institutional, and political levels are examined (Bronfenbrenner, 1979, 1992). SEL

research grounded in the social-ecological systems theory focuses on exploring the development of children's SEL competencies within different levels of social environments and organizations, ranging from the proximal environments, such as schools and families (micro-level system), to the more distal environments, such as government policies and support (meso-level system), and attitudes and ideologies of the society (macro-level system) (Bornstein and Lamb, 2015). In applying the social-ecological systems theory, we sought to understand what constraints have impeded the SEL implementation and how the processes of SEL implementation could be facilitated in a high-needs elementary school setting.

Materials and methods

Research design

A case study design (Stake, 2005) was utilized to explore the school educators' perspectives on SEL in a US high-needs elementary school setting. Two sources of data, interviews and field notes, were used for this study. Data collection and analysis were followed by qualitative research traditions (Miles et al., 2014).

Participants and contexts

This study was in partnership with 15 school educators (**Table 1**), including 10 K-5 teachers, one special education teacher, one social worker, and three school leaders at one high-needs elementary school in the US School educators' consent was received following university IRB regulations for this study. All participants and the school were given pseudonyms.

Whitehead school is recognized as a high-needs rural elementary school since 98% of school students come from families with incomes below the poverty line and qualify for free or reduced-price meals (Higher Education Act, 2020). Whitehead school is operated in partnership with the school district and a local university, which serves 375 K-5 students with diverse ethnic backgrounds (59% African American, 20% Caucasian, 11% Hispanic, 10% Multi-Racial). Whitehead school is a "Lab School" created by North Carolina lawmakers through a state provision act in 2018. "A lab school shall provide an opportunity for research, demonstration, student support, and expansion of the teaching experience and evaluation regarding management, teaching, and learning." (Public School First NC, 2022).

A lab school operates much like a charter but is managed by a collaborating university. The Lab School act is intended to support high-need and low-performing schools, and improve student academic outcomes. Lab schools, like charter schools, can employ experimental teaching methods and are afforded more flexibility in designing and implementing their curriculum, their choice of calendar, and staffing models. When we talk about Whitehead school, we refer to it as a partnership school, since it had a close partnership with a university.

School-wide restorative practices have been adopted in Whitehead school since the academic year of 2019 to develop students' SEL competencies. In 2021, all educators in the school also participated in a 2-day restorative practices professional development to better facilitate students' SEL in the classrooms. Restorative practices focus on establishing "environments where members of the community take responsibility to repair harm when it occurs" (Gonzalez, 2012, pp. 300–301). The restorative circle was a frequent pedagogical practice utilized in the classrooms at the Whitehead school. Facilitated by the teachers daily, either at the beginning or the end of the classes, every student had a chance to share and speak on a specific topic by turns with a talking piece in the 10–15 min restorative circle. For example, teachers utilized the restorative circle to solve students' conflicts in class, listen to students' reflections and feedback, and help students with academic goal setting and planning.

TABLE 1 Participants information.

Name	Position	Race/Ethnicity	Years of experience
Samantha	Grade 5 teacher	Caucasian	24
Melissa	Grade 1 teacher	Caucasian	4
Teresa	Principal	African American	23
Jessica	Grade 5 teacher	Caucasian	24
Connor	Assistant principal	African American	10
Kinsley	Social worker	African American	10
Trinity	Grade 4 teacher	Hispanic	3
Julianna	Grade 5 teacher	Caucasian	19
Barbara	Grade 3 teacher	Caucasian	17
James	PE teacher	Caucasian	6
Angela	Grade 3 teacher	Caucasian	9
Martin	School director	Caucasian	51
Helen	Grade 2 special ed teacher	Caucasian	3
Tony	Grade 4 teacher	Caucasian	17
Tayler	Grade2 teacher	Caucasian	5

Data collection

Whitehead was visited twelve times by the researchers over the Spring and Fall semesters in 2021 to build rapport with the participants and school context, observe classes, and interview educators. Trained data collectors interviewed the school educators regarding their perspectives on SEL before, during, and after school hours. Fifteen individual interviews and one focus group were conducted. Each individual interview lasted for 40–45 min. The focus group interview lasted for 65 min. Semi-structured questions were asked about the school educators' specific SEL practices, as well as challenges related to their SEL practices. Example questions include: "Is social and emotional learning important in your class (or the school) and why?" and "What problems/issues or challenges do you see coming from the implementation of social and emotional learning?"

Non-participant class observations using organized methods of taking field notes were conducted in this study

(Emerson et al., 2011). The second and the third authors observed the classes and took field notes during 17 different days of visit (60–90 min) at the Whitehead elementary school over the two semesters. After each observation, the authors talked to the teacher regarding their perspectives on SEL implementation in their classrooms. A total of 15 individual interviews, one focus group, and 17 sets of field notes were written during the visit to the school.

Data analysis

Inductive analysis and constant comparison were used for data analysis (Miles et al., 2014). The process started with transcribing interviews, followed by importing all the data into NVivo 12 plus for further organization and management. Open coding was employed first. Open coding is the process of assigning labels to statements or events in the data and summarizing them in a word or short phrase (Miles et al., 2014). Open coding formed the first cycle of data analysis (Emerson et al., 2011), which produced nodes or thematic descriptions of the school educators' perspectives of SEL implementation at the Whitehead school. The second stage of analysis involved axial coding (Emerson et al., 2011; Miles et al., 2014), which aimed to identify conceptual links, discover relationships among categories, and generate themes by constant comparison and triangulation of the interview data from different educators and field notes.

Trustworthiness of the data analysis was achieved by guaranteeing the findings have credibility, dependability, confirmability, and transferability (Lincoln and Guba, 1985; Miles et al., 2014). Credibility was achieved through the extended periods of time staying with the teachers and school leaders during three-weekend professional development workshops on restorative practices. The continual presence would reduce possible distortions in data credibility. Consistent member checks also achieved credibility during the data analysis process. All the interview transcripts were reviewed by the school educators so that they would have the opportunity to adapt and modify any parts of the transcripts. The dependability of the findings was achieved by having a colleague who is familiar with this research but not directly involved in the study (the fourth author). This colleague reviewed and challenged our interpretations of the interview data and the themes that were subsequently drawn, resulting in a more reflective process for the data analysis. Confirmability of the findings was addressed by providing a reflexive, self-critical account through an iterative peer debriefing process with research colleagues and school educators. The non-participant observations, interviews, and field notes triangulated the findings. Transferability is a messy concept, and it was difficult to determine whether the findings found in Whitehead school could be found in other school contexts (Miles et al., 2014). However, findings from this study do contribute to the paucity of SEL studies in the context of rural high-needs elementary schools. Trustworthiness was strengthened by utilizing different data analysis strategies, constantly challenging the interpretations of the findings, establishing conceptual relations, and uncovering key themes through frequent peer debriefing within the researcher team.

Findings

The purpose of this study was to explore educators' buy-in and constraints toward SEL in a high-needs elementary school setting. Six themes were drawn from the interviews with the school educators and observations in the school: *prerequisite for academic success, essential skills for life success, lack of time, lack of preparation and development, home-school disconnection,* and *pushback from students.* The themes of *prerequisite for academic success* and *essential skills for life success* implicated the school educators' buy-in of SEL. The school educators' perceived constraints to SEL implementation were presented in the other four themes.

Educators' buy-in of SEL

Teaching and learning in schools have strong social, emotional, and academic components (Zins et al., 2004). Educators at Whitehead Elementary expressed a unified buy-in around the value of SEL and believed students' SEL competencies fostered by the RP approach are pivotal for their academic and life success.

Prerequisite for academic success

Educators at the Whitehead elementary school perceived the importance of promoting students' SEL development for their academic success. Samantha, a grade 5 teacher, shared a strong sense of buy-in toward SEL from the current staff and the necessity of incorporating SEL into the overall classroom atmosphere: "the people [teachers] we have this year have more of a 'hang in there' buy-in …the need for [SEL] is so great …they [students] can't learn academics if these basic needs of comfort, safety, and security aren't met first." For Melissa (Grade 1 Teacher), SEL competencies were vital prerequisites for students' academic achievement: "If [students] are not able to act well socially, it gets in the way of academics." Teresa (School Leader) added to Melissa's point, "[SEL] is not only helping those kids develop social skills and develop comradery with their peers. It's definitely helping them academically." She continued: "I see the possibility is definitely a decline in the number of days missed from school for disciplinary behavior. I see another benefit of just the ability to have conversations and to communicate in a positive way." Student's behavioral changes relating to academic performance were also observed after the one-year implementation of the schoolwide SEL implementation of restorative practices: "I am hearing a lot less, like, during math, I used to hear 'I can't, I can't, I'm not going to try,' and I feel like that's slowly changing." (Jessica, Grade 5 Teacher). Connor (School Leader) commented: "They [students] need to turn those behaviors toward academic learning … until we teach them how to be in relationship with one another, we can't accomplish some of the other goals."

Essential skills for everyday life

Educators acknowledge the importance of buy-in in SEL, believing SEL competencies were essential for students' life success, including relationship building, conflict resolution, social awareness, and self-awareness. For Kinsley (Social Worker), SEL provided an excellent opportunity for the students to learn trust,

which is the foundation of relationship building: "SEL is about trust. For trust with my kids and myself, they have to be able to know that they can come to me openly and know upfront." She also expressed the importance of educational efforts to promote SEL in schools: "If they [students] aren't emotionally prepared, if they aren't given the opportunity to learn differently, they're not going to be successful." Trinity (Grade 4 Teacher) perceived SEL as "understanding yourself and how you interact with other people in a positive way," which is essential for building a positive relationship with others. In the class, Julianna (Grade 5 Teacher) observed that students "enjoyed talking about themselves, they enjoyed learning about their peers, they enjoyed learning about their teachers. so that consistency is something that I think that they enjoy along with the relationship building of things." Julianna also added that "it's definitely important for our students to learn about kindness, compassion, about treating, treating one another with respect."

Conflict resolution was another essential SEL quality perceived by the educators at the Whitehead Elementary school. Barbara (Grade 3 Teacher) commented after the one-year schoolwide SEL program, "the [students'] conversations have changed. I mean, before it was, let's argue about it, let's yell and let's fuss and fight, and now it's, okay, let's talk this out, let's figure this." James (PE Teacher) explained his understanding of the importance of SEL as "teaching kids how to have feelings and how to nurture those feelings and how to, you know, repair harm when there is a negative or bad feeling." He further addressed that SEL taught the students "work together collaboratively" and learned "how to be kind" and "how to resolve conflict." The strategy of teaching peaceful conflict resolution was observed during the classes: "Conflict corner was one of the commonly used teaching strategies to develop students' SEL in classes. It is a place where students in conflicts could talk about the happenings, share their feelings, and reach a peaceful solution under the mediation with teachers" (Field Note, Research Team, YS).

The educators perceived social awareness and self-awareness as the other two essential qualities for students' life success. Connor accentuated the importance of teaching the students "to be productively interactive" and preparing them to be "not only good students but good citizens, good people to be around." Teresa and Barbara respectively added to Connor's point that SEL skills prepared students "to be a part of a class community, to be a part of a school community, to be a part of a city community." Students needed to "understand that 'I'm upset' and 'Let's not overreact on things that maybe we shouldn't." The improvement of social awareness and self-awareness was also reflected in the observation field notes during one of the PE classes: "Students stayed behind the lines and played fair. When students made mistakes, they were not upset. Some smiled and laughed, while others refocused quickly and moved to another pin to protect" (Field Note, Research Team, DH).

Educators' constraints to SEL implementation

Despite the evident and strong buy-in expressed by the educators, many constraints still impede and challenge the SEL implementation at the Whitehead elementary school. The educators perceived *lack of time, lack of preparation and development, home-school disconnection,* and *pushback from students* as salient constraints for SEL implementation at the Whitehead elementary school.

Lack of time

Lack of time was the most frequently mentioned constraint for SEL implementation at the Whitehead. Tayler (Grade 2 Teacher) commented that time is "a big constraint," as "you can't get to every person [in one class], so that is a definite constraint." She further explained that teaching SEL is a "slow and gradual process" and "goes against the instant gratification that we're so accustomed to." Angela (Grade 3 Teacher) added to Teresa's point that "I think sometimes you have to look at where's that sweet spot, but then there are so many other things you want to put in that sweet spot too, so being able to balance all in the class is very important." Barbara reiterated lack of time is "the biggest one" challenge for SEL implementation as "it takes a lot of time. It takes a lot of energy too. Some days I don't want to talk about it. Some days I just want to, you know, go sit down and move on." For Tony (Grade 4 Teacher), the lack of time to teach SEL was attributed to "academic pressure," which made "teachers feel like they don't have the time to focus on SEL in an intentional way." Connor expressed his concern about the lack of time for effective SEL implementation, "we still don't have enough time to do that [SEL] because that's hard work, right? You can't just give them a list and say, here, go do this. They've got to feel it. And that takes time."

Lack of preparation and development

Lack of professional preparation and development was another salient constraint frequently discussed by the educators at the Whitehead elementary school. Samantha mentioned limited preparation opportunities for teaching SEL: "I do not think any of my experience has prepared me for what I am dealing with these kids [SEL]. And this is my 19th year. I don't even have the ending for that." In a similar manner as Samantha, James commented on the professional development for SEL as "it [SEL] is not really a professional development that I've ever seen or been offered." Despite the lack of professional preparation and development for SEL, James also shared that the counselor and social worker helped the teachers for SEL in an informal way, "those conversations [about SEL] are definitely happening. But I don't think they're happening in a controlled like 'Hey, let's sit down and talk about it. I think it's done pretty much on teachers' free time or after school." Due to the lack of preparation and development, teachers were hesitant or uncertain about how to teach SEL in the classes: "During the class observations, the teacher spoke to students individually or removed them from the immediate situations. It appears that the teacher is a little uncertain as to how much restorative practice is being conducted in classrooms." (Field Note, Research Team, DH).

Recently, the school organized two professional development workshops on SEL for the teachers, but "there's been no kind of meaningful follow-up, which I think is another piece that's missing" (Martin, School Leader). Connor added to the comments for the recent SEL workshops "you [SEL researchers] not only have to work with teachers in that workshopping setting, but you also have to work with them in their classrooms. It's coaching and just ongoing discussions."

Home-school disconnection

Educators recognized home-school disconnection as another constraint for SEL implementation at the Whitehead. James observed and reflected: "A lot of our students go home to a very non-nurturing kind of destructive area or household. And I think it's hard for them because when they come here to us, they still have their defense up." He accentuated that: "I think our kids here are heavily, heavily influenced from their home life . . . the number one quote that is always said is 'My mama told me that if I get hit, I need to hit back'." Helen (Grade 2 Special Ed Teacher) complained about the consequences of the home-school disconnection on students' SEL development "teachers want to use the same [SEL] strategies at school, at home, but parents are so overwhelmed too that that's hard to keep consistent." The principal also added to Helen's point with specific examples of what students might experiences in their home environments: "Students don't always feel heard at home . . . because that's just not the way family dynamics are set up." Trinity proposed a way to improve the "home-school disconnect" by trying to "connect with parents and do some education on what the school process of [SEL] is" and teaching the parents "some of this [SEL] language and stuff that we're using at school in case they wanted to use that at home." A field note also reflected the educators' perceptions of building a home-school connection on students' SEL development during a summer restorative practice workshop: "The teachers and school leaders in the workshop acknowledged the importance of building common SEL language between teachers, students, and parents" (Field Note, Research Team, YS).

Pushback from students

Though the students overall have demonstrated some positive SEL progresses, there also has been pushback from students, which provided challenges to the student's social and emotional development at school. Connor cited students' confrontational nature owing to their challenging backgrounds beyond the school: "a lot of our kids are very streetwise, they use a lot of confrontation, they use language with one another that is oppositional, and they are suspicious of authority figures." Teresa added to Connor's point: "Some of our kids are very angry. They're very defensive. They can be very aggressive with one another. They can be very aggressive with the adults in the buildings." Samantha shared:

I'm almost desensitized to reacting because there's so much all the time with so many of them. You might have one situation in your school over a few years, and here it's almost all of our kids all of the time.

Despite teachers' efforts, behavior issues among students also have been observed outside of their classrooms. James commented on his students' behaviors in PE classes as "some kids do not understand how to control their voice and how to say things. So, they come across as very abrupt, very rude"; "One student firmly refused to walk to the line, and the PE teacher had to talk with his classroom teacher about his behavior" (Observations, Research Team, SB). Kinsley suggested that "sometimes your kids are just off," but teachers have to "set specific expectations that are clear and concise" so that students "know what our expectations are." During classroom observations, we found:

Younger students in third and fourth graders who have been instructed with restorative practices within the last academic year were polite, cooperative, and self-disciplined in the classes (Field Note, Research Team, YS).

Teachers also mentioned other external social media that might negatively influence students' SEL development, such as "news media and all of the conflict and chaos that's going on in the society" (Teresa).

Discussion

The purpose of this study was to explore educators' perspectives or voices of SEL in a high-needs rural elementary school setting, our research at this lab school/partnership school provides an understanding of what works and what needs improvement when schools attempt to develop students' SEL skills. Six themes were drawn from the interviews with the school educators and observations in the school: *prerequisite for academic success, essential skills for life success, lack of time, lack of preparation and development, home-school disconnection,* and *pushback from students*. The study confirmed the importance of SEL and revealed nested levels of constraints for school educators' SEL implementation from the school educators' perspectives.

School educators in this study confirmed the importance of SEL, acknowledging the interrelated and dynamic relationships between students' social, emotional, and academic competencies. Previous studies have shown quantitative evidence to conclude that SEL competencies are associated with students' academic success in schools (Durlak et al., 2011; Sklad et al., 2012; Taylor et al., 2017; Corcoran et al., 2018). Adding to the previous quantitative evidence where a general reciprocal relationship between SEL competencies and academic success was recognized, this study adds more in-depth qualitative evidence about the relationship between students' SEL competencies and academic success perceived by these elementary school educators. Drawing on the human developmental Cascades Theory (Capaldi, 1992; Dodge et al., 2008), we found evidence from the educators' voices that students' SEL competencies can be fundamental to students' academic success and life skill development. In other words, positive changes and development of students' SEL competencies have complemented and enhanced the students' potential for academic and later life success. Based on this finding, the study advocates more SEL programs and practices should be available for students in high-needs rural elementary schools. The research strongly advocates for school-based SEL research making use of qualitative research methods concentrated on the school context and relevant family and community connections to transform our understanding of the need for quality SEL implementation (Jagers et al., 2019).

In adopting a social-ecological systems perspective, we found nested levels of contextual constraints and process challenges the school educators had to confront when infusing SEL into the current teaching practice at the Whitehead. Constraints of lack of class time and pushback from students were recognized by the school educators in the immediate school context. Those constraints were categorized as the micro-level. A lack of professional preparation and development was recognized by the school educators as a significant constraint at the meso-level. In addition, school educators also perceived home-school disconnection as a salient constraint for students' SEL development at the meso-level.

Due to the pressures of the current school agendas, educators reported a lack of time as the most salient barrier to teaching SEL in the classrooms (Jones et al., 2017; Oberle and Schonert-Reichl, 2017; Humphrey et al., 2018). To combat the barrier of limited time, Ottmar et al. (2015) conducted an SEL intervention, where teachers were trained to create a well-managed and positive social environment in the classrooms. The study suggested the infusion of SEL into existing school curricula to overcome the constraint of limited class time caused by high demands for academic performance. In addition to the strategy of infusing SEL into the current school curriculum, we also suggest that SEL practices could be "sequenced, active, focused, and explicit" in the organization and sequencing of selection activities to achieve greater students' learning outcomes in SEL (Durlak et al., 2011).

Educators in high-need schools tend to have job-related stressors and burnout (LoCasale-Crouch et al., 2007; Pas et al., 2012). Those negative feelings and experiences can be deteriorated with students' pushback, which may lead to poor quality classroom instructions and may aggravate school and classroom climate (Hamre and Pianta, 2005; Downer et al., 2007). We advocate that high-need elementary school leaders reduce students' pushback by promoting schoolwide policies for SEL and using consistent SEL languages, which may "shift the norms, culture, and climate of their school" (Brackett et al., 2019, p. 154).

Lack of professional preparation and development emerged as a salient challenge for school educators to implement SEL successfully and was considered a meso-level constraint. It has been recognized that school educators' SEL competencies, pedagogical skills, and their understandings of students' SEL states and lives outside of school have significant influences on school-based SEL programs' process and effectiveness (Weissberg et al., 2015). However, previous studies have reported that the majority of school educators, especially teachers, received very few pre-service preparation and in-service professional learning experiences that focus explicitly on the content and pedagogical knowledge for students' SEL development (Bridgeland et al., 2013; Schonert-Reichl, 2017; McKown, 2019). The inclusion of children's development, teacher-student relationship, and positive learning environment in the pre-service teacher education program has been suggested as a way to empower school educators to facilitate implement SEL practices or programs at schools (Schonert-Reichl, 2017). These contextual factors have been recognized to be highly associated with the effectiveness of school-based SEL programs (Stoiber, 2011). In-service teachers continued professional development explicitly focusing on SEL should also highlight the importance of the context, colleague collaboration, roles of modeling, expert support, and opportunities for reflection and feedback (Immordino-Yang et al., 2019, Jagers et al., 2019).

Home-school disconnection stands out as another salient barrier at the meso-level of the social-ecological system for students' SEL. To provide students contextual opportunities to practice the SEL values learned at schools consistently in their homes, effective communications must be built between program providers, school personnel, and families (Payton et al., 2000). Families have played a critical role in students' SEL development since norms and values held by family members have a significant influence on their children's SEL (Meléndez and Martinek, 2015). The extent of family involvement in the school often determines the success of the SEL program (Holt et al., 2017; Jagers et al., 2019). To avoid fragmented SEL programs or practices "through which students pass like pinballs in a pinball machine" (Elias, 2019, p. 234), we suggest high-needs rural elementary schools adopt a multi-level SEL approach (Weissberg et al., 2015; Gordon et al., 2016; Brackett et al., 2019). Knowledge and strategies for developing SEL competencies are shared among multi-level social organizations, such as schools, families, and communities. Consequently, students will have the opportunity to learn and practice SEL competencies consistently across those social contexts. We also encourage affordable technologies (McKown, 2017; Williamson, 2017) to be employed for more effective communications regarding effective strategies to develop SEL competencies between the school and the family.

The major limitation of this study is the transferability of the findings. Since Whitehead is a lab school that is operated in partnership with the school district and a local university. It could be argued that this school has more resources and opportunities to conduct SEL programs compared to other public high-needs elementary schools. In addition, the number of participants in this study was relatively small. Those limitations might raise concerns about the transferability of the findings to other settings. Despite those limitations, we would argue that our findings would contribute to future SEL studies in the context of rural high-needs elementary schools.

Conclusion

The purpose of this study was to explore educators' buy-in and constraints toward SEL in a high-needs rural elementary school setting. In the process of observing classes and interviewing educators, we sought to better understand what works and what needs to be improved. The study provides evidence to support the importance of SEL and revealed nested levels of constraints for school educators' implementing SEL from the "voices" of school educators. The study recommends future school-based SEL research utilizing a qualitative research approach to gain more in-depth knowledge and understanding of the needs for contextually relevant SEL implementation. The study also calls for collaborative efforts and shared strategies to facilitate "legitimate" long-term partnerships between universities and schools, families, and communities, particularly in rural areas, in promoting a more holistic vision of the social and emotional development of our children (Meyers et al., 2015; San Antonio, 2018; Jagers et al., 2019).

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the University of North Carolina at Greensboro. The patients/participants provided their written informed consent to participate in this study.

Author contributions

BD conducted the research and wrote the manuscript. YS collected the data and drafted the manuscript. DH and SB collected and analyzed part of the data and contributed to the writing and proofreading process. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2023.1100667/full#supplementary-material

References

Adolph, K. E., and Robinson, S. R. (2013). "The road to walking: what learning to walk tells us about development," in *Oxford library of psychology. The Oxford handbook of developmental psychology (Vol. 1): body and mind*, ed. P. D. Zelazo (Oxford: Oxford University Press), 403–443. doi: 10.1093/oxfordhb/9780199958450.013.0015

Anyon, Y., Nicotera, N., and Veeh, C. A. (2016). Contextual influences on the implementation of a schoolwide intervention to promote students' social, emotional, and academic learning. *Child. Schools* 38, 81–88. doi: 10.1093/cs/cdw008

Avramidis, E., and Norwich, B. (2002). Teachers' attitudes towards integration/inclusion: a review of the literature. *Eur. J. Spec. Needs Educ.* 17, 129–147. doi: 10.1080/08856250210129056

Blair, C., and Raver, C. C. (2015). School readiness and self-regulation: a developmental psychobiological approach. *Annu. Rev. Psychol.* 66, 711–731. doi: 10.1146/annurev-psych-010814-015221

Bornstein, M. H., and Lamb, M. E. (2015). *Developmental science: an advanced textbook*, 7th Edn. Mahwah, NJ: Erlbaum. doi: 10.4324/9780203112373

Brackett, M. A., Bailey, C. S., Hoffmann, J. D., and Simmons, D. N. (2019). RULER: a theory-driven, systemic approach to social, emotional, and academic learning. *Educ. Psychol.* 54, 144–161. doi: 10.1080/00461520.2019.1614447

Bridgeland, J., Bruce, M., and Hariharan, A. (2013). *The missing piece: a national teacher survey on how SEL can empower children and transform schools. Collaborative for academic, social, and emotional learning*. Chicago, IL: Civic Enterprises.

Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.

Bronfenbrenner, U. (1992). "Ecological systems theory," in *Six theories of child development: revised formulations and current issues*, ed. R. Vasta (London: Jessica Kingsley Publishers), 187–249.

Bronfenbrenner, U., and Morris, P. A. (1998). "The ecology of developmental processes," in *Handbook of child psychology, Vol. 1: theoretical models of human development*, 5th Edn, eds W. Damon and R. M. Lerner (New York, NY: Wiley), 993–1023.

Capaldi, D. M. (1992). Co-occurrence of conduct problems and depressive symptoms in early adolescent boys: a 2-year follow-up at Grade 8. *Dev. Psychopathol.* 4, 125–144. doi: 10.1017/S0954579400005605

Collaborative for Academic, Social, and Emotional Learning [CASEL] (2023). *What is SEL?*. Chicago, IL: CASEL.

Collie, R. J., Shapka, J. D., and Perry, N. E. (2012). School climate and social–emotional learning: predicting teacher stress, job satisfaction, and teaching efficacy. *J. Educ. Psychol.* 104:1189. doi: 10.1037/a0029356

Corcoran, R. P., Cheung, A. C. K., Kim, E., and Xie, C. (2018). Effective universal school-based social and emotional learning programs for improving academic achievement: a systematic review and meta-analysis of 50 years of research. *Educ. Res. Rev.* 25, 56–72. doi: 10.1016/j.edurev.2017.12.001

Dodge, K. A., Greenberg, M. T., Malone, P. S., and Conduct Problems Prevention Research Group (2008). Testing an idealized dynamic cascade model of the development of serious violence in adolescence. *Child Dev.* 79, 1907–1927. doi: 10.1111/j.1467-8624.2008.01233.x

Downer, J. T., Rimm-Kaufman, S. E., and Pianta, R. C. (2007). How do classroom conditions and children's risk for school problems contribute to children's behavioral engagement in learning? *School Psychol. Rev.* 36, 413–432. doi: 10.1080/02796015.2007.12087938

Durlak, J. A., and DuPre, E. P. (2008). Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am. J. Commun. Psychol.* 41, 327–350. doi: 10.1007/s10464-008-9165-0

Durlak, J. A., Weissberg, R. P., and Pachan, M. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *Am. J. Commun. Psychol.* 45, 294–309. doi: 10.1007/s10464-010-9300-6

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

Dyson, B., Howley, D., and Shen, Y. (2019). Teachers' perspectives of social and emotional learning in aotearoa new zealand primary schools. *J. Res. Innov. Teach. Learn.* 12, 68–84. doi: 10.1108/JRIT-02-2019-0024

Dyson, B., Howley, D., and Shen, Y. (2021). 'Being a team, working together, and being kind': primary students' perspectives of cooperative learning's contribution to their social and emotional learning. *Phys. Educ. Sport Pedagogy* 26, 137–154. doi: 10.1080/17408989.2020.1779683

Ee, J., and Cheng, Q. L. (2013). Teachers' perceptions of students' social emotional learning and their infusion of SEL. *J. Teach. Teach. Educ.* 210, 1–14. doi: 10.12785/jtte/010201

Elias, M. J. (2019). What if the doors of every schoolhouse opened to social-emotional learning tomorrow: reflections on how to feasibly scale up high-quality SEL. *Educ. Psychol.* 54, 233–245. doi: 10.1080/00461520.2019.1636655

Elias, M. J., and Haynes, N. M. (2008). Social competence, social support, and academic achievement in minority, low-income, urban elementary school children. *School Psychol. Q.* 23, 474–495. doi: 10.1037/1045-3830.23.4.474

Elias, M. J., Zins, J. E., Weissberg, R. P., Frey, K. S., Greenberg, M. T., Haynes, N. M., et al. (1997). *Promoting social and emotional learning: guidelines for educators*. Alexandria, VA: Association for Supervision and Curriculum development.

Ellis, A. K. (2003). *Exemplars of curriculum theory*. New York, NY: Routledge.

Emerson, R. M., Fretz, R. I., and Shaw, L. L. (2011). *Writing ethnographic fieldnotes*. Chicago, IL: University of Chicago Press. doi: 10.7208/chicago/9780226206868.001.0001

Ennis, C. D., and Chen, A. (1995). Teachers' value orientations in urban and rural school settings. *Res. Q. Exerc. Sport* 66, 41–50. doi: 10.1080/02701367.1995.10607654

Evans, L. (2014). Leadership for professional development and learning: enhancing our understanding of how teachers develop. *Cambridge J. Educ.* 44, 179–198. doi: 10.1080/0305764X.2013.860083

Fleming, J., and Bay, M. (2004). "Social and emotional learning in teacher preparation standards," in *Building school success through social and emotional learning: implications for practice and research*, eds J. E. Zins, R. P. Weissberg, M. C. Wang, and H. J. Walberg (New York, NY: Teachers College Press), 94–110.

Gonzalez, T. (2012). Keeping kids in schools: restorative justice, punitive discipline, and the school to prison pipeline. *J. Law Educ.* 41, 281–335.

Gordon, B., Jacobs, J. M., and Wright, P. M. (2016). Social and emotional learning through a teaching personal and social responsibility based after-school program for disengaged middle-school boys. *J. Teach. Phys. Educ.* 35, 358–369. doi: 10.1123/jtpe.2016-0106

Greenberg, M. T., Weissberg, R. P., O'Brien, M. U., Zins, J. E., Fredericks, L., Resnik, H., et al. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *Am. Psychol.* 58:466. doi: 10.1037/0003-066X.58.6-7.466

Hamre, B. K., and Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Dev.* 76, 949–967. doi: 10.1111/j.1467-8624.2005.00889.x

Higher Education Act, (2020). *Higher Education Act, 20 USC 1021*. Available online at: https://www.govinfo.gov/content/pkg/USCODE-2020-title20/pdf/USCODE-2020-title20-chap28-subchapII-sec1021.pdf (accessed November 13, 2021).

Holt, N. L., Neely, K. C., Slater, L. G., Camiré, M., Côté, J., Fraser-Thomas, J., et al. (2017). A grounded theory of positive youth development through sport based on results from a qualitative meta-study. *Int. Rev. Sport Exerc. Psychol.* 10, 1–49. doi: 10.1080/1750984X.2016.1180704

Hulvershorn, K., and Mulholland, S. (2018). Restorative practices and the integration of social emotional learning as a path to positive school climates. *J. Res. Innov. Teach. Learn.* 11, 110–123. doi: 10.1108/JRIT-08-2017-0015

Humphrey, N., Barlow, A., and Lendrum, A. (2018). Quality matters: implementation moderates student outcomes in the PATHS curriculum. *Prev. Sci.* 19, 197–208. doi: 10.1007/s11121-017-0802-4

Immordino-Yang, M. H., Darling-Hammond, L., and Krone, C. R. (2019). Nurturing nature: how brain development is inherently social and emotional, and what this means for education. *Educ. Psychol.* 54, 185–204. doi: 10.1080/00461520.2019.1633924

Jagers, R. J., Rivas-Drake, D., and Williams, B. (2019). Transformative social and emotional learning (SEL): toward SEL in service of educational equity and excellence. *Educ. Psychol.* 54, 162–184. doi: 10.1080/00461520.2019.1623032

Jones, K., and Cater, M. (2020). An investigation of principals' social and emotional learning beliefs and attitudes. *Qual. Rep.* 25, 3204–3223. doi: 10.46743/2160-3715/2020.3575

Jones, S. M., and Bouffard, S. M. (2012). Social and emotional learning in schools: from programs to strategies and commentaries. *Soc. Policy Rep.* 26, 1–33. doi: 10.1002/j.2379-3988.2012.tb00073.x

Jones, S. M., and Doolittle, E. J. (2017). Social and emotional learning: introducing the issue. *Fut. Child.* 27, 3–11. doi: 10.1353/foc.2017.0000

Jones, S. M., Barnes, S. P., Bailey, R., and Doolittle, E. J. (2017). Promoting social and emotional competencies in elementary school. *Fut. Child.* 27, 49–72. doi: 10.1353/foc.2017.0003

Jones, S. M., Bouffard, S. M., and Weissbourd, R. (2013). Educators' social and emotional skills vital to learning. *Phi Delta Kappan* 94, 62–65. doi: 10.1177/003172171309400815

Jones, S. M., McGarrah, M. W., and Kahn, J. (2019). Social and emotional learning: a principled science of human development in context. *Educ. Psychol.* 54, 129–143. doi: 10.1080/00461520.2019.1625776

Katz, D., Mahfouz, J., and Romas, S. (2020). Creating a foundation of well-being for teachers and students starts with SEL curriculum in teacher education programs. *Northw. J. Teach. Educ.* 15:5. doi: 10.15760/nwjte.2020.15.2.5

Lincoln, Y. S., and Guba, E. G. (1985). *Naturalistic inquiry*. London: Sage. doi: 10.1016/0147-1767(85)90062-8

LoCasale-Crouch, J., Konold, T., Pianta, R., Howes, C., Burchinal, M., Bryant, D., et al. (2007). Observed classroom quality profiles in state-funded pre-kindergarten programs and associations with teacher, program, and classroom characteristics. *Early Childh. Res. Q.* 22, 3–17. doi: 10.1016/j.ecresq.2006.05.001

Masten, A. S., and Cicchetti, D. (2010). Developmental cascades. *Dev. Psychopathol.* 22, 491–495. doi: 10.1017/S0954579410000222

Masten, A. S., Roisman, G. I., Long, J. D., Burt, K. B., Obradoviæ, J., Riley, J. R., et al. (2005). Developmental cascades: linking academic achievement and externalizing and internalizing symptoms over 20 years. *Dev. Psychol.* 41, 733–746. doi: 10.1037/0012-1649.41.5.733

McKown, C. (2017). Social-emotional assessment, performance, and standards. *Fut. Child.* 27, 157–178. doi: 10.1353/foc.2017.0008

McKown, C. (2019). Challenges and opportunities in the applied assessment of student social and emotional learning. *Educ. Psychol.* 54, 205–221. doi: 10.1080/00461520.2019.1614446

Meléndez, A., and Martinek, T. (2015). Life after project effort: applying values acquired in a responsibility-based physical activity program. *Rev. Int. Ciencias Del Deporte* 11, 258–280. doi: 10.5232/ricyde2015.04105

Meyers, A. B., Tobin, R. M., Huber, B. J., Conway, D. E., and Shelvin, K. H. (2015). Interdisciplinary collaboration supporting social-emotional learning in rural school systems. *J. Educ. Psychol. Consul.* 25, 109–128. doi: 10.1080/10474412.2014.929956

Miles, M. B., Huberman, A. M., and Saldana, J. (2014). *Qualitative data analysis*. London: Sage.

Oberle, E., and Schonert-Reichl, K. A. (2017). "Social and emotional learning: Recent research and practical strategies for promoting children's social and emotional competence in schools," in *Handbook of social behavior and skills in children*, ed. J. L. Matson (Berlin: Springer), 175–197. doi: 10.1007/978-3-319-64592-6_11

Oberle, E., Schonert-Reichl, K. A., Hertzman, C., and Zumbo, B. D. (2014). Social–emotional competencies make the grade: predicting academic success in early adolescence. *J. Appl. Dev. Psychol.* 35, 138–147. doi: 10.1016/j.appdev.2014.02.004

Osher, D., Kidron, Y., Brackett, M., Dymnicki, A., Jones, S., and Weissberg, R. P. (2016). Advancing the science and practice of social and emotional learning: looking back and moving forward. *Rev. Res. Educ.* 40, 644–681.

Ottmar, E. R., Rimm-Kaufman, S. E., Larsen, R. A., and Berry, R. Q. (2015). Mathematical knowledge for teaching, standards-based mathematics teaching practices, and student achievement in the context of the responsive classroom approach. *Am. Educ. Res. J.* 52, 787–821. doi: 10.3102/0002831215579484

Palacios, A. F., and Lemberger-Truelove, M. E. (2019). A counselor-delivered mindfulness and social–emotional learning intervention for early childhood educators. *J. Humanistic Couns.* 58, 184–203. doi: 10.1002/johc.12119

Pas, E. T., Bradshaw, C. P., and Hershfeldt, P. A. (2012). Teacher-and school-level predictors of teacher efficacy and burnout: identifying potential areas for support. *J. School Psychol.* 50, 129–145. doi: 10.1016/j.jsp.2011.07.003

Payton, J. W., Wardlaw, D. M., Graczyk, P. A., Bloodworth, M. R., Tompsett, C. J., and Weissberg, R. P. (2000). Social and emotional learning: a framework for promoting mental health and reducing risk behavior in children and youth. *J. School Health* 70, 179–185. doi: 10.1111/j.1746-1561.2000.tb06468.x

Public School First NC. (2022). *The facts on lab schools*. Raleigh, NC: Public School First NC.

San Antonio, D. M. (2018). Collaborative action research to implement social-emotional learning in a rural elementary school: helping students become "little kids with big words. *Can. J. Action Res.* 19, 26–47. doi: 10.33524/cjar.v19i2.384

Schonert-Reichl, K. A. (2017). Social and emotional learning and teachers. *Fut. Child.* 27, 137–155. doi: 10.1353/foc.2017.0007

Sklad, M., Diekstra, R., Ritter, M. D., Ben, J., and 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? *Psychol. Schools* 49, 892–909. doi: 10.1002/pits.21641

Smith, L. B., and Thelen, E. (2003). Development as a dynamic system. *Trends Cogn. Sci.* 7, 343–348. doi: 10.1016/S1364-6613(03)00156-6

Stake, R. E. (2005). "Qualitative case studies," in *The Sage handbook of qualitative research*, eds N. K. Denzin and Y. S. Lincoln (Thousand Oaks, CA: Sage Publications Ltd), 443–466.

Stoiber, K. C. (2011). Translating knowledge of social-emotional learning and evidence-based practice into responsive school innovations. *J. Educ. Psychol. Consul.* 21, 46–55. doi: 10.1080/10474412.2011.549039

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

Thelen, E. (1989). "Self-organization in developmental processes: can systems approaches work?," in *The Minnesota symposia on child psychology: Vol. 22. Systems and development*, eds M. R. Gunnar and E. Thelen (Hillsdale, NJ: Erlbaum), 336–374. doi: 10.1002/9780470753507.ch18

Tudge, J. R. H., Mokrova, I., Hatfield, B. E., and Karnik, R. B. (2009). Uses and misuses of bronfenbrenner's bioecological theory of human development. *J. Fam. Theory Rev.* 1, 198–210. doi: 10.1111/j.1756-2589.2009.00026.x

Turnbull, B. (2002). Teacher participation and buy-in: implications for school reform initiatives. *Learn. Environ. Res.* 5, 235–252. doi: 10.1023/A:1021981622041

Walker, T. (2020). *Social-emotional learning should be priority during COVID-19 crisis*. NEAToday. Available online at: http://neatoday.org/2020/04/15/social-emotional-learning-during-covid/ (accessed April 14, 2020).

Weissberg, R. P., Durlak, J. A., Domitrovich, C. E., and Gullotta, T. P. (eds) (2015). "Social and emotional learning: past, present, and future," in *Handbook of social and emotional learning: Research and practice*, (New York, NY: The Guilford Press), 3–19.

Wigelsworth, M., Lendrum, A., Oldfield, J., Scott, A., Ten Bokkel, I., Tate, K., et al. (2016). The impact of trial stage, developer involvement and international transferability on universal social and emotional learning programme outcomes: a meta-analysis. *Cambridge J. Educ.* 46, 347–376. doi: 10.1080/0305764X.2016.1195791

Williamson, B. (2017). Decoding ClassDojo: psycho-policy, social-emotional learning and persuasive educational technologies. *Learn. Media Technol.* 42, 440–453. doi: 10.1080/17439884.2017.1278020

Zins, J. E., and Elias, M. J. (2007). Social and emotional learning: promoting the development of all students. *J. Educ. Psychol. Consul.* 17, 233–255. doi: 10.1080/10474410701413152

Zins, J. E., Weissberg, R. P., Wang, M. C., and Walberg, H. J. (2004). *Building academic success on social and emotional learning: what does the research say?*. New York, NY: Teachers College Press.

frontiers | Frontiers in Psychology

TYPE Perspective
PUBLISHED 06 March 2023
DOI 10.3389/fpsyg.2023.1011039

OPEN ACCESS

EDITED BY
Andres Molano,
University of Los Andes,
Colombia

REVIEWED BY
Baiba Martinsone,
University of Latvia,
Latvia
Celene Domitrovich,
Georgetown University,
United States

*CORRESPONDENCE
Dana Charles McCoy
dana_mccoy@gse.harvard.edu

SPECIALTY SECTION
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

RECEIVED 03 August 2022
ACCEPTED 15 February 2023
PUBLISHED 06 March 2023

CITATION
McCoy DC and Hanno EC (2023) Systemic barriers and opportunities for implementing school-based social–emotional learning interventions in low-income and conflict-affected settings.
Front. Psychol. 14:1011039.
doi: 10.3389/fpsyg.2023.1011039

Systemic barriers and opportunities for implementing school-based social–emotional learning interventions in low-income and conflict-affected settings

Dana Charles McCoy* and Emily C. Hanno

Human Development, Learning, and Teaching, Harvard Graduate School of Education, Cambridge, MA, United States

Children living in low-income and conflict-affected settings face unique systemic risk factors that shape their social, emotional, and mental well-being. However, little is known about how these and other systemic factors may impede or support the delivery of social–emotional learning (SEL) interventions in these contexts. In this article, we draw from our experience delivering and evaluating a classroom-based SEL curriculum in Rio de Janeiro, Brazil to surface systemic barriers and opportunities for implementing SEL interventions in low-income, conflict-affected settings. Specifically, we identify (1) culture, (2) timing, and (3) government support and stability as factors underlying SEL program demand, dosage, quality, and effectiveness. We provide recommendations for improving implementation of SEL programs in low-income and conflict-affected contexts, including the importance of building pro-active partnerships, using qualitative research, and investing in adaptation to both understand and address systemic barriers.

KEYWORDS

social–emotional learning, low-and middle-income countries, violence, conflict, implementation

1. Introduction

Nearly 90 percent of children live in a low-or middle-income country (LMIC; World Bank, 2019) and one in six lives in a conflict zone (Kamøy et al., 2021). Children growing up in low-income and conflict-affected settings face unique risks that may jeopardize their social–emotional well-being, including exposure to trauma and reduced access to protective resources (Black et al., 2017; Murphy et al., 2017). Indeed, 10 to 20 percent of children and adolescents living in LMICs are affected by a mental health problem, with many cases thought to be preventable (Kieling et al., 2011).

School-based social–emotional learning (SEL) supports – including direct instruction in SEL strategies and/or teacher training in positive behavior management and stress reduction – have been shown to meaningfully improve children's social–emotional wellbeing and mental health (Durlak et al., 2011; Wigelsworth et al., 2016; Taylor et al., 2017; Blewitt et al., 2018). Importantly, this evidence largely comes from high-income

countries, with substantially less known about school-based SEL programming in non-Western, low-income, and/or violence-afflicted settings (Barry et al., 2013), which are characterized by a complex system of risks, resources, and cultural imperatives that likely also affect program implementation.

In this Perspective article, we draw from our experiences evaluating a school-based SEL intervention in Rio de Janeiro, Brazil to describe possible barriers and opportunities for implementing SEL programming in low-income, conflict-affected settings. Notably, whereas much of the literature on intervention implementation has focused on the individual-, school-, or community-level factors that shape SEL program dosage, fidelity, and reach (e.g., teacher characteristics, school climate; Durlak and DuPre, 2008; Durlak, 2016), here we emphasize broader macro-and exo-systemic considerations that may either promote or interfere with SEL implementation in low-income, conflict-affected contexts. In doing so, our goal is to identify potential paths forward for supporting children's social–emotional wellbeing in these under-represented contexts. To support our arguments, we incorporate evidence from our own work alongside findings from a small but growing set of published SEL program evaluations in LMICs and high-violence settings (e.g., Ştefan and Miclea, 2013; Huang et al., 2017; Baker-Henningham and Walker, 2018; Bilir Seyhan et al., 2019; Torrente et al., 2019; Aber et al., 2021; Tubbs Dolan et al., 2022). Nevertheless, we acknowledge that our points are largely speculative and need to be confirmed with future research.

2. Overview of the Programa Compasso evaluation

This article draws from our experiences implementing Programa Compasso ("Compass Program") within a randomized control trial conducted in 90 primary schools across Rio de Janeiro, Brazil in 2017. A middle-income country, Brazil is characterized by robust social services for families and high economic inequality, violence, and instability. In particular, Brazil's homicide rate is about five times the global average (27 vs. 5 per 100,000 people, respectively), ranking it among the most violent countries in the world (UN Office on Drugs and Crime, 2021).

Programa Compasso is a universal, school-wide SEL curriculum that includes 22 weekly, 50 min lessons delivered to students by classroom teachers. These scripted lessons provide direct instruction in emotion knowledge, self-regulation, executive function, empathy, and social problem solving strategies that is reinforced by activities, games, and materials. Lessons were adapted from the United States-based Second Step program by a Brazilian NGO, Instituto Vila Educação. After piloting the Programa Compasso curriculum in 17 schools in São Paulo in 2015, Instituto Vila Educação made slight modifications to the lessons and expanded the program to include student workbooks to reinforce content at home, a parent engagement component, and teacher trainings focused on improving implementation.

The evaluation of Programa Compasso was funded by a Brazilian education-focused foundation and included 90 schools that were randomized within matched pairs at the beginning of 2017 to either an intervention or a business-as-usual control condition. A total of 3,018 students from 90 third-and 90 fifth-grade classrooms took part in the evaluation, which included teachers' reports of student behavior problems and group-administered direct assessments of student executive function and emotion knowledge as outcomes. Results of the 2017 evaluation showed no average impacts of the Programa Compasso intervention after 1 year on the five outcomes tested. We did, however, observe small, positive program impacts ($d = 0.15$ *SD*s) on students' labeling of emotional expressions and inhibitory control in communities characterized by below-average levels of violence. Although budgetary limitations and local data collection restrictions prevented us from collecting detailed implementation data to formally contextualize these impacts, findings from a voluntary end-of-year teacher survey suggested that responding treatment teachers delivered an average of just 13 of the 22 intended lessons. Furthermore, consistent with prior SEL evaluations in conflict-affected settings (e.g., Tubbs Dolan et al., 2022), student attendance was generally quite low during the intervention period, suggesting limited take-up. For additional details of the intervention, study design, and results, see McCoy et al. (2021).

3. Barriers and opportunities for ensuring SEL program implementation

3.1. Culture

Based on our experience with Programa Compasso, perhaps the most salient systemic influence on the implementation of SEL programming in low-income, conflict-affected settings is *culture*. Cultural values prioritizing SEL in Brazil provided an opportunity for our study to take place by fomenting initial demand for SEL services (Durlak and DuPre, 2008). Brazilians generally adhere to horizontal collectivist values, emphasizing group well-being and prosocial behavior alongside individual equality (Carlo et al., 2007; Martinez et al., 2020). Although Brazilians report that the learning of these values begins at home, they also believe that the education system is key to equipping children –particularly from low-income backgrounds –with skills to get along with others and endure the hardships of everyday life (Dessen and Torres, 2002). Indeed, Brazil's national learning standards –codified in the Base Nacional Comum Curricular (BNCC) –explicitly emphasize SEL skills as outcomes of public education, including responsibility and citizenship, empathy and cooperation, self-knowledge and self-care, and critical and creative thinking (Movimento Pela Base Nacional Comum, 2018). In our experience, these collective values around (1) the importance of SEL and (2) the central role that schools play in its socialization generated the community appetite that allowed our work to be funded, implemented, and taken-up in Rio. Had we attempted this work in a different cultural context (e.g., areas of sub-Saharan Africa where parental demand prioritizes schools' academic rigor; Bidwell et al., 2014; Wolf, 2020), we speculate that gaining community buy-in would have been more difficult.

Culture also positively *and* negatively shaped the structure and implementation of Programa Compasso itself. Prior meta-analysis quantifying the "cross-cultural transportability" of SEL programming suggests that interventions implemented outside the

countries in which they were developed tend to be less effective for improving certain outcomes than those implemented in their country of origin (Wigelsworth et al., 2016). Accordingly, considering cultural relevance is important for optimizing the success of school-based SEL programming in LMICs and conflict-affected settings. To maximize the cultural appropriateness of Programa Compasso, Instituto Vila Educação added a student workbook and parent meetings to the original US-based Second Step curriculum to more explicitly reflect the centrality of the family system in Brazil (Carlo et al., 2007) and teachers' beliefs that the program would not work without investments from parents. The core lessons from Second Step were also modified, but mostly in minor ways (e.g., replacing references to skiing with football/soccer) to avoid tampering with the program's "active ingredients" (Durlak, 2016).

Although many teachers praised the final intervention content and structure as "relevant" and "productive," some reported that the program remained "decontextualized from [their] own reality." One educator, for example, questioned the importance of teaching students to individually regulate (read: control) their emotions, saying instead that Brazilian children should "embrace" emotions, sharing and co-regulating them with peers and caregivers. Teachers also sometimes pushed back against the scripted nature of the lessons, instead pursuing *ad hoc* approaches to teaching the SEL topic of the week. We speculate several different opportunities could have mitigated these challenges. For example, more substantive –or "deep structure" (Ahluwalia et al., 1999) –adaptations to Programa Compasso's lesson content could have improved cultural alignment. Alternatively, helping teachers to understand how existing lessons could be used to support context-specific goals (e.g., co-regulation) may have improved their motivation to implement the program as-is, with greater fidelity.

Finally, our findings highlight the importance of viewing culture not just as a stable, broad-scale influence, but also as a force that shapes SEL program implementation and effectiveness more locally (Vélez-Agosto et al., 2017). Our evaluation showed that Programa Compasso improved child outcomes within neighborhoods characterized by lower-than-average levels of violence; however, in higher-violence communities, the program showed no impacts. Independent of the broader cultural values of Brazil as a whole, these findings could reflect more localized variability in the socialization practices used in safer versus less safe environments. Indeed, most SEL programs –including Programa Compasso –take an "approach" orientation to teaching conflict resolution, encouraging children to stop, think, and discuss their feelings with others. Although these mainstream strategies seem to have been modestly effective in low-violence neighborhoods, they may have contradicted the avoidant strategies often taught to protect children's physical safety in conflict-affected settings (e.g., to quickly disengage from conflict, run away, etc.; Kliewer et al., 2006), limiting their applicability, take-up, and effectiveness in Rio's more dangerous communities. Once again, these results reinforce the importance of aligning SEL programmatic strategies with cultural values. Importantly, however, they also encourage taking a narrower, more localized view of culture to avoid fallacies regarding cultural homogeneity (e.g., within all LMICs, within Brazil, etc.).

3.2. Timing

A second noteworthy factor affecting SEL implementation in low-income, conflict-affected settings is *timing*. Around the start of our study in 2017, the federal Ministry of Education ratified the BNCC, increasing political appetite for curricular approaches targeting SEL-related learning standards. Simultaneously, Brazil was experiencing widespread gang violence and police shootouts amidst an economic recession and several highly publicized government corruption scandals. In particular, Rio experienced a 26 percent surge in community crime (Fonseca and Alper, 2018), forcing schools to close for so-called "violence days" during 99 of the first 107 school days in 2017 (de Oliveira, 2017). Even when children were attending school, educators reported concerns with students' social–emotional wellbeing, noting that children "are here, but their head is always outside" (Londoño, 2017). Finally, a simultaneous move from part-to full-day schooling for many primary schools in Rio effectively doubled the "supply" of instructional hours available to meet the demand for SEL programming inspired by the BNCC and rise in community violence. Collectively, these "opportunities" opened doors for us to implement Programa Compasso in Rio. Nevertheless, the stress these same factors placed on educators may have also negatively affected their capacity to deliver the intervention with sufficient dose or fidelity. For example, more than 60 percent of teachers in our sample said that violence affects their school at least "a little," and approximately 70 percent said that either they or their students had trouble getting to school because of violence.

Overall, these findings suggest that the timing of broader political, cultural, or societal events can present as both opportunities *and* barriers for program implementation. In low-income and conflict-affected settings, destabilizing events like the ones we observed in Rio (e.g., bursts of violence, policy shifts) are especially common, and often result in calls for supporting the social–emotional needs of children. Nevertheless, as we observed in our study, the stress and instability created by these events may also limit the bandwidth of individuals tasked with delivering SEL supports. Such tensions were also observed globally amidst the COVID-19 pandemic, as widespread teacher burnout coincided with broader demand for and availability of SEL services (Reimers et al., 2020; Gultom et al., 2022). Taking advantage of the opportunities to affect change presented by these broader social shocks, while also compensating for the additional stress that they place on program implementers, could be one path forward for successful implementation of SEL programming in LMICs and high-violence areas.

Such shocks may also open opportunities for creative SEL solutions that either supplement or replace traditional school-based approaches. Community-or technology-based programming (e.g., parent groups, apps) may be particularly useful for reaching children when they cannot attend school safely. Virtual SEL programs have been especially popular since the onset of the COVID-19 pandemic (Katzman and Stanton, 2020). In 2021, for example, our team worked with Rio's Ministry of Education to deliver lessons on stress management remotely via television. Although these non-school-based approaches to SEL hold promise for addressing access gaps during times of crisis, evidence regarding their effectiveness is still emerging.

3.3. Government support and stability

A third systemic factor that we observed to underpin SEL implementation is *government support and stability*. In many LMIC and conflict-affected settings, government officials (e.g., staff from Ministries of Education) influence not only whether an SEL program is taken up, but whether it is sustained and how it is implemented. Government turnover is a major issue globally, and especially in contexts characterized by instability and public mistrust. In Brazil, it is common for new governments to abolish or severely restructure programs established by their predecessors. In the case of Programa Compasso, the 2016 municipal elections in Rio led to staffing changes in the Ministry of Education that coincided with the start of our study. Although we were fortunate to receive permission from the new government to continue our research, the initial enthusiasm we received from the Ministry of Education was tempered as new officials took office and focused on their own agendas. We speculate that these changes ultimately affected implementation, with new government staff providing less oversight, guidance, and encouragement for schools to deliver Programa Compasso than their predecessors.

Even in non-election years, inconsistencies in government oversight and support can affect SEL implementation by breeding mistrust from program implementers. Some teachers in our study voiced resistance to Programa Compasso solely because it was mandated by officials who they perceived as unfamiliar with and unsupportive of their day-to-day work. Despite the shift to full-day teaching, teachers reported being over-worked, having limited time, and needing social–emotional services for *themselves* before they could support their students. Studies have shown similar patterns of government mistrust in Brazil, with many teachers pushing for increased autonomy regarding resource allocation and curricular planning (Lennert da Silva and Mølstad, 2020).

Despite these barriers, durable partnerships between NGOs and public officials whose jobs are not tied to a particular political party or election result (i.e., "comissionados" in Brazil) could help to sustain implementation in the face of government instability. Although we lacked such partnerships in Rio, we have seen in other areas of Brazil that collaborations focused outwardly on advocacy, awareness-raising, and empowerment may be particularly effective. For example, efforts led by a coalition of foundations to educate political candidates and the public in Ceará, Brazil about the importance of the early years have helped to ensure the popularity –and longevity –of early childhood programs in the state (Fundação Maria Cecilia Souto Vidigal, 2021). Furthermore, internal dialogs focused on building trust and equality between government staff and program implementers could help to overcome the issues of mistrust that we observed in our study. In particular, such partnerships can support implementers' (e.g., teachers') understanding of the value of a given SEL program while also improving their capacity to make improvements aligned with government goals. Indeed, experimental evidence from Brazil has shown that shifting decision-making authority from governments to teachers can reduce teacher turn-over and improve student learning and social–emotional outcomes (Piza et al., 2020).

4. Recommendations and paths forward

Overcoming the challenges highlighted above (along with additional challenges common to low-income, high-violence contexts but not explicitly considered here) requires creative solutions. Beyond directly addressing systemic barriers (e.g., through broad-scale violence reduction efforts), there are several steps that researchers and practitioners can take *before* implementing SEL programs that may improve delivery and uptake in LMICs and conflict-affected settings. First, investments in durable, trusting partnerships with decisionmakers and advocates (e.g., government officials, NGO staff, funders) are critical for gauging initial demand for SEL programming, for promoting a sense of co-ownership that maintains this demand and associated supports over time, and for overseeing implementation (Durlak and DuPre, 2008). Such partnerships take time to build and sustain, and must be based on mutual trust, responsiveness, patience, and flexibility (Aber et al., 2021). Critically, teachers and other implementers should be included in these partnerships to maximize their buy-in as active contributors to the programming that they are ultimately expected to deliver, and to ensure that efforts to oversee implementation are not perceived as reducing their autonomy. Second, understanding the cultural and political appetite for SEL programming, as well as whether the timing is right for proceeding, is a must. Implementing SEL interventions in unwelcoming contexts is likely to be a Sisyphean task. Conducting qualitative research with a variety of parties (e.g., government officials, teachers, and families) can identify potential systemic roadblocks like those described above, along with more localized opportunities and barriers within communities or school systems (Tinajero et al., 2016). Working with teachers to sensitize them to the benefits of SEL programs –both for their students *and* their own well-being –is also critical. Third, adaptation of imported programmatic content and structures is needed to maximize political and cultural relevance. As noted above, adaptation should ensure alignment not only with broad-scale cultural norms (e.g., collectivist values), but also more localized, community-and school-specific priorities. Once again, this adaptation should occur in partnership with program implementers and beneficiaries, as well as researchers familiar with local SEL program best practices (Durlak, 2016).

5. Discussion

Demand for SEL programming in low-income, conflict-affect settings is high. Nevertheless, the same systemic factors that increase children's risk for social–emotional challenges in these settings also shape SEL service selection, delivery, and take-up. Our experiences implementing and evaluating a classroom-based SEL program for primary students in Rio de Janeiro, Brazil highlighted (1) culture, (2) timing, and (3) government support and stability as key systemic factors underlying SEL program demand, implementation, and effectiveness in LMICs and conflict-affected settings. Understanding and addressing these factors through partnerships with government, NGO, school, community, and family collaborators is critical for optimizing the potential of SEL programming *before* implementation begins.

BOX 1 Key steps for successful implementation of SEL programs in low-income, conflict-affected settings.

1. Invest in partnerships

- Build relationships between interested parties to share information, equalize power dynamics, and develop co-ownership.
- Ensure involvement of multiple parties, including government officials (ideally whose positions are not tied to a particular political party or election result), NGO staff, funders, researchers, community leaders, program implementers (e.g., school leaders, teachers), and program beneficiaries (e.g., students, families).
- Allow plenty of time to build partnerships *before* making key decisions about program implementation.

2. Understand the context

- Conduct qualitative research to gauge initial demand for SEL programming in the particular setting, as well as barriers and opportunities for ongoing implementation.
- Pay particular attention to systemic barriers and opportunities related to (1) culture, (2) timing, and (3) government support and stability.
- Involve multiple interested parties to understand contextual needs at multiple "levels," ranging from broader government systems to specific community/school priorities.

3. Adapt as needed

- Adjust program content and structures to address key cultural factors prior to implementation. Ensure that cultural adaptations are attuned to both broad-scale cultural norms and localized priorities (e.g., of a given type of community or school).
- Invest in high-quality, autonomy-focused training that (1) educates program implementers (e.g., teachers) regarding the value of SEL programming for their students *and* themselves, and (2) guides them regarding how content should and should not be further adapted to meet their needs without compromising core program ingredients.
- Involve multiple experts in the adaptation process, including local researchers, implementers, and beneficiaries.
- Consider the timing of broader social events (e.g., elections, policy shifts, outbreaks of violence) when deciding when and how to implement. Address any specific barriers emerging from these events before proceeding.
- Do not be afraid to pivot or cancel implementation entirely if conditions are not right.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

DM conceived the manuscript structure, devised the argument, and drafted the manuscript. EH contributed to the manuscript structure and argument, and provided critical edits to the manuscript text. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by a series of grants from the Harvard Lemann Brazil Research Fund. The funder had no role in the design, execution, analysis, or reporting of the research reported in this article.

Acknowledgments

The authors are grateful for the support of the David Rockefeller Center for Latin American Studies, the Rio de Janeiro Ministry of Education, Committee for Children, Instituto Vila Educação, the Rede Aplicada FGV (Applied Research and Knowledge Network), and CNPq. The authors are particularly grateful to Marcela Almeida of Instituto Vila Educação, who provided helpful inputs and feedback regarding this article. The authors would also like to thank study collaborators from the São Paulo School of Economics, including Vladimir Ponczek, Cristine Pinto, Gabriela Fonseca, and Natália Marchi. Finally, The authors would like to express our gratitude to the government officials, principals, pedagogical coordinators, teachers, and students who participated in the evaluation of the Programa Compasso intervention described in this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Aber, J. L., Dolan, C. T., Kim, H. Y., and Brown, L. (2021). Children's learning and development in conflict-and crisis-affected countries: building a science for action. *Dev. Psychopathol.* 33, 506–521. doi: 10.1017/S0954579420001789

Ahluwalia, J., Baranowski, T., Braithwaite, R., and Resnicow, K. (1999). Cultural sensitivity in public health: defined and demystified. *Ethn. Dis.* 9, 10–21.

Baker-Henningham, H., and Walker, S. (2018). Effect of transporting an evidence-based, violence prevention intervention to Jamaican preschools on teacher and class-wide child behaviour: a cluster randomised trial. *Glob. Ment. Health* 5:e7. doi: 10.1017/gmh.2017.29

Barry, M. M., Clarke, A. M., Jenkins, R., and Patel, V. (2013). A systematic review of the effectiveness of mental health promotion interventions for young people in low and middle income countries. *BMC Public Health* 13:e7, 1–19. doi: 10.1186/1471-2458-13-835

Bidwell, K., Watine, L., and Perry, K., (2014). *Exploring Early Education Programs in Peri-Urban Settings in Africa*. Innovations for Poverty Action: Accra, Ghana.

Bilir Seyhan, G., Ocak Karabay, S., Arda Tuncdemir, T. B., Greenberg, M. T., and Domitrovich, C. (2019). The effects of promoting alternative thinking strategies preschool program on teacher–children relationships and children's social competence in Turkey. *Int. J. Psychol.* 54, 61–69. doi: 10.1002/ijop.12426

Black, M. M., Walker, S. P., Fernald, L. C., Andersen, C. T., DiGirolamo, A. M., Lu, C., et al. (2017). Early childhood development coming of age: science through the life course. *Lancet* 389, 77–90. doi: 10.1016/S0140-6736(16)31389-7

Blewitt, C., Fuller-Tyszkiewicz, M., Nolan, A., Bergmeier, H., Vicary, D., Huang, T., et al. (2018). Social and emotional learning associated with universal curriculum-based interventions in early childhood education and care centers: a systematic review and meta-analysis. *JAMA Netw. Open* 1:e185727. doi: 10.1001/jamanetworkopen.2018.5727

Carlo, G., Koller, S., Raffaelli, M., and De Guzman, M. R. (2007). Culture-related strengths among Latin American families: a case study of Brazil. *Marriage Fam. Rev.* 41, 335–360. doi: 10.1300/J002v41n03_06

de Oliveira, N. (2017). *In Rio, 381 Schools Closed One or More Days this Year Because of Shootings*. Rio de Janeiro, Brasil: Agência Brasil.

Dessen, M. A., and Torres, C. V. (2002). Family and socialization factors in Brazil: an overview. *Online Readings Psychol. Cult.* 6, 304–325.

Durlak, J. A. (2016). Programme implementation in social and emotional learning: basic issues and research findings. *Camb. J. Educ.* 46, 333–345. doi: 10.1080/0305764X.2016.1142504

Durlak, J. A., and DuPre, E. P. (2008). Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am. J. Community Psychol.* 41, 327–350. doi: 10.1007/s10464-008-9165-0

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

Fundação Maria Cecilia Souto Vidigal, (2021). Annual report 2020. Available at: https://issuu.com/fmcsv/docs/fmcsv-annual-report-2020

Fonseca, P., and Alper, A. (2018). *Brazil Army Ordered to take Over Security in Violent Rio de Janeiro*. Rio de Janeiro, Brasil: Reuters.

Gultom, S., Endriani, D., and Harahap, A. S. (2022). Less emotion but more fatigue: social-emotional learning (SEL) competencies, and compassion fatigue among educators during the COVID-19 pandemic. *Kinestetik: J. Ilmiah Pendidikan Jasmani* 6, 146–158. doi: 10.33369/jk.v6i1.21034

Huang, K. Y., Nakigudde, J., Rhule, D., Gumikiriza-Onoria, J. L., Abura, G., Kolawole, B., et al. (2017). Transportability of an evidence-based early childhood intervention in a low-income African country: results of a cluster randomized controlled study. *Prev. Sci.* 18, 964–975. doi: 10.1007/s11121-017-0822-0

Kamøy, K., Podieh, P., and Salarkia, K. (2021). Stop the War on Children: A crisis of recruitment. *Save the Children International*, Washington, DC.

Katzman, N., and Stanton, M. (2020). The integration of social emotional learning and cultural education into online distance learning curricula: now imperative during the COVID-19 pandemic. *Creat. Educ.* 11, 1561–1571. doi: 10.4236/ce.2020.119114

Kieling, C., Baker-Henningham, H., Belfer, M., Conti, G., Ertem, I., Omigbodun, O., et al. (2011). Child and adolescent mental health worldwide: evidence for action. *Lancet* 378, 1515–1525. doi: 10.1016/S0140-6736(11)60827-1

Kliewer, W., Parrish, K. A., Taylor, K. W., Jackson, K., Walker, J. M., and Shivy, V. A. (2006). Socialization of coping with community violence: influences of caregiver coaching, modeling, and family context. *Child Dev.* 77, 605–623. doi: 10.1111/j.1467-8624.2006.00893.x

Lennert da Silva, A. L., and Mølstad, C. E. (2020). Teacher autonomy and teacher agency: a comparative study in Brazilian and Norwegian lower secondary education. *Curric. J.* 31, 115–131. doi: 10.1002/curj.3

Londoño, E. (2017). In Rio de Janeiro, 'complete vulnerability' as violence surges. *New York Times.*

Martinez, I., Garcia, F., Veiga, F., Garcia, O. F., Rodrigues, Y., and Serra, E. (2020). Parenting styles, internalization of values and self-esteem: a cross-cultural study in Spain, Portugal and Brazil. *Int. J. Environ. Res. Public Health* 17:2370. doi: 10.3390/ijerph17072370

McCoy, D. C., Hanno, E. C., Ponczek, V., Pinto, C., Fonseca, G., and Marchi, N. (2021). Um compasso Para aprender: a randomized trial of a social-emotional learning program in homicide-affected communities in Brazil. *Child Dev.* 92, 1951–1968. doi: 10.1111/cdev.13579

Movimento Pela Base Nacional Comum, (2018). *General Competencies of the National Learning Standards* São Paulo, Brasil: Movimento Pela Base Nacional Comum.

Murphy, K. M., Rodrigues, K., Costigan, J., and Annan, J. (2017). Raising children in conflict: An integrative model of parenting in war. *Peace Confl. J. Peace Psychol.* 23, 46–57. doi: 10.1037/pac0000195

Piza, C., Zwager, A., Ruzzante, M., Dantas, R., and Loureiro, A. (2020). *Supporting Teacher Autonomy to Improve Education Outcomes*. Washington, DC: World Bank.

Reimers, F., Schleicher, A., Saavedra, J., and Tuominen, S. (2020). Supporting the continuation of teaching and learning during the COVID-19 pandemic. *OECD* 1, 1–38.

Ştefan, C. A., and Miclea, M. (2013). Effects of a multifocused prevention program on preschool children's competencies and behavior problems. *Psychol. Sch.* 50, 382–402. doi: 10.1002/pits.21683

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

Tinajero, A. R., Cohen, N. J., and Ametorwo, S. (2016). No data, no problem, no action: parenting programs in low-income countries. Making the social–emotional outcomes more visible. *Child Care Health Dev.* 42, 117–124. doi: 10.1111/cch.12301

Torrente, C., Aber, J. L., Starkey, L., Johnston, B., Shivshanker, A., Weisenhorn, N., et al. (2019). Improving primary education in the Democratic Republic of the Congo: end-line results of a cluster-randomized wait-list controlled trial of learning in a healing classroom. *J. Res. Educ. Effect.* 12, 413–447. doi: 10.1080/19345747.2018.1561963

Tubbs Dolan, C., Kim, H. Y., Brown, L., Gjicali, K., Borsani, S., Houchaimi, S. E., et al. (2022). Supporting Syrian refugee children's academic and social-emotional learning in national education systems: a cluster randomized controlled trial of nonformal remedial support and mindfulness programs in Lebanon. *Am. Educ. Res. J.* 59, 419–460. doi: 10.3102/00028312211062911

UN Office on Drugs and Crime, (2021). *Intentional Homicides (per 100,000 people)*. The World Bank. Washington, D.C.

Vélez-Agosto, N. M., Soto-Crespo, J. G., Vizcarrondo-Oppenheimer, M., Vega-Molina, S., and García Coll, C. (2017). Bronfenbrenner's bioecological theory revision: moving culture from the macro into the micro. *Perspect. Psychol. Sci.* 12, 900–910. doi: 10.1177/1745691617704397

Wigelsworth, M., Lendrum, A., Oldfield, J., Scott, A., Ten Bokkel, I., Tate, K., et al. (2016). The impact of trial stage, developer involvement and international transferability on universal social and emotional learning programme outcomes: A meta-analysis. *Camb. J. Educ.* 46, 347–376. doi: 10.1080/0305764X.2016.1195791

Wolf, S. (2020). "Me I don't really discuss anything with them": parent and teacher perceptions of early childhood education and parent-teacher relationships in Ghana. *Int. J. Educ. Res.* 99:101525. doi: 10.1016/j.ijer.2019.101525

World Bank. (2019). *Population, ages 0-14, total*. World Bank.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 10 March 2023
DOI 10.3389/fpsyg.2023.1059138

OPEN ACCESS

EDITED BY
Stephanie M. Jones,
Harvard University,
United States

REVIEWED BY
Brianna Devlin,
University of Oregon,
United States
Benjamin Dreer,
University of Erfurt,
Germany
Joshua Brown,
Fordham University,
United States

*CORRESPONDENCE
Summer S. Braun
✉ ssbraun@ua.edu

SPECIALTY SECTION
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

RECEIVED 30 September 2022
ACCEPTED 10 February 2023
PUBLISHED 10 March 2023

CITATION
Braun SS, Bradshaw CP, Beahm LA, Budavari AC, Downer J, Ialongo NS and Tolan PH (2023) Predicting implementation of the PAX Good Behavior Game + MyTeachingPartner interventions. *Front. Psychol.* 14:1059138. doi: 10.3389/fpsyg.2023.1059138

Predicting implementation of the PAX Good Behavior Game + MyTeachingPartner interventions

Summer S. Braun[1*], Catherine P. Bradshaw[2], Lydia A. Beahm[2], Alexa C. Budavari[2], Jason Downer[2], Nicholas S. Ialongo[3] and Patrick H. Tolan[2]

[1]Department of Psychology and Center for Youth Development and Intervention, University of Alabama, Tuscaloosa, AL, United States, [2]School of Education and Human Development, University of Virginia, Charlottesville, VA, United States, [3]Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States

Introduction: Effective classroom management is critical to creating a classroom environment in which social, emotional, and academic learning can take place. The present study investigated the association between early career, early elementary teachers' occupational health (job stress, burnout, and perceived teaching ability) and perceptions of program feasibility in relation to their implementation dosage and quality of two evidence-based classroom management programs implemented together: the PAX Good Behavior Game (GBG) and MyTeachingPartner (MTP) intervention.

Methods: Teachers provided information on their occupational health at the start of the school year and were then randomized to the PAX GBG + MTP condition or control condition. Teachers' perceptions of the feasibility of the program, implementation dosage, and implementation quality of the intervention were measured at the end of the school year for the 94 intervention teachers.

Results: Teachers participated in more MTP coaching cycles when they reported that the combined PAX GBG + MTP program was feasible. Although there were no main effects of occupational health on implementation, the associations between job stress and implementation quality were moderated by perceptions of feasibility.

Discussion: Findings highlight the complexity of factors influencing the implementation of evidence-based programs in school settings.

KEYWORDS

implementation science, good behavior game, feasibility, teacher well-being, occupational health

1. Introduction

Social and emotional learning programs and classroom management interventions are known to improve children's social and emotional skills, attitudes, behavior, and academic performance (Durlak et al., 2011; Korpershoek et al., 2016). Yet, effective implementation of such evidence-based intervention programs is critical to producing these targeted outcomes (Kam et al., 2003; Durlak and DuPre, 2008; Domitrovich et al., 2010). Schools are inherently complex systems, and there is noted variation in the implementation of such programs within this context (Witt et al., 1997; Ringwalt et al., 2009; Sanetti and Kratochwill, 2009; Hicks et al., 2014; Sanetti and Collier-Meek, 2019). To advance the science of implementation, it is essential

to understand the factors influencing the adoption of evidence-based programs. The present study aimed to contribute to the practical understanding of implementation in the school context by exploring the effects of teachers' occupational health and their perceptions of feasibility of two evidenced-based interventions employed together: PAX Good Behavior Game (GBG) and MyTeachingPartner (MTP).

1.1. Importance of evidence-based programs for youth

Evidence-based programs are those which have undergone rigorous scientific testing and are found to have beneficial effects for the target population (What Works Clearinghouse, 2017). Evidence-based social and emotional learning programs are known to have widespread positive effects on children: improving social and emotional skills, attitudes towards self and others, positive social behavior, mental health, and academic performance, as well as preventing conduct problems and emotional distress (Durlak et al., 2011). Yet, effective classroom management is critical to creating a classroom environment where such learning can occur (Evertson and Weinstein, 2011). In fact, classroom management interventions themselves have been shown to benefit not only classroom behavior but also students' academic outcomes and social and emotional development (Korpershoek et al., 2016). Further, evidence-based social and emotional learning and classroom management interventions can have effects long after program participation, even in areas not directly targeted by the intervention (e.g., graduation; Taylor et al., 2017). Based on evidence that early-career teachers struggle specifically with classroom management (Kukla-Acevedo, 2009; Wei et al., 2010), which is foundational to student learning, this project combined two interventions with a strong evidence-base for improving classroom management: a classroom management program, the PAX GBG, and a teacher coaching intervention, MTP.

1.1.1. PAX GBG

The GBG is an evidence-based program originally developed by Barrish et al. (1969) that aims to promote teachers' classroom behavior management, increase on-task behaviors, and decrease disruptive behaviors (Bowman-Perrott et al., 2016; Smith et al., 2021). The GBG incorporates several behavior management strategies such as positive behavior praise, explicit instruction, feedback, and positive reinforcement, thus making it a high-quality practice for teachers to implement in their classrooms. The GBG is an interdependent group contingency that requires all students in the group to meet the requirements of the contingency as a group and individually (Embry, 2002). As such, students typically encourage their peers to meet expectations, thus reducing some of the demands on teachers (Hopman et al., 2018). GBG includes identifying the target behavior (e.g., completing the worksheet silently), posting the GBG expectations, dividing the class into equal teams, and awarding points to teams meeting expectations or removing points for infractions (Barrish et al., 1969; Embry, 2002). The team with the most points receives a non-tangible group reinforcement. The PAX GBG augmented the original version of GBG by integrating additional activities and components to improve compliance and classroom management (e.g., soliciting student input on classroom expectations; Embry, 2002). Previous research has demonstrated positive effects of the GBG and the PAX GBG for both students (e.g., reduced aggressive/disruptive behaviors and improved academic outcomes; Ialongo et al., 2019; Johansson et al., 2020; Smith et al., 2021; Leidig et al., 2022) and teachers (e.g., decreased emotional exhaustion; Hopman et al., 2018).

1.1.2. MTP

MTP is an evidence-based coaching intervention that targets effective classroom management through the quality of teachers' interactions with students (Allen et al., 2015). Throughout the program, coaches provide video-based, individualized feedback to teachers as they develop classroom management skills and the capacity to provide emotional and instructional support to students. Previous studies of MTP have demonstrated positive effects on the quality of student-teacher interactions, peer interactions, social functioning, behavioral engagement, and academic outcomes (Pianta et al., 2008b; Allen et al., 2011, 2015; Mikami et al., 2011; Gregory et al., 2014).

Given the evidence behind the PAX GBG and MTP, recent research has combined both programs in an effort to support the development of early-career teachers' classroom management skills and capacity for high-quality interactions with students (Tolan et al., 2020). In this combined approach, the PAX GBG and MTP work in tandem, incorporating unique and overlapping classroom management strategies that aim to improve teachers' interactions with students, their classroom management practices, and subsequent student outcomes. Early-career teachers are a particularly suitable population for these interventions because, in comparison to more experienced teachers, they are actively developing new habits, may be more open to feedback, and may need new classroom management skills, given that many pre-service preparation programs are known to provide insufficient training in this area (Freeman et al., 2014).

1.2. Importance of implementation

Importantly, effective implementation of evidence-based programs such as the PAX GBG and MTP is critical to their success. A meta-analysis of 542 studies of interventions for youth concluded that implementation had profound effects on outcomes; programs implemented well resulted in effect sizes several times higher than those with poorer implementation (Durlak and DuPre, 2008). Despite their potential convenience for improving public health, school-based interventions are often at risk for poor implementation (Domitrovich et al., 2008; Sanetti et al., 2014). The emergence of the field of implementation science has brought an explicit focus on understanding and in turn, addressing the barriers that jeopardize the effective implementation of such programs (Eccles and Mittman, 2006; Durlak, 2015).

1.3. Conceptualization of implementation

Implementation refers to the content of the program and how it is delivered in a specific setting (Durlak and DuPre, 2008). Durlak and Dupre (2008) describe eight dimensions of implementation: (1) Fidelity, also known as adherence or compliance, is the extent to which a program aligns with the originally intended curriculum; (2) Dosage is the amount of the original program that was provided,

often measured by the number of program sessions delivered; (3) Quality is how well, clearly, and correctly the program was delivered; (4) Participant responsiveness is the extent to which the program stimulates interest and garners the attention of the participants; (5) Program differentiation refers to the uniqueness of the program from other interventions; (6) Monitoring of comparison conditions is the documentation of the services received by those outside of the intervention group; with researchers primarily focused on the intervention condition, the control group often goes unmonitored, yet knowing the activities of both groups is important when drawing conclusions about the comparative effect of a program; (7) Program reach refers to the proportion of involvement of individuals in a population and the representativeness of program participants, which is particularly important when considering program scale-up; (8) Adaptation refers to the changes made to the program that result in differences between that implementation and the original intervention. Most school-based implementation research has focused on the first two dimensions – fidelity and dosage – little is known about the effects of these other dimensions on program outcomes (Gould et al., 2016). The present study expands the field of implementation science by examining whether teachers' occupational health and perceptions of program feasibility influence both the dosage and quality of implementation of the PAX GBG and MTP.

1.4. Predictors of implementation

Domitrovich et al. (2008) offered a multi-level conceptual framework outlining factors influencing the implementation of school-based interventions to guide implementation research in this context. Informed by ecological systems models (e.g., Bronfenbrenner, 1994), the Domitrovich et al. conceptual model posits that the implementation of school-based programs is impacted by influences specific to the context in which the program is being implemented. These influences are described in three main categories: individual-level factors relating to the program implementer (e.g., occupational health, perceptions of the program), school-level factors (e.g., school culture, resources), and macro-level factors (e.g., federal, state, and district policies). As most school-based intervention programs are implemented by classroom teachers instead of an external provider such as a clinician or school counselor (Forman et al., 2009), it is important to understand the role that teachers play in impacting implementation. As such, the present study focused on how the characteristics of the teachers implementing the PAX GBG + MTP may impact their dosage and quality of implementation of the programs. We focus on teachers' occupational health and their perceptions of the feasibility of the intervention due to their theoretical and empirical relevance to teachers' capacity to implement intervention programming, described in detail below.

1.4.1. Teachers' occupational health

Teachers' occupational health refers to their evaluations of various aspects of their job (van Horn et al., 2004). The multifaceted construct incorporates affective, cognitive, professional, social, and psychosomatic dimensions (van Horn et al., 2004). Notably, Kazdin (1993) posited that the absence of dysfunction does not reflect the presence of optimal functioning. Thus, it is important to consider both negative experiences of distress (e.g., stress) and positive experiences of well-being (e.g., perceived ability) in assessing occupational health. Indeed, researchers have used a variety of measures to assess this construct, including job stress, burnout, self-efficacy, and others (van Horn et al., 2004; Bakker and Rodríguez-Muñoz, 2010). Although research often considers teachers' occupational health as an outcome of an intervention (e.g., Ross et al., 2012) teachers' occupational health may also influence their classroom practice. Specifically, the conceptual model of the Prosocial Classroom (Jennings and Greenberg, 2009) posits that teachers' occupational health and well-being likely support effective implementation of intervention programs, while feelings of stress and associated experiences jeopardize implementation. We focus on three salient experiences of occupational health in the present study: teachers' experiences of job stress, burnout, and perceptions of teaching ability.

1.4.1.1. Job stress

Teachers report one of the highest levels of stress of any profession (Johnson et al., 2005). The Jobs Demands-Resource model of occupational stress posits that when the demands of a job exceed the resources of the individual and organization, stress can occur (Bakker and Demerouti, 2007). Theoretically, when teachers are under stress their emotional resources, attention, and cognitive energy are devoted to coping, leaving fewer resources for maintaining healthy relationships with students, supporting student development, and effectively implementing programs (Boekaerts, 1993; Roeser et al., 2012, 2021). Although empirical studies investigating these associations are still emerging, initial evidence provides support for this theory. Stressed teachers report more barriers to implementing evidence-based programs, such as lack of time to implement the program than their less-stressed counterparts (McGoey et al., 2014). Further, another study found a well-being intervention for teachers reduced stress and also improved their implementation of an evidence-based program for their students (Larson et al., 2018). Finally, one recent study found high levels of teacher stress to be associated with poor implementation quality of a mindfulness curriculum for students (Braun et al., 2023). However, this association was attenuated when teachers were provided with in-depth training, suggesting that highly stressed teachers may need more hands-on support than is typical to implement interventions well.

1.4.1.2. Burnout

Prolonged exposure to stressors (e.g., job demands exceeding resources) are often associated with experiences of burnout (Maslach and Jackson, 1981). Burnout is often characterized by three dimensions: emotional exhaustion, depersonalization (e.g., feeling disconnected), and personal accomplishment (e.g., feelings of competence in the classroom; Maslach and Jackson, 1981). Burnout is an unfortunately common experience for teachers (García-Carmona et al., 2019; Salmela-Aro et al., 2019). Teachers' feelings of burnout have been known to be associated with other salient experiences for themselves (e.g., depression, job dissatisfaction, disengagement; Leiter and Durup, 1994; Bakker and Schaufeli, 2000; Pines and Keinan, 2005), their classroom practices (e.g., poor student-teacher interactions, classroom management characterized by harsh discipline; Reinke et al., 2013) and student outcomes (e.g., impaired academic

achievement; Chang and Davis, 2009; Breeman et al., 2015; Herman et al., 2020).

One process through which burnout may have such effects on students is by reducing their capacity to effectively implement evidence-based intervention programs, as posited by the Prosocial Classroom model (Jennings and Greenberg, 2009). Indeed, research from the field of social and emotional learning has supported this conceptual model; teachers' feelings of burnout have been related to lower dosage implementation of several different school-based programs, including the PAX GBG (Ransford et al., 2009; Domitrovich et al., 2015; Swift et al., 2017). Although one study found no main effect of burnout, the effect of burnout on implementation dosage was moderated by teacher-coach alliance: burnout was associated with a lower dosage of the PAX GBG, specifically when teacher-coach alliance was low (Wehby et al., 2012). Notably, the negative effect of burnout was found in the PAX GBG dosage but not the quality of implementation (Domitrovich et al., 2015). The present study is the first to assess the role of burnout in the implementation of the combined PAX GBG + MTP interventions.

1.4.1.3. Perceived ability

Self-efficacy refers to teachers' belief in their capability to successfully accomplish a specific teaching task (Tschannen-Moran et al., 1998). Two primary domains comprise self-efficacy: (1) self-perception of teaching competence (i.e., a teacher's assessment of their own skills and knowledge), and (2) beliefs about the demands of a specific teaching task (e.g., a teacher's context-specific assessment of external resources and barriers). Teachers' sense of self-efficacy is associated with other indices of occupational health (e.g., burnout, job satisfaction; Brouwers and Tomic, 2000; Skaalvik and Skaalvik, 2010, 2014). In addition, teachers' self-efficacy is associated with the use of more supportive classroom management practices, higher quality interactions with students, and student achievement (Swars et al., 2006; Zee and Koomen, 2016). Of focus in the present study is teachers' perceived teaching ability. Perceived teaching ability captures the teachers' views toward their own abilities as a teacher, which may be likened to the self-perception of teaching competence domain of self-efficacy.

Although no research has investigated the association between teachers' perceived ability and intervention implementation specifically, previous studies have explored the association between overall self-efficacy and implementation of school-based interventions. Teachers' self-efficacy has been associated with both the quality (Rohrbach et al., 1993; Kallestad and Olweus, 2003) and dosage (Ransford et al., 2009; Clayback et al., 2022) of school-based interventions. Some studies have focused specifically on the association between self-efficacy for classroom management and implementation, with mixed results. One study of early childhood teachers' implementation of a SEL program found self-efficacy for classroom management predictive of dosage, but not the quality of implementation (Thierry et al., 2022). Perhaps most relevant to the present investigation is a study of the PAX GBG, which found self-efficacy for classroom management was unrelated to implementation dosage and quality (Domitrovich et al., 2015). Further research is needed to clarify whether distinct aspects of self-efficacy, such as perceived teaching ability (vs. efficacy for classroom management, etc.), are associated with implementation dosage and quality.

1.4.2. Perceptions of program feasibility

Perceptions of the feasibility of a program are a core component of the social validity of an intervention. Social validity refers to the extent to which an intervention is useable, valuable, and favorably viewed by interested parties (Kazdin, 1977; Wolf, 1978; Horner et al., 2005). Although teachers' ratings of the feasibility of a specific program may be averaged to reflect the feasibility of the program as a whole, teachers themselves may vary in the extent to which they personally find the program to be feasible to implement (Han and Weiss, 2005). This teacher-level variation in perceptions of feasibility has important implications for implementation: Conceptual understandings and empirical evidence indicate that teachers who have positive perceptions of an intervention attend more training sessions (dosage) and implement the program with higher fidelity (Han and Weiss, 2005; Clayback et al., 2022). With regard to the social validity of PAX GBG specifically, previous research has found that teachers who perceive the program more favorably implement the program with greater fidelity and quality (Wehby et al., 2012). The predictive utility of social validity in the context of the combined PAX GBG + MTP has not yet been assessed and is important to consider as the combined high-quality implementation of these programs could be profound. Further, as social validity is in response to the program itself, these perceptions are potentially malleable, and results could inform amendments to the program to maximize social validity if found to be an important predictor of implementation. To this end, the present study focused on teachers' perceptions of the feasibility of the combined PAX GBG + MTP program (i.e., how easy it was to use).

1.4.3. The potential moderating effect of perceptions of feasibility on the association between stress and implementation

Although stress is theorized to be a barrier to implementation, it is possible that the effects of these predictors of implementation are more complex; a selection of factors may work together to impact implementation (Jennings and Greenberg, 2009; Ransford et al., 2009). For example, Dreer et al. (2017) found teachers' perceptions of the program to moderate the association between teachers' readiness to engage in the program and their commitment to utilizing new skills, where teachers experienced the greatest commitment to utilizing new skills when they were both ready to engage with the program and had positive perceptions of the program. In the context of the present study, teachers' positive perceptions of the feasibility of the programs could serve to buffer against the negative effect of stress on implementation. Conceptually, teachers' perceptions of feasibility may motivate teachers' engagement with the program (Wehby et al., 2012), despite their stress and function as a protective factor to lessen the impact of stress on implementation. In contrast, high levels of stress and perceptions that the program is difficult to implement may indicate compounding risks for poor implementation. Yet, most research examining the predictors of teachers' implementation has focused on the main effects (e.g., Domitrovich et al., 2015). Although few studies have investigated these more complex associations, one such study did find that teachers with high levels of burnout and negative perceptions of a social and emotional learning program exhibited the lowest implementation dosage and quality (Ransford et al., 2009), suggesting that these associations may be more complex than initially proposed (Jennings and Greenberg, 2009). Thus, additional research probing these more complex effects will contribute to our understanding of how

combinations of factors may work together to impact implementation. Although we may hypothesize that perceptions of feasibility may buffer against the negative effects of burnout in the same way as it might for stress, as burnout emerges from experiences of chronic stress (Maslach and Jackson, 1981), stress, rather than burnout, is likely a more salient experience for early career teachers. Thus, of interest in this study was whether early career teachers' perceptions of feasibility may attenuate the negative effects of stress on implementation.

1.5. Present study

The present study aimed to expand our understanding of the role that teachers play in the implementation of evidence-based programs for youth by investigating predictors of implementation. Guided by the Domitrovich et al. (2015) conceptual model of implementation of school-based interventions and the Prosocial Classroom model (Jennings and Greenberg, 2009), the present study explored the association between teachers' occupational health and perceptions of feasibility in relation to their dosage and quality of implementation of the PAX GBG + MTP program. Specifically, we addressed the following two specific research questions: RQ_1 Do teachers' own occupational health and perceptions of the feasibility (i.e., ease of use) of the program impact their implementation of the PAX GBG + MTP? RQ_2) Is the effect of teachers' stress on their implementation of the PAX GBG + MTP moderated by their perceptions of the feasibility of the program? We hypothesized that low levels of job stress, low levels of burnout, high levels of perceived ability, and positive perceptions of feasibility would be associated with greater dosage and implementation quality. We also hypothesized that the negative association between stress and implementation would be weaker for teachers who had positive perceptions of feasibility.

2. Methods

2.1. Study design and recruitment

The present study draws from a longitudinal, teacher-level randomized controlled trial of the PAX GBG + MTP program. Early career teachers (≤ 3 years of teaching experience) hired by three participating public school districts in Kindergarten-3rd grade were identified by the districts. These teachers were recruited into the project by project staff during district-wide professional development events for early-career teachers held prior to the start of the 2013 school year. To limit heterogeneity in teaching demands, eligible teachers included those in early grades (Kindergarten-3rd Grade) and excluded Teach for America engaged teachers, given the variation in their educational backgrounds from typical teachers. Project staff conducted all recruitment sessions, either through attendance at new teacher training and orientation events or through individual or small group sessions. Participation was voluntary, and teachers provided written informed consent consistent with IRB procedures approved at the investigators' universities and school divisions. Participating teachers received an honorarium (e.g., gift cards) for their participation and completion of data collection activities. This recruitment and randomization procedure was repeated the following 2 years (i.e., the fall of 2014 and 2015, respectively), for a total of three cohorts.

Recruitment efforts resulted in 272 interested teachers, of which 236 teachers consented to participate. Of those, eight withdrew before randomization, 15 were ineligible due to being assigned to an ineligible classroom (i.e., not Kindergarten-3rd Grade, special education classroom, resource class), not being permitted to attend training (based on the principal's decision), having already been trained in the PAX GBG or MTP, leaving the participating districts, or leaving the teaching profession altogether. Of those eligible, 25 left the project prior to baseline data collection. The final sample in the RCT intent-to-treat analyses included 188 teachers (69% of initially interested teachers; 80% of those who consented) recruited from 72 schools (Median number of teachers per school = 2, Range = 1–13 teachers). Cohort 1 consisted of 56 teachers (30 control condition, 26 intervention condition) from 34 schools, Cohort 2 consisted of 51 (25 control condition, 26 intervention condition) teachers from 30 schools, and Cohort 3 consisted of 81 teachers (39 control condition, 42 intervention condition) from 36 schools. Note that the same school could be represented in multiple cohorts if they had new teachers in subsequent years, as was the case in several instances. See Tolan et al. (2020) for the full consort diagram. Attrition during the first study year was low, with 11% (10 from control, 11 from intervention) discontinuing participation before the Year 1 post-intervention timepoint.

2.2. Participants

Due to the present study's focus on implementation outcomes, only the 94 teachers randomized to the intervention condition were included in the analytic sample in this study. Randomization was effective as there were no significant differences in baseline demographics, occupational health, nor implementation outcomes between teachers in the intervention and control conditions (see Downer et al., 2023). The majority of teachers in the intervention condition were female (93%) and White (80%), with 1–3 years of teaching experience, and most teachers were in their first year of teaching (60%). Teachers were approximately evenly distributed across the Kindergarten-3rd Grade classes.

2.3. Procedure

Baseline (Time 1) data collection occurred in the fall of the school year as close to the beginning of school as possible, in October. Post-intervention (Time 2) data collection occurred 7 months later, in May, shortly before the end of the school year. At each timepoint, teachers completed an online survey, and trained observers conducted classroom observations. Following baseline data collection, teachers were randomly assigned (blocking on school and district) to the intervention or control conditions. School-level demographic data were obtained from the state department of education.

2.3.1. Classroom observation procedures

Observations were conducted in accordance with the protocol for the Classroom Assessment Scoring System (CLASS; Pianta et al., 2008a). At each timepoint, certified CLASS observers conducted 4–6, 15-min observation cycles. Observations were conducted over two separate days balancing observations in the morning and the

afternoon. Immediately following each observation cycle, observers stepped out of the classroom and completed the CLASS ratings. There were no calculations of inter-rater agreement following the CLASS training and certification process, but previous research has demonstrated high inter-rater agreement for the CLASS (79–94% within 1 point; Pianta et al., 2008b; Mantzicopoulos et al., 2018).

2.3.2. PAX GBG

The GBG allows teachers to utilize social learning principles within a team-based, game-like context to reduce aggressive, disruptive, and off-task behavior and facilitate academic instruction. The current project used the PAX GBG, an augmented version of GBG which integrates ancillary components known to improve compliance and classroom management (Embry, 2002).

Prior to the implementation of the PAX GBG at the beginning of the school year, the teachers and students collaborated to define their vision of a "PAX" (Latin, meaning peaceful or ideal) classroom. Toward that end, they identified the behaviors that were necessary for creating a focused, productive, and peaceful classroom. During this collaboration, the teacher explained to the students that the positive behaviors they listed were referred to as "PAX" behaviors, and the negative behaviors were referred to as "spleems." After jointly defining PAX and spleems, teachers assigned students to one of three or four teams. The teams worked cooperatively to maintain PAX behavior in the classroom. Teachers gave points to the team when a member displayed a spleem. Teachers were trained to respond unemotionally to rule-breaking and when marking points against a child's team. At the end of the game period, all teams with three or fewer spleems won the game. The students were rewarded for displaying self-control, emotion regulation, and group regulation while not attending to or reinforcing the misbehavior of others. The team-based nature of the game allowed teachers to take advantage of positive peer pressure to improve academic and pro-social student behavior at the individual as well as at the classroom level.

2.3.2.1. Training in the PAX GBG

Professional development was provided to the intervention teachers over the course of one weekend day, during which teachers received intensive training and practice using the PAX GBG approach aligned with the MTP framework. Seventy-seven percent of the intervention teachers completed this training, with the remainder attending a small group or one-on-one make-up trainings. Teachers were asked to play approximately three PAX GBG games each school day with increasing length and in increasingly varied settings over the course of the year.

2.3.3. MTP

MTP is grounded in an evidence-based framework for thinking about teacher-student interactions that contribute to student behavior and achievement, called Teaching through Interactions (Hamre et al., 2013). This framework emphasizes that interactions should be emotionally supportive, well-organized, and cognitively enriching. The Teaching through Interactions framework is based on these three core domains of classroom interactions as captured by an observational approach called the Classroom Assessment Scoring System (CLASS): emotional support, instructional support, and classroom organization (Pianta et al., 2008a). During the intervention, the CLASS is used as a lens for viewing and providing feedback on a teacher's practice in the classroom.

2.3.3.1. Training in MTP

Teachers in the intervention condition also participated in 1 day of training in MTP. Following the training, teachers participated in biweekly MTP coaching cycles throughout the training year, with initial contact in-person that then shifted to web-mediated training. See Tolan et al. (2020) for a detailed description of coaching steps. Over the course of the school year, these coaching cycles focused on all three CLASS domains as well as elements of the PAX GBG that would help teachers optimize their implementation of the PAX GBG by attending to their interactions with students. The coaching cycles were intended to be collaborative, supportive, constructive, and to help teachers develop CLASS and PAX GBG knowledge, improve observation skills, develop analysis skills, feel supported in these endeavors, and increase their sense of agency and efficacy in the classroom. Teachers were asked to incorporate new strategies related to both the CLASS domains and PAX GBG elements into their teaching practice to improve both their implementation of the PAX GBG and their overall teaching practice.

2.4. Measures

2.4.1. Implementation outcomes

The present study reports implementation dosage and quality data for the PAX GBG and MTP collected at post-intervention.

2.4.1.1. Implementation dosage

The number of *coaching cycles completed* is an indicator of the dosage of MTP teachers received. These data were collected on each participating teacher through the online MTP coaching platform. The number of *PAX GBG games played* is an indicator of the dosage of the PAX GBG the students received. For the school year following the training, teachers self-reported on the number of PAX GBG games played each week, which was averaged across the school year.

2.4.1.2. Implementation quality

Following the CLASS procedure above, after each classroom observation period, observers provided a rating from 1 to 7 (1 = Low, 7 = High) for each of the 11 CLASS dimensions. In accordance with contemporary uses of the CLASS (e.g., Mashburn et al., 2008), three scales were created to reflect the core CLASS domains: classroom management, emotional support, and instructional support. Teachers' scores within each timepoint (baseline, post-intervention) were averaged across cycles of observation. Cronbach's α for all domains and timepoints were acceptable (ranging from 0.84–0.92). As MTP aims to improve the quality of teachers' interactions with students using the CLASS as a guide to anchor coaches' feedback to teachers, the CLASS was used as an indicator of the quality of teachers' implementation of MTP.

2.4.1.2.1. Emotional support

Emotional Support was calculated as the average of the positive climate, negative climate, teacher sensitivity, and regard for student perspectives dimensions.

2.4.1.2.2. Instructional support

Instructional support reflects teachers' facilitation of academic learning, measured as the average of the quality of feedback, concept development, and language modeling.

2.4.1.2.3. Classroom organization

Classroom organization, also referred to as classroom management, assesses the quality of teachers' interactions with students while they are managing the students in the room. It is the average of the behavior management, productivity, and instructional learning formats dimensions.

2.4.2. Occupational health and perceptions of feasibility

Main predictors of interest included indices of teachers' occupational health collected at baseline, namely job stress, burnout, and perceived ability, along with teachers' perceptions of intervention feasibility collected post-intervention.

2.4.2.1. Job stress

Teachers' experiences of job stress were assessed using items from the National Institute for Occupational Safety and Health survey of work-related stress (National Institute for Occupational Safety and Health, 1999). Teachers rated five items about their current feelings of stress (e.g., "In my job, I feel like I am under great stress") on a 1 to 4 scale (1 = Strongly Disagree; 4 = Strongly Agree; $\alpha = 0.82$).

2.4.2.2. Burnout

Teachers' experiences of burnout were assessed using four items from the emotional exhaustion subscale of the Maslach Burnout Inventory (Maslach et al., 1996). Teachers rated four items about their feelings of emotional exhaustion ("I feel burned out from my work," "I feel like I am at the end of my rope," I feel emotionally drained from my work," and "I feel used up at the end of the work day") on a 1 to 4 scale (1 = Strongly Disagree; 4 = Strongly Agree; $\alpha = 0.85$). This subscale was abbreviated for use in this study due to practical considerations to reduce participant burden.

2.4.2.3. Perceived ability

Teachers' perceptions of their ability as a teacher were assessed using the Perceived Ability subscale of the Factors Influencing Teaching Choice (FIT-Choice) measure (Watt and Richardson, 2007). Teachers rated three items about their perceived ability ("I have the qualities of a good teacher," "I have good teaching skills," and "Teaching is a career suited to my abilities") on a 1 to 4 scale (1 = Strongly Disagree; 4 = Strongly Agree; $\alpha = 0.69$).

2.4.2.4. Perceptions of feasibility

Perceptions of the feasibility of the combined PAX GBG + MTP intervention were assessed using items from the Teacher Perceptions of the Intervention Attributes scale (Domitrovich et al., 2015) with adapted wording to be relevant to the PAX GBG and MTP programs. Teachers rated five items assessing their perceptions of how feasible the combined program was to implement (e.g., "The GBG + MTP coaching process was easy to participate in") on a 1 to 4 scale (1 = Strongly Disagree; 4 = Strongly Agree; $\alpha = 0.72$).

2.4.3. Demographics

At the teacher-level, teachers self-reported the grade that they taught and their years of teaching experience. School-level demographic data regarding the enrollment of the school and percent of students eligible for free and reduced-priced meals (FARMS) were obtained from the state department of education.

2.5. Analytic plan

2.5.1. Missing data

Full information maximum likelihood (FIML) was used to incorporate all participants with baseline data (including those who did not provide data at post-intervention) into the intent-to-treat analyses. This approach accounts for the missing data while obtaining minimally biased estimates (e.g., Little et al., 2014; Witkiewitz et al., 2014).

2.5.2. Preliminary analyses

All analyses were run in R studio. Preliminary analyses included descriptive statistics and bivariate correlations among study measures.

2.5.3. RQ1: The association between occupational health, perceptions of feasibility, and implementation

Multiple linear regression models with cluster robust standard errors were employed to test the association between teachers' occupational health (i.e., stress, burnout, perceived ability) and perceptions of feasibility and their implementation of the PAX GBG + MTP. Because FIML was invoked to account for missing data, models were run in the latent framework using the lavaan package (Rosseel, 2012). A separate model was run for each outcome. Cluster robust standard errors were used to account for the nesting of teachers in schools in these analyses, where multilevel modeling is not appropriate given the average cluster size was so small ($M = 2.44$ teachers/school). Models predicting the number of coaching cycles completed used estimator = "MLR" to account for the non-normal distribution of this count outcome. Measures of occupational health and perceptions of feasibility were grand mean centered. Models controlled for grade level (continuous, where 0 = Kindergarten) and years teaching (continuous, where 0 = 1st year), as the participants in this study ranged from Kindergarten-3rd Grade teachers and were in their 1st-3rd years of teaching. Models predicting implementation quality (CLASS outcomes) controlled for teachers' CLASS scores at baseline, which were grand mean centered. School-level covariates included the school enrollment and the percent of students eligible for FARMs, which were standardized, and grand mean centered, respectively. Centering in this way results in an intercept that can be interpreted as the predicted level of implementation for a teacher who is experiencing an average level of job stress, burnout, perceptions of their teaching ability, and perceptions of feasibility, and is a Kindergarten teacher in their 1st year in the classroom, who is in a school of average enrollment and eligibility for FARMs.

2.5.4. RQ2: The moderating effect of perceptions of feasibility on the association between stress and implementation

To test the potentially moderating role of perceptions of feasibility, the interaction of stress and perceptions of feasibility was added to each of the models above.

3. Results

3.1. Preliminary analyses

Descriptive statistics of measures are provided in Table 1. Teachers demonstrated a relatively high dosage of the MTP elements of the

combined program. On average, teachers completed 8.24 coaching cycles, which exceeded the target number of cycles of 8, which previous research has shown to impact teacher practice and student outcomes (Allen et al., 2011, 2015). Implementation of the PAX GBG elements of the combined program was also relatively high, with teachers playing an average of 9 games per week (i.e., an indicator of dosage).

Bivariate correlations among study measures are provided in Table 2. Notably, burnout was significantly negatively correlated with the number of MTP coaching cycles completed ($r = -0.21$, $p = 0.048$), whereas perceptions of feasibility were significantly positively correlated with the number of MTP coaching cycles completed ($r = 0.36$, $p = 0.001$). These correlations indicated that teachers who reported higher levels of burnout at baseline completed fewer coaching cycles than their peers who were more burned out, and that teachers who reported the program was more feasible to implement (i.e., easy to use) completed more coaching cycles than their peers who reported lower levels of program feasibility.

3.2. RQ1: The association between occupational health, perceptions of feasibility, and implementation

3.2.1. Coaching cycles completed

Teachers' perceptions of the feasibility of the program were associated with attending more coaching cycles ($B = 1.66$, $SE = 0.53$, $p = 0.002$). Grade was negatively associated with the number of coaching cycles completed ($B = -0.43$, $SE = 0.19$, $p = 0.03$; Table 3), such that teachers in lower grades attended more coaching sessions. No other effects were significant.

3.2.2. Number of games played

Occupational health and perceptions of feasibility were unrelated to the number of PAX GBG games played. At the school level, the percent of students eligible for FARMS was associated with playing more games ($B = 5.39$, $SE = 1.47$, $p < 0.001$). No other effects were significant.

3.2.3. Emotional support

Occupational health and perceptions of feasibility were unrelated to observations of teachers' emotional support. Teachers' emotional support at baseline was strongly associated with their emotional support at post-intervention ($B = 0.40$, $SE = 0.12$, $p < 0.001$). At the school level, the percent of students eligible for FARMS was marginally negatively associated with emotional support ($B = -0.66$, $SE = 0.39$, $p = 0.09$). No other effects were significant.

3.2.4. Instructional support

Teachers' instructional support at baseline was strongly associated with their emotional support at post-intervention ($B = 0.40$, $SE = 0.12$, $p < 0.001$). No other effects were significant.

3.2.5. Classroom organization

Occupational health and perceptions of feasibility were unrelated to observations of teachers' classroom organization. Grade level was marginally associated with classroom organization ($B = 0.34$, $SE = 0.09$, $p < 0.001$) such that teachers in higher grades were observed to have higher levels of classroom organization. Teachers' classroom organization at baseline was strongly associated with their classroom organization at post-intervention ($B = 0.34$, $SE = 0.09$, $p < 0.001$). At the school level, school enrollment ($B = -0.11$, $SE = 0.05$, $p = 0.03$) and the percent of students eligible for FARMS ($B = -0.74$, $SE = 0.34$, $p = 0.03$) were negatively associated with classroom organization such that the larger the school and the more students eligible for FARMS, the lower the observed classroom organization. No other effects were significant.

3.3. RQ2: The moderating effect of perceptions of feasibility on the association between stress and implementation

3.3.1. Coaching cycles completed

The interaction between stress and feasibility was not significant in predicting the number of coaching cycles completed ($B = 0.17$, $SE = 1.04$, $p = 0.87$; Table 4).

TABLE 1 Descriptive statistics.

	N	*Missing*	Mean	*SD*	Minimum	Maximum	Range
Implementation outcomes (Time 2)							
Implementation dosage							
Coaching cycles completed	86	8%	8.24	2.3	1	11	10
Number of games played	83	11%	9.16	2.81	2.11	15.43	13.32
Quality of implementation							
Emotional support	83	11%	4.88	0.68	2.39	6.25	3.86
Classroom organization	83	11%	5.04	0.71	3.38	6.28	2.90
Instructional support	83	11%	2.35	0.65	1.33	4.22	2.89
Occupational health (Time 1)							
Job stress	91	3%	2.63	0.61	1.00	3.80	2.80
Burnout	91	3%	2.71	0.69	1.00	4.00	3.00
Perceived ability	91	3%	3.34	0.44	2.00	4.00	2.00
Feasibility (Time 2)							
Easy to use	78	16%	3.16	0.56	1.60	4.00	2.40

TABLE 2 Bivariate correlations among study measures.

		Implementation outcomes					Occupational health			Feasibility	Demographics			
		Implementation dosage		Quality of implementation							Teacher level		School level	
		1	2	3	4	5	6	7	8	9	10	11	12	13
Implementation outcomes (Time 2)														
Implementation dosage														
1	Coaching Cycles Completed													
2	Games Played	**0.26**												
Quality of implementation														
3	Emotional support	0.19	−0.09											
4	Classroom organization	0.08	0.03	**0.52**										
5	Instructional support	0.05	−0.11	**0.75**	**0.46**									
Occupational health (Time 1)														
6	Job Stress	−0.17	−0.02	0.12	0.13	0.00								
7	Burnout	**−0.21**	0.00	0.06	0.17	−0.08	**0.84**							
8	Perceived Ability	0.10	0.03	−0.04	−0.06	−0.06	**−0.25**	**−0.29**						
Feasibility (Time 2)														
9	Feasibility	**0.36**	−0.02	−0.01	0.00	−0.06	0.05	−0.05	0.08					
Demographics														
Teacher level														
10	Grade	−0.15	−0.05	0.09	0.16	0.14	0.17	0.19	−0.12	0.20				
11	Years teaching	−0.16	0.05	0.13	0.01	0.03	0.19	0.15	−0.04	−0.12	−0.09			
School level														
12	Enrollment	−0.02	−0.10	0.13	−0.07	−0.10	−0.01	0.05	0.08	−0.16	0.04	0.09		
13	Eligible for FARMS	−0.10	**0.35**	**−0.25**	−0.14	**−0.31**	0.08	0.05	0.18	−0.13	**−0.21**	0.12	−0.15	

Bold indicates significant at $p < 0.05$. FARMS = Free and Reduced-Priced Meals.

3.3.2. Number of games played

The interaction between stress and feasibility was not significant in predicting the average number of PAX GBG games played each week ($B = 0.51$, $SE = 0.86$, $p = 0.55$).

3.3.3. Emotional support

The interaction between stress and feasibility was not significant in predicting observations of teachers' emotional support ($B = -0.33$, $SE = 0.27$, $p = 0.21$).

3.3.4. Instructional support

The interaction between stress and feasibility was significant in predicting observations of teachers' instructional support ($B = -0.77$, $SE = 0.26$, $p = 0.003$). This effect, visualized in Figure 1A, indicates that teachers who reported high levels of stress and lower levels of program feasibility (i.e., perceptions that the program was harder to use) implemented the program with higher quality than those who were highly stressed and reported the program was more feasible to implement (i.e., easy to use).

3.3.5. Classroom organization

The interaction between stress and feasibility was marginally significant in predicting observations of teachers' classroom organization ($B = -0.41$, $SE = 0.24$, $p = 0.09$). This effect, visualized in Figure 1B, shows that although job stress was relatively unrelated to implementation quality for those teachers who reported lower levels of program feasibility (i.e., perceptions that the program was harder to use), the opposite was true for teachers who reported that the interventions were feasible to implement. That is, teachers who had greater perceptions of program feasibility (i.e., perceptions that the program was easy to use) and experienced higher stress had poorer quality implementation than those who had greater perceptions of program feasibility and experienced low stress.

TABLE 3 Main effects models: predicting implementation dosage and quality of PAX GBG+MTP.

	Implementation dosage				Implementation quality					
	Coaching cycles completed		Number of games played		Emotional support		Instructional support		Classroom organization	
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Intercept	9.10*	0.41	9.13*	0.50	4.74*	0.12	2.31*	0.15	4.89*	0.11
Occupational health										
Job stress	−0.08	0.53	−0.38	0.86	0.13	0.21	0.07	0.28	0.18	0.22
Burnout	−0.35	0.55	0.16	0.87	0.00	0.17	0.16	0.22	−0.14	0.16
Perceived ability	−0.01	0.61	−0.29	0.70	0.09	0.19	0.04	0.15	−0.03	0.12
Feasibility										
Feasibility	1.66*	0.53	0.11	0.51	−0.07	0.13	−0.07	0.13	−0.20	0.15
Demographics and covariates										
Teacher level										
Grade	−0.43*	0.19	0.07	0.31	0.07	0.05	0.04	0.07	0.11+	0.06
Years teaching	−0.29	0.32	−0.08	0.34	0.09	0.08	0.00	0.09	0.00	0.09
CLASS (where appropriate)					0.40*	0.12	0.40*	0.12	0.34*	0.09
School level										
Enrollment	0.12	0.20	−0.16	0.34	0.04	0.06	−0.09	0.07	−0.11*	0.05
FARMS	−0.44	1.01	5.39*	1.47	−0.66+	0.39	−0.27	0.32	−0.74*	0.34
R Squared	0.24		0.15		0.29		0.17		0.30	

* indicates significant at $p < 0.05$, + indicates significant at $p < 0.10$. Cohort was omitted in final models because its inclusion did not substantively change the pattern of results. FARMS = Free and Reduced-Priced Meals.

TABLE 4 Moderation models: predicting implementation dosage and quality of PAX GBG+MTP.

	Implementation dosage				Implementation quality					
	Coaching cycles completed		Number of games played		Emotional support		Instructional support		Classroom organization	
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Intercept	9.12*	0.40	9.16*	0.50	4.72*	0.12	2.28*	0.15	4.86*	0.12
Occupational health										
Job stress	−0.05	0.53	−0.32	0.89	0.09	0.21	−0.01	0.27	0.13	0.22
Burnout	−0.35	0.56	0.14	0.88	0.02	0.17	0.18	0.22	−0.13	0.16
Perceived ability	−0.01	0.61	−0.28	0.70	0.08	0.19	0.03	0.16	−0.05	0.12
Feasibility										
Feasibility	1.64*	0.51	0.06	0.53	−0.04	0.12	0.00	0.13	−0.16	0.15
Interactions										
Stress*Feasibility	0.17	1.04	0.51	0.86	−0.33	0.27	−0.77*	0.26	−0.41+	0.24
Demographics and covariates										
Teacher level										
Grade	−0.44*	0.18	0.05	0.31	0.08	0.06	0.05	0.07	0.12*	0.06
Years teaching	−0.30	0.31	−0.11	0.34	0.11	0.08	0.03	0.09	0.02	0.09
CLASS (where appropriate)					0.40*	0.11	0.48*	0.11	0.35*	0.09
School level										
Enrollment	0.12	0.20	−0.16	0.35	0.03	0.06	−0.11+	0.07	−0.11*	0.05
FARMS	−0.47	1.02	5.32*	1.43	−0.61	0.40	−0.10	0.34	−0.67+	0.36
R Squared	0.24		0.15		0.30		0.26		0.33	

* indicates significant at $p < 0.05$, + indicates significant at $p < 0.10$. Cohort was omitted in final models because its inclusion did not substantively change the pattern of results. FARMS = Free and Reduced-Priced Meals.

4. Discussion

The present study investigated the associations among teachers' occupational health and perceptions of the feasibility of the combined PAX GBG+MTP program and two dimensions of implementation: dosage and quality. Contrary to our hypotheses, there were no main effects of any occupational health indicator on implementation dosage or quality. However, teachers who reported that the program elements were easier to use did complete more coaching cycles. In addition, the effect of teachers' job stress on two dimensions of implementation quality, instructional support and classroom organization, was moderated by teachers' perceptions of how feasible the program was, suggesting that the effect of job stress on implementation may be more nuanced than initially proposed (Jennings and Greenberg, 2009). The present study expands the field of implementation science in education research by: (1) investigating several potential predictors of implementation with a particular focus on teachers' occupational health and perceptions of program feasibility; (2) considering two aspects of implementation: dosage and quality; and (3) testing these effects within the context of the combination of two evidence-based programs (the PAX GBG and MTP) with strong potential for scaling up and which have seldom been combined or tested together. Given that early-career teachers may have fewer resources (e.g., training, on the job experience) compared to more experienced teachers, evidence-based programs for early-career teachers may be particularly useful resources for meeting the job demands they face.

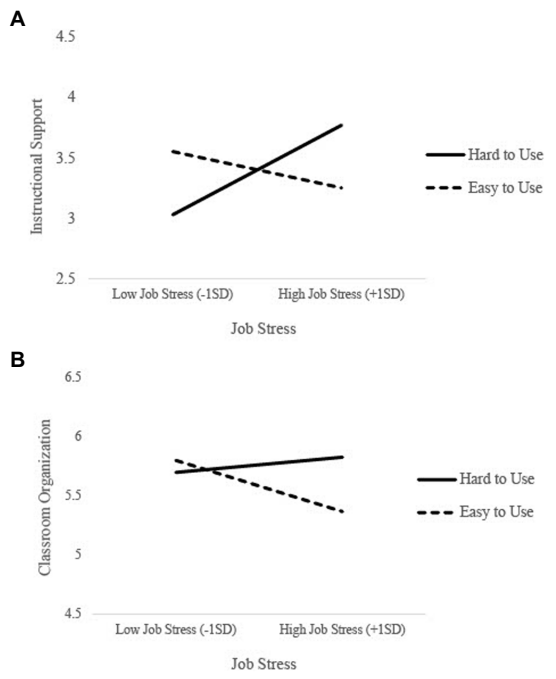


FIGURE 1
The effect of job stress on implementation quality was moderated by teachers' perception of feasibility of PAX GBG+MTP. Teachers experiencing high levels of stress and felt the program was harder to use demonstrated greater **(A)** instructional support, and **(B)** classroom organization than those who were highly stressed and reported the program was easy to use.

4.1. Teachers' occupational health did not predict implementation

Regarding RQ_1, although the bivariate correlations indicated that burnout at baseline was negatively associated with the number of coaching cycles completed, this association did not hold in the more complex models. It is worth noting that neither stress, burnout, nor perceived ability at baseline predicted implementation dosage or quality as assessed at post-program. Thus, the hypothesis that greater occupational health at baseline would be associated with greater implementation dosage and quality of the interventions was not supported. The lack of significant associations between occupational health and implementation found in the current study is mostly inconsistent with previous findings of school-based intervention programs, including the PAX GBG, which suggested that higher levels of occupational health were associated with greater implementation dosage and quality (Ransford et al., 2009; Wehby et al., 2012; Domitrovich et al., 2015). These previous studies have included qualitatively different populations than the current study, including teachers of up to 5th grade, and those beyond their first 3 years in the classroom. Yet, results are consistent with existing studies of the PAX GBG which found no association between self-efficacy for behavior management and implementation (Domitrovich et al., 2015), and a recent study of early childhood educators which found inconsistent associations between stress and implementation (Clayback et al., 2022).

The non-significant associations between stress, burnout, perceived ability, and implementation outcomes in this study should be interpreted in the context of several key considerations. First, occupational health may yet be related to implementation, just the present study's assessment of stress, burnout, and perceived ability may not be the relevant occupational health indicators that are important for implementation. For example, perceived ability, a specific aspect of self-efficacy measured in this study, may be too nuanced; with previous research indicating that general self-efficacy is associated with the quality and dosage of school-based interventions (Kallestad and Olweus, 2003; Ransford et al., 2009), it may be that general self-efficacy or specific self-efficacy around implementing new programs, rather than perceived teaching ability, may be related to implementation. Specifically, positive dimensions of occupational health such as job satisfaction and feelings of personal accomplishment are known to be salient experiences for teachers (e.g., Maslach et al., 2001; Hakanen et al., 2006). These indices of occupational health were not measured in the present study but may influence teachers' implementation of intervention programs. Future research assessing positive indices of occupational health (e.g., general self-efficacy, job satisfaction, etc.) will shed light on this possibility.

Moreover, the current project was focused solely on early career teachers. Experiences of burnout may be less salient than other measures of occupational health in this population teachers given it results from experiences of chronic stress, which these teachers may not have had time to experience yet. In addition, there may be other factors besides occupational health that exert a greater influence on implementation quality for early career teachers, such as administrative support or openness to interventions. Future research should explore other such teacher-specific factors that may influence implementation quality among early career teachers (Domitrovich et al., 2008). Further, because these teachers are still

actively developing their teaching practices, efforts to optimize implementation during this period may particularly impactful. For example, interventions addressing barriers to implementation could have a positive effect on these teachers' implementation of evidence-based programs across their career. It is also worth noting that the teachers in this study tended to report relatively high levels of occupational health. For example, teachers reported high levels of ability with limited variation ($M = 3.34$; $SD = 0.44$), potentially precluding the opportunity to detect significant differences across the spectrum of ability. Levels of burnout in this sample ($M = 2.71$; $SD = 0.69$) were also lower than those in other studies (e.g., Roeser et al., 2021), perhaps due to their early career status. Regardless, the limited variability in these measures of occupational health may have also limited their predictive utility; it may be that higher levels of stress and burnout are necessary in order to impair implementation. It is also possible that the indicators of occupational health measured in this study may have effects on other domains of implementation described by Durlak and DuPre (2008) that were not assessed in this study (e.g., participant responsiveness). Finally, the moderation of the effect of stress on implementation quality by perceptions of feasibility suggests that the association between occupational health and implementation may be more nuanced than direct effects, a finding which we explore further in the subsequent sections.

4.2. Teachers' perceptions of the feasibility of the PAX GBG+MTP program predicted MTP dosage

The hypothesis that greater perceptions of program feasibility would be associated with greater implementation was partially supported. Teachers' perceptions of how feasible the program was, an indicator of the social validity of the intervention, were not predictive of implementation quality, but were predictive of implementation dosage, assessed here as the number of coaching cycles completed. These results are consistent with existing evidence that positive perceptions of the program are associated with greater implementation (e.g., Wehby et al., 2012; Clayback et al., 2022). Implementation dosage is an important outcome to consider as existing literature has found dosage of MTP to be related to program outcomes (e.g., Pianta et al., 2014, 2022). This finding is informative for interventionists as it indicates that designing programs in ways that are simple to implement could be an effective strategy to increase sustained engagement in the program and subsequent targeted outcomes. Findings from successfully implemented school-based interventions have highlighted that school administration can be important champions for interventions (Forman et al., 2009). In this vein, it may be advantageous for school leadership to frame these programs as easy to use and easy to integrate into teaching, which could set the program up for success from the start (Forman et al., 2009).

Importantly, we measured perceptions of program feasibility alongside post-intervention assessments of implementation. This decision was made to not overburden participants with another survey in the middle of the school year. However, this design decision precludes definitive conclusions about the directionality of this association; yet, it is anticipated that teachers' perceptions of program feasibility preceded their implementation of the games and their attendance in coaching cycles. Finally, although we conceptualized teachers' perceptions of how easy the program elements were to use as an assessment of feasibility and social validity (Kazdin, 1977; Wolf, 1978), it could also be conceptualized as a component of implementation, namely participant responsiveness, a little studied dimension of implementation described in Durlak and DuPre (2008). The conceptualization of the construct underlies important design and analytic decisions, such as situating it as a predictor of implementation or an implementation outcome.

4.3. The association between stress and implementation quality was moderated by perceptions of feasibility

Regarding RQ_2, a significant interaction between teacher-reported stress at baseline and perceptions of feasibility at post-intervention emerged in models predicting both instructional support and classroom organization, indicating that the effect of stress on implementation quality differed according to perceptions of the feasibility of the program. Although our hypothesis regarding feasibility moderating the effect of stress on implementation was supported, the direction of these effects was contrary to our hypotheses. We hypothesized that the negative association between stress and implementation would be weaker for teachers who had positive perceptions of feasibility. Yet, results indicated that highly stressed teachers demonstrated greater instructional support and classroom organization when they found the program was *harder* to use compared to teachers who found it easy to use. For instructional support, we found that teachers reporting low levels of stress had higher implementation quality when they found the program easy to use compared to low-stress teachers who found it hard to use. The findings among low-stress teachers are consistent with previous literature, which has found that positive perceptions of social validity are associated with increased implementation (Wehby et al., 2012; McNeill, 2019). Contrary to previous research, the findings among highly stressed teachers may be capturing a particular subset of highly conscientious teachers who devoted more time and effort to learning and implementing the program, thus making it more difficult to use due to the high resource burden. It may also be the case that highly stressed teachers may have perceived the program as more valuable or useful due to its perceived complexity and difficulty, thus leading these teachers to implement increased instructional support and classroom management techniques. The complexity of these findings is aligned with recent evidence that the association between teachers' stress and the implementation of a mindfulness-based program for students was moderated by the amount of training they received (Braun et al., 2023). Together, these results suggest that there is more nuance to these associations than suggested in conceptual models, such as the Prosocial Classroom Model, in that the effect of stress on implementation may differ according to other teacher-, school-, and program-specific factors (Dreer et al., 2017; Jennings and Greenberg, 2009). Results should also be interpreted in light of the timing of these

measures, which is elaborated more in the Limitations and Future Research Directions section.

4.4. Teacher- and school-level demographics and implementation

Although not a main focus of these analyses, the effects of teacher- and school-level demographics included as covariates yielded findings also worth discussing. Teacher-level demographics of grade level and years of teaching experience were primarily unrelated to teachers' implementation dosage and quality. The exception was that teachers of lower grade levels completed more MTP coaching cycles. Teaching in the lower grade levels, particularly, is highly relational and high-quality interactions with students are as important as didactic instruction (Pianta and Stuhlman, 2004; Burchinal et al., 2008). Teachers of younger students could have been more motivated to attended MTP coaching because the relational content was particularly salient given the age of their students.

At the school level, teachers working in schools where more (vs. fewer) students were eligible for FARMS played a higher number of games, indicating increased implementation dosage in these schools. At the same time, implementation quality across all CLASS domains was lower for teachers in schools with more (vs. fewer) students eligible for FARMS. These results are consistent with existing research demonstrating that students experiencing the greatest socioeconomic need have teachers with lower-quality interactions (e.g., St Clair and Stone, 2016). Similarly, teachers in larger schools were observed to have lower levels of classroom organization. These findings indicate that although teachers in schools with high levels of FARMS may recognize the need for such interventions and employ more PAX GBG games than their peers from other schools, the quality of their implementation of MTP may be lower. These findings highlight that predictors of implementation dosage are not necessarily redundant with predictors of implementation quality, suggesting that researchers should continue to investigate dimensions of implementation as related yet separate outcomes. These teacher- and school-level findings could be useful in identifying teachers who may be at risk for a lower dosage of implementation and lower quality implementation of interventions.

4.5. Limitations and future research directions

There were certain limitations of the perceived ability measure used in the current analysis, evidenced by the relatively low internal consistency of the measure (i.e., $\alpha = 0.69$). This may be due to the fact that the items were drawn from a scale intended to measure individual's motivations for becoming a teacher, such that the items only capture the teaching competence domain of self-efficacy. Based on these findings, future research should incorporate measures that assess both domains of self-efficacy in order to capture both internal and external influences on teachers' perceived self-efficacy. Despite this limitation, the current findings demonstrate the importance of incorporating task- and context-specific measures of self-efficacy when examining factors that influence implementation quality.

An assumption underlying the interpretations of the feasibility findings is that perceptions of feasibility were stable across the course of the intervention. As feasibility was only assessed at post-intervention in this study, we were unable to test the variability nor directionality of these effects. Perceptions of the feasibility, or more generally, the social validity, of interventions could shift over the course of the program (Clayback et al., 2022). Future research should administer measures of social validity throughout the intervention in order to understand the potentially bidirectional influence between social validity and implementation dosage and quality, and what might predict more favorable changes in social validity over the course of the intervention.

In addition, the current findings should be interpreted within the context of early career, elementary school teachers since the identified associations with implementation quality and dosage may be specific to this population. Furthermore, these findings should be contextualized within the sociodemographic makeup of the sample, given that the sample was predominantly white (80%) and female (93%). Future research should build upon these findings in order to clarify whether similar factors influence implementation among teachers who teach middle and high school, have a greater number of years of experience, and are from more sociodemographically diverse backgrounds. Future research could employ enriched samples to improve racial/ethnic and gender identity diversity in order to capture the broader experience of all teachers. Further, despite the early career status of teachers in this study, participants reported slightly higher averages of burnout than stress. Future research could continue to explore whether other indices of occupational health (e.g., burnout) may also interact with perceptions of feasibility to impact teachers' implementation of evidence-based programs.

The present study provides some support for the Prosocial Classroom Model and model of factors impacting the implementation of school-based interventions (Domitrovich et al., 2008; Jennings and Greenberg, 2009). However, the interactions between occupational health and perceptions of feasibility found in this study also highlight that those models may be too simplified for the complexity of school-based research. Based on these findings, future research should continue to explore the multitude of program-, teacher- and school-level factors that influence the quality of intervention implementation among teachers with a range of experience and across varying intervention programs.

4.6. Implications for practice

Higher dosages of coaching cycles frequently lead to improved implementation fidelity and, ultimately, better student outcomes (Becker et al., 2013; Pas et al., 2022). Therefore, it is important for educators to be motivated to participate in coaching cycles. Results of this study indicate that teachers who perceived PAX GBG + MTP as feasible also participated in more coaching cycles. As such, efforts to increase perceptions of program feasibility may result in greater program dosage. One way to increase perceptions of feasibility is to ensure the program aligns with the school's core values (Forman et al., 2009). If teachers feel as though the program is a good "fit" to their own goals and philosophies, they are more likely to view the program in a positive way (Forman et al., 2009). Additionally, researchers may consider sharing findings regarding the positive perceptions of PAX GBG + MTP with teachers interested in implementing the program, as teachers respond well to learning new information from other teachers (Forman et al., 2009; Beahm et al., 2021).

Although the results provided no evidence that teachers' occupational health predicted their dosage and quality of implementation of the PAX GBG + MTP, we are cautious in our interpretation of these findings given that these associations have been found in previous

research (e.g., Ransford et al., 2009; Domitrovich et al., 2015). Regardless of its predictive utility for teachers' implementation of evidence-based programs, experiences of occupational health are salient and meaningful experiences for teachers. When indicating that teachers are suffering from poor occupational health, schools should be motivated to intervene not just because poor occupational health could impact teaching practices and implementation of evidence-based programs but also from a compassionate perspective to alleviate suffering.

The teacher and school demographics included as covariates in this study shed light on who and in what contexts implementation is notably high. Given that teachers of higher grades completed fewer coaching cycles, these teachers may be in need of greater support from coaching staff in order to increase engagement in the program. Although teachers in schools where a higher percentage of students were eligible for FARMS had greater implementation dosage, they simultaneously had lower implementation quality. These findings indicate that these teachers may be in need of additional support, potentially beyond the existing scope of the PAX GBG + MTP program, in order to reach high-quality implementation of PAX GBG + MTP. Taken together, future research should continue to explore teacher and school characteristics that influence both implementation quality and dosage in order to improve student and teacher outcomes.

5. Conclusion

The present study advances the field of implementation science in school-based research by investigating the association between teachers' occupational health and perceptions of program feasibility in relation to the dosage and quality of implementation of two evidence-based programs implemented together. Results provided some support for conceptual models of factors that influence the implementation of school-based interventions (Domitrovich et al., 2015), and highlight the complexity of optimizing implementation in this context. With the growing emphasis on the implementation of evidence-based programs in schools, efforts to scale-up such programs with fidelity should continue to attend to teacher- and school-level contextual factors. This study provides additional empirical evidence of particular characteristics that that may hinder implementation, while identifying potential factors such as program feasibility that may be important for facilitating the implementation of evidence-based programs.

Data availability statement

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by University of Virginia IRB. The patients/participants provided their written informed consent to participate in this study.

Author contributions

SB: conceptualization of the paper, formal analysis, and writing—original draft. CB, NI, PT, and JD: conceptualization of the project, funding acquisition, and writing—review and editing. LB and AB: writing—original draft. All authors contributed to the article and approved the submitted version.

Funding

The research reported here was supported by the Institute of Education Sciences, US Department of Education, through grants R305A190162 and R305A130107 (to the University of Virginia).

Acknowledgments

The authors would like to thank the coaches who supported implementation of this project and April Lawson for her assistance with project management.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Author disclaimer

The opinions expressed are those of the authors and do not represent views of the US Department of Education.

References

Allen, J. P., Hafen, C. A., Gregory, A. C., Mikami, A. Y., and Pianta, R. (2015). Enhancing secondary school instruction and student achievement: replication and extension of the my teaching partner-secondary intervention. *J. Res. Educ. Effect.* 8, 475–489. doi: 10.1080/19345747.2015.1017680

Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., and Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science* 333, 1034–1037. doi: 10.1126/science.1207998.An

Bakker, A. B., and Demerouti, E. (2007). The job demands-resources model: state of the art. *J. Manag. Psychol.* 22, 309–328. doi: 10.1108/02683940710733115

Bakker, A. B., and Derks, D. (2010). "Positive occupational health psychology," in *Occupational health psychology*, eds. S. Leka and J. Houdmont (Wiley Blackwell) 194–224.

Bakker, A. B., and Schaufeli, W. B. (2000). Burnout contagion processes among teachers. *J. Appl. Soc. Psychol.* 30, 2289–2308. doi: 10.1111/j.1559-1816.2000.tb02437.x

Barrish, H. H., Saunders, M., and Wolf, M. M. (1969). Good behavior game: effects of individual contingencies for group consequences on disruptive behavior in a classroom. *J. Appl. Behav. Anal.* 2, 119–124. doi: 10.1901/jaba.1969.2-119

Beahm, L. A., Yan, X., and Cook, B. G. (2021). Where do teachers go for behavior management strategies? *Educ. Treat. Child.* 44, 201–213. doi: 10.1007/s43494-021-00046-2

Becker, K. D., Bradshaw, C. P., Domitrovich, C., and Ialongo, N. S. (2013). Coaching teachers to improve implementation of the good behavior game. *Admin. Pol. Ment. Health* 40, 482–493. doi: 10.1007/s10488-013-0482-8

Boekaerts, M. (1993). Being concerned with well-being and with learning. *Educ. Psychol.* 28, 149–167. doi: 10.1207/s15326985ep2802_4

Bowman-Perrott, L., Burke, M. D., Zaini, S., Zhang, N., and Vannest, K. (2016). Promoting positive behavior using the good behavior game: A meta-analysis of single-case research. *J. Posit. Behav. Interv.* 18, 180–190. doi: 10.1177/1098300715592355

Braun, S. S., Greenberg, M. T., Roeser, R. W., Taylor, L. J., Montero-Marin, J., Crane, C., et al. (2023). Teachers' stress, training, and implementation of a school-based mindfulness program: Results from a cluster randomized controlled trial. Manuscript submitted for publication.

Breeman, L. D., Wubbels, T., Lier, P A C.Van, Verhulst, F. C., Ende, J.Van, Der, Maras A., et al. (2015). Teacher characteristics, social classroom relationships, and children's social, emotional, and behavioral classroom adjustment in special education. *J. Sch. Psychol.*, 53, 87–103. doi: 10.1016/j.jsp.2014.11.005

Bronfenbrenner, U. (1994). "Ecological models of human development" in *International encyclopedia of education*, vol. *3. 2nd* ed (Netherland: Elsevier), 37–43.

Brouwers, H., and Tomic, W. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teach. Teach. Educ.* 16, 239–253. doi: 10.1016/S0742-051X(99)00057-8

Burchinal, M., Howes, C., Pianta, R., Bryant, D., Early, D., Clifford, R., et al. (2008). Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher-child interactions and instruction. *Appl. Dev. Sci.* 12, 140–153. doi: 10.1080/10888690802199418

Chang, M. L., and Davis, H. A. (2009). *Understanding the role of teacher appraisals in shaping the dynamics of their relationships with students: deconstructing teachers' judgments of disruptive behavior/students*. New York, NY: Advances in teacher emotion research: The Impact on Teachers' Lives.

Clayback, K. A., Williford, A. P., and Vitiello, V. E. (2022). Identifying teacher beliefs and experiences associated with curriculum implementation fidelity in early childhood education. *Prev. Sci.* 24, 27–38. doi: 10.1007/s11121-022-01414-z

Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Buckley, J. A., Olin, S., Romanelli, L. H., et al. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: A conceptual framework. *Adv. School Ment. Health Promot.* 1, 6–28. doi: 10.1080/1754730X.2008.9715730

Domitrovich, C. E., Gest, S. D., Jones, D., Gill, S., and Derousie, R. M. S. (2010). Implementation quality: lessons learned in the context of the head start REDI trial. *Early Child. Res. Q.* 25, 284–298. doi: 10.1016/j.ecresq.2010.04.001

Domitrovich, C. E., Pas, E. T., Bradshaw, C. P., Becker, K. D., Keperling, J. P., Embry, D. D., et al. (2015). Individual and school organizational factors that influence implementation of the PAX good behavior game intervention. *Prev. Sci.* 16, 1064–1074. doi: 10.1007/s11121-015-0557-8

Downer, J. T., Braun, S. S., Bradshaw, C. P., Elreda, L., Ialongo, N. S., Tolan, P. H., et al. (2023). Testing the combined effects of the PAX Good Behavior Game and MyTeachingPartner coaching for early career teachers: Impacts on teacher-student interaction quality and teachers' occupational health. Manuscript submitted for publication.

Dreer, B., Dietrich, J., and Kracke, B. (2017). From in-service teacher development to school improvement: factors of learning transfer in teacher education. *Teach. Dev.* 21, 208–224. doi: 10.1080/13664530.2016.1224774

Durlak, J. A. (2015). "What everyone should know about program implementation" in *Handbook of social and emotional learning*. eds. J. A. Durlak, C. E. Domitrovich, R. P. Weissberg and T. P. Gullotta (United States: Guilford), 395–405.

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

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

Eccles, M. P., and Mittman, B. S. (2006). Welcome to implementation science. *Implement. Sci.* 1, 1–3. doi: 10.1186/1748-5908-1-1

Embry, D. D. (2002). The good behavior game: A best practice candidate as a universal behavioral vaccine. *Clin. Child. Fam. Psychol. Rev.* 5, 273–297. doi: 10.1023/A:1020977107086

Evertson, C. M., and Weinstein, C. S. (Eds.). (2011). *Handbook of classroom management: Research, practice, and contemporary issues*. UK: Routledge.

Forman, S. G., Olin, Æ. S. S., Hoagwood, K. E., Crowe, M., and Saka, N. (2009). Evidence-based interventions in schools: developers' views of implementation barriers and facilitators. *Sch. Ment. Heal.* 1, 26–36. doi: 10.1007/s12310-008-9002-5

Freeman, J., Simonsen, B., Briere, D. E., and MacSuga-Gage, A. S. (2014). Preservice teacher training in classroom management: A review of state accreditation policy and teacher preparation programs. *Teach. Educ. Spec. Educ.* 37, 106–120. doi: 10.1177/0888406413507002

García-Carmona, M., Marín, M. D., and Aguayo, R. (2019). Burnout syndrome in secondary school teachers: A systematic review and meta-analysis. *Soc. Psychol. Educ.* 22, 189–208. doi: 10.1007/s11218-018-9471-9

Gould, L. F., Dariotis, J. K., Greenberg, M. T., and Mendelson, T. (2016). Assessing fidelity of implementation (FOI) for school-based mindfulness and yoga interventions: A systematic review. *Mindfulness* 7, 5–33. doi: 10.1007/s12671-015-0395-6

Gregory, A., Allen, J. P., Kimami, A., Hafen, C. A., and Pianta, R. C. (2014). Effects of a professional development program on behavioral engagement of students in middle and high school. *Psychol. Sch.* 51, 143–163. doi: 10.1002/pits

Hakanen, J. J., Bakker, A. B., and Schaufeli, W. B. (2006). Burnout and work engagement among teachers. *J. Sch. Psychol.* 43, 495–513. doi: 10.1016/j.jsp.2005.11.001

Hamre, B. K., Pianta, R. C., Downer, J. T., DeCoster, J., Mashburn, A. J., Jones, S. M., et al. (2013). Teaching through interactions: Testing a developmental framework of teacher effectiveness in over 4,000 classrooms. *Elementary School J.* 113, 461–487. doi: 10.1086/669616

Han, S. S., and Weiss, B. (2005). Sustainability of teacher implementation of school-based mental health programs. *J. Abnorm. Child Psychol.* 33, 665–679. doi: 10.1007/s10802-005-7646-2

Herman, K. C., Prewett, S. L., Eddy, C. L., Savala, A., and Reinke, W. M. (2020). Profiles of middle school teacher stress and coping: concurrent and prospective correlates. *J. Sch. Psychol.* 78, 54–68. doi: 10.1016/j.jsp.2019.11.003

Hicks, T. B., Shahidullah, J. D., Carlson, J. S., and Palejwala, M. H. (2014). Nationally certified school psychologists' use and reported barriers to using evidence-based interventions in sc. *Sch. Psychol. Q.* 29, 469–487. doi: 10.1037/spq0000059

Hopman, J. A. B., Van, Lier P. A. C., Ende, J.Van, Der, Struiksma, Wubbels, T., Verhulst, F. C., et al. (2018). Impact of the good behavior game on special education teachers. *Teach. Teach. Theory Pract.*, 24, 350–368. doi: 10.1080/13540602.2017.1379389

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., and Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Except. Child.* 71, 165–179. doi: 10.1177/001440290507100203

Ialongo, N. S., Domitrovich, C., Embry, D., Greenberg, M. T., Lawson, A., Becker, K. D., et al. (2019). A randomized controlled trial of the combination of two school-based universal prevention programs. *Dev. Psychol.* 55, 1313–1325. doi: 10.1037/dev0000715

Jennings, P. A., and Greenberg, M. T. (2009). The prosocial classroom: teacher social and emotional competence in relation to student and classroom outcomes. *Rev. Educ. Res.* 79, 491–525. doi: 10.3102/0034654308325693

Johansson, M., Biglan, A., and Embry, D. (2020). The PAX good behavior game: one model for evolving a more nurturing society. *Clin. Child. Fam. Psychol. Rev.* 23, 462–482. doi: 10.1007/s10567-020-00323-3

Johnson, S., Cooper, C., Cartwright, S., Donald, I., Taylor, P. J., and Millet, C. (2005). The experience of work-related stress across occupations. *J. Manag. Psychol.* 20, 178–187. doi: 10.1108/02683940510579803

Kallestad, J. H., and Olweus, D. (2003). Predicting teachers' and schools' implementation of the Olweus bullying prevention program: A multi-level study. *Prev. Treat.* 6. doi: 10.1037/1522-3736.6.1.621a

Kam, C.-M., Greenberg, M. T., and Walls, C. T. (2003). Examining the role of implementation quality in school-based prevention using the PATHS curriculum. *Prev. Sci.* 4, 55–63. doi: 10.1023/A:1021786811186

Kazdin, A. E. (1977). Assessing the clinical or applied importance of behavior change through social validation. *Behav. Modif.* 1, 427–452. doi: 10.1177/014544557714001

Kazdin, A. E. (1993). Adolescent mental health. *Am. Psychol.* 48, 127–141. doi: 10.1037/0003-066X.48.2.127

Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., and Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Rev. Educ. Res.* 86, 643–680. doi: 10.3102/0034654315626799

Kukla-Acevedo, S. (2009). Leavers, movers, and stayers: the role of workplace conditions in teacher mobility decisions. *J. Educ. Res.* 102, 443–452. doi: 10.3200/JOER.102.6.443-452

Larson, M., Cook, C. R., Fiat, A., and Lyon, A. R. (2018). Stressed teachers don't make good implementers: examining the interplay between stress reduction and intervention fidelity. *Sch. Ment. Heal.* 10, 61–76. doi: 10.1007/s12310-018-9250-y

Leidig, T., Casale, G., Wilbert, J., Hennemann, T., Volpe, R. J., Briesch, A., et al. (2022). Individual, generalized, and moderated effects of the good behavior game on at-risk primary school students: A multi-level multiple baseline study using behavioral progress monitoring. *Front. Educ.*, 1–15. doi: 10.3389/feduc.2022.917138

Leiter, M. P., and Durup, J. (1994). The discriminant validity of burnout and depression: A confirmatory factor analytic study. *Anxiety Stress Coping* 7, 357–373. doi: 10.1080/10615809408249357

Little, T. D., Jorgensen, T. D., Lang, K. M., and Moore, E. W. G. (2014). On the joys of missing data. *J. Pediatr. Psychol.* 39, 151–162. doi: 10.1093/jpepsy/jst048

Mantzicopoulos, P., French, B. F., Patrick, H., Watson, J. S., and Ahn, I. (2018). The stability of kindergarten teachers' effectiveness: A generalizability study comparing the framework for teaching and the classroom assessment scoring system. *Educ. Assess.* 23, 24–46. doi: 10.1080/10627197.2017.1408407

Maslach, C., and Jackson, S. (1981). The measurement of experienced burnout. *J. Occup. Behav.* 2, 99–113. doi: 10.1002/job.4030020205

Maslach, C., Jackson, S. E., and Leiter, M. P. (1996). *Maslach burnout inventory manual*.

Maslach, C., Schaufeli, W. B., and Leiter, M. P. (2001). Job burnout. *Annu. Rev. Psychol.* 52, 397–422. doi: 10.1146/annurev.psych.52.1.397

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., et al. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, 79, 732–749.

McGoey, K. E., Rispoli, K. M., Venesky, L. G., Schaffner, K. F., McGuirk, L., and Marshall, S. (2014). A preliminary investigation into teacher perceptions of the barriers to behavior intervention implementation. *J. Appl. Sch. Psychol.* 30, 375–390. doi: 10.1080/15377903.2014.950441

McNeill, J. (2019). Social validity and teachers' use of evidence-based practices for autism. *J. Autism Dev. Disord.* 49, 4585–4594. doi: 10.1007/s10803-019-04190-y

Mikami, A. Y., Gregory, A., Allen, J. P., Pianta, R. C., and Lun, J. (2011). Effects of a teacher professional development intervention on peer relationships in secondary classrooms. *Sch. Psychol. Rev.* 40, 367–385. doi: 10.1080/02796015.2011.12087704

National Institute for Occupational Safety and Health. (1999). *Stress at work*.

Pas, E. T., Borden, L., Debnam, K. J., Lucia, D.De, and Bradshaw, C. P. (2022). Exploring profiles of coaches' fidelity to double Check's motivational interviewing-embedded coaching: outcomes associated with fidelity. *J. Sch. Psychol.*, 92, 285–298. doi: 10.1016/j.jsp.2022.04.003

Pianta, R. C., Decoster, J., Cabell, S., Burchinal, M., Hamre, B. K., Downer, J. T. et al. (2014). Dose-response relations between preschool teachers' exposure to components of professional development and increases in quality of their interactions with children. *Early Child. Res. Q.* 29, 499–508. doi: 10.1016/j.ecresq.2014.06.001

Pianta, R. C., La Paro, K. M., and Hamre, B. (2008a). *Classroom assessment scoring system: Manual K-3*. United States: Paul H Brookes Publishing.

Pianta, R. C., Lipscomb, D., and Ruzek, E. (2022). Indirect effects of coaching on pre-K students' engagement and literacy skill as a function of improved teacher-student interaction. *J. Sch. Psychol.* 91, 65–80. doi: 10.1016/j.jsp.2021.12.003

Pianta, R. C., Mashburn, A. J., Downer, J. T., Hamre, B. K., and Justice, L. (2008b). Effects of web-mediated professional development resources on teacher-child interactions in pre-kindergarten classrooms. *Early Child. Res. Q.* 23, 431–451. doi: 10.1016/j.ecresq.2008.02.001

Pianta, R. C., and Stuhlman, M. W. (2004). Teacher-child relationships and children's success in the first years of school. *Sch. Psychol. Rev.* 33, 444–458. doi: 10.1080/02796015.2004.12086261

Pines, A. M., and Keinan, G. (2005). Stress and burnout: the significant difference. *Personality and Individual Difference* 39, 625–635. doi: 10.1016/j.paid.2005.02.009

Ransford, C. R., Greenberg, M. T., Domitrovich, C. E., Small, M., and Jacobson, L. (2009). The role of teachers' psychological experiences and perceptions of curriculum supports on the implementation of a social and emotional learning curriculum. *Sch. Psychol. Rev.* 38, 510–532.

Reinke, W. M., Herman, K. C., and Stormont, M. (2013). Classroom-level positive behavior supports in schools implementing SW-PBIS: identifying areas for enhancement. *J. Posit. Behav. Interv.* 15, 39–50. doi: 10.1177/1098300712459079

Ringwalt, C., Vincus, A. A., Hanley, S., Ennett, S. T., Bowling, J. M., and Rohrbach, L. A. (2009). The prevalence of evidence-based drug use prevention curricula in U.S. middle schools in 2005. *Prev. Sci.* 10, 33–40. doi: 10.1007/s11121-008-0112-y

Roeser, R. W., Mashburn, A. J., Skinner, E. A., Taylor, C., Choles, J. R., Rickert, N. P., et al. (2021). Mindfulness training improves middle school teachers' occupational health, well-being, and interactions with students in their most stressful classrooms. *J. Educ. Psychol.* 114, 408–425. doi: 10.1037/edu0000675

Roeser, R. W., Skinner, E., Beers, J., and Jennings, P. A. (2012). Mindfulness training and teachers' professional development: an emerging area of research and practice. *Child Dev. Perspect.* 6, 167–173. doi: 10.1111/j.1750-8606.2012.00238.x

Rohrbach, L. A., Graham, J. W., and Hansen, W. B. (1993). Diffusion of a school-based substance abuse prevention program: predictors of program implementation. *Prev. Med.* 22, 237–260. doi: 10.1006/pmed.1993.1020

Ross, S. W., Romer, N., and Horner, R. H. (2012). Teacher well-being and the implementation of school-wide positive behavior interventions and supports. *J. Posit. Behav. Interv.* 14, 118–128. doi: 10.1177/1098300711413820

Rosseel, Y. (2012). Lavaan: an R package for structural equation modeling. *J. Stat. Softw.* 48, 1–36. doi: 10.18637/jss.v048.i02

Salmela-Aro, K., Hietajärvi, L., and Lonka, K. (2019). Work burnout and engagement profiles among teachers. *Front. Psychol.* 10, 1–8. doi: 10.3389/fpsyg.2019.02254

Sanetti, L., and Collier-Meek, M. (2019). Increasing implementation science literacy to address the research-to-practice gap in school psychology. *School Psychol. Res.* 76, 33–47. doi: 10.1016/j.jsp.2019.07.008

Sanetti, L. M., Hagermoser Dobey, L. M., and Gallucci, J. (2014). Treatment integrity of interventions with children in school psychology international from 1995 – 2010. *Sch. Psychol. Int.* 35, 370–383. doi: 10.1177/0143034313476399

Sanetti, L. M. H., and Kratochwill, T. R. (2009). Treatment integrity assessment in the schools: an evaluation of the treatment integrity planning protocol. *Sch. Psychol. Q.* 24, 24–35. doi: 10.1037/a0015431

Skaalvik, E. M., and Skaalvik, S. (2010). Teacher self-efficacy and teacher burnout: A study of relations. *Teach. Teach. Educ.* 26, 1059–1069. doi: 10.1016/j.tate.2009.11.001

Skaalvik, E. M., and Skaalvik, S. (2014). Teacher self-efficacy and perceived autonomy: relations with teacher engagement, job satisfaction, and emotional exhaustion. *Psychol. Reports: Employment Psychol. Market.* 114, 68–77. doi: 10.2466/14.02.PR0.114k14w0

Smith, S., Barajas, K., Ellis, B., Moore, C., McCauley, S., and Reichow, B. (2021). A meta-analytic review of randomized controlled trials of the good behavior game. *Behav. Modif.* 45, 641–666. doi: 10.1177/0145445519878670

St Clair, L., and Stone, T. (2016). Who gets the better educators in afterschool? An analysis of teaching and learning interactions and student economic status. *Sch. Community J.* 26, 71–81.

Swars, S. L., Daane, C. J., and Giesen, J. (2006). Mathematics anxiety and mathematics teacher efficacy: what is the relationship in elementary preservice teachers? *Sch. Sci. Math.* 106, 306–315. doi: 10.1111/j.1949-8594.2006.tb17921.x

Swift, L. E., Hubbard, J. A., Bookhout, M. K., Grassetti, S. N., Smith, M. A., and Morrow, M. T. (2017). Teacher factors contributing to dosage of the KiVa anti-bullying program. *J. Sch. Psychol.* 65, 102–115. doi: 10.1016/j.jsp.2017.07.005

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

Thierry, K. L., Vincent, R. L., and Norris, K. (2022). Teacher-level predictors of the fidelity of implementation of a social-emotional learning curriculum. *Early Educ. Dev.* 33, 92–106. doi: 10.1080/10409289.2020.1849896

Tolan, P., Molloy, L. E., Bradshaw, C. P., Downer, J. T., and Ialongo, N. (2020). Randomized trial testing the integration of the good behavior game and MyTeachingPartner: the moderating role of distress among new teachers on student outcomes. *J. Sch. Psychol.* 78, 75–95. doi: 10.1016/j.jsp.2019.12.002

Tschannen-Moran, M., Hoy, A. W., and Hoy, W. K. (1998). Teacher efficacy: its meaning and measure. *Rev. Educ. Res.* 68, 202–248. doi: 10.3102/00346543068002202

van Horn, J. E., Taris, T. W., Schaufeli, W. B., and Schreurs, P. J. G. (2004). The structure of occupational well-being: A study among Dutch teachers. *J. Occup. Organ. Psychol.* 77, 365–375. doi: 10.1348/0963179041752718

Watt, H. M. G., and Richardson, P. W. (2007). Motivational factors influencing teaching as a career choice: development and validation of the FIT-choice scale. *J. Exp. Educ.* 75, 167–202. doi: 10.3200/JEXE.75.3.167-202

Wehby, J. H., Maggin, D. M., Partin, T. C. M., and Robertson, R. (2012). The impact of working alliance, social validity, and teacher burnout on implementation fidelity of the good behavior game. *Sch. Ment. Heal.* 4, 22–33. doi: 10.1007/s12310-011-9067-4

Wei, R. C., Darling-Hammond, L., and Adamson, F. (2010). *Professional development in the United States: Trends and challenges*. The Stanford Center for Opportunity Policy in Education.

What Works Clearinghouse. (2017). What works clearinghouse standards handbook. Available at: https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_standards_handbook_v4_draft.pdf

Witkiewitz, K., Falk, D. E., Kranzler, H. R., Litten, R. Z., and Hallgren, K. A. (2014). Methods to analyze treatment effects in the presence of missing data for a continuous heavy drinking outcome measure when participants drop out from treatment in alcohol clinical trials. *Alcohol. Clin. Exp. Res.* 38, 2826–2834. doi: 10.1111/acer.12543

Witt, J. C., Noell, G. H., LaFleur, L. H., and Mortenson, B. P. (1997). Teacher use of interventions in general education settings: measurement and analysis of the independent variable. *J. Appl. Behav. Anal.* 30, 693–696. doi: 10.1901/jaba.1997.30-693

Wolf, M. M. (1978). Social validity: the case for subjective measurement or how applied behavior analysis is finding its heart. *J. Appl. Behav. Anal.* 11, 203–214. doi: 10.1901/jaba.1978.11-203

Zee, M., and Koomen, H. M. Y. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Rev. Educ. Res.* 86, 981–1015. doi: 10.3102/0034654315626801

frontiers | Frontiers in Psychology

TYPE Perspective
PUBLISHED 03 April 2023
DOI 10.3389/fpsyg.2023.1034633

OPEN ACCESS

EDITED BY
Stephanie M. Jones,
Harvard University,
United States

REVIEWED BY
Emily Hanno,
Manpower Demonstration Research Corporation,
United States
Sophie Barnes,
Harvard University,
United States

*CORRESPONDENCE
Deborah Jane Leong
✉ DLeong@toolsofthemind.org

SPECIALTY SECTION
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

RECEIVED 01 September 2022
ACCEPTED 03 March 2023
PUBLISHED 03 April 2023

CITATION
Bodrova E, Leong DJ and Yudina E (2023) Play is a play, is a play, is a play... or is it? Challenges in designing, implementing and evaluating play-based interventions.
Front. Psychol. 14:1034633.
doi: 10.3389/fpsyg.2023.1034633

Play is a play, is a play, is a play... or is it? Challenges in designing, implementing and evaluating play-based interventions

Elena Bodrova[1], Deborah Jane Leong[1]* and Elena Yudina[2]

[1]Tools of the Mind, Denver, CO, United States, [2]Moscow School of Social and Economic Sciences, Moscow, Russia

When a social-emotional learning (SEL) intervention is implemented in an early childhood classroom, it often involves play. Some interventions even list play as its main component. However, the advocates of play arguing for the return of play in early childhood education (ECE) classrooms still have difficulty convincing the proponents of more rigorous academic instruction. These proponents cite research pointing to the insufficient evidence of the positive effect of play on children's short- and longer-term social, emotional, cognitive, and behavioral outcomes as well as their overall well-being. We believe that there are multiple issues with play-based interventions' design, implementation, and evaluation that might account for this insufficient evidence. In our paper, we discuss the numerous ways play does (or does not) feature in SEL interventions and how it might affect the outcomes of these interventions. We also examine the methodological challenges of having child-controlled play as a component of an SEL intervention. While we are not proposing a specific protocol for re-evaluation of the results of existing interventions, we outline some ways such re-evaluation can be possible in the future, along with the development and evaluation of new play-based SEL interventions.

KEYWORDS

social-emotional learning, play, classroom intervention, early childhood, children

1. Introduction

With early childhood being the formative period for the development of children's social and emotional skills, it is now recognized that the programs targeting these skills in the early years have the greatest potential to promote children's well-being and healthy development (Blewitt et al., 2018). Over the past decades, the number of these programs has been growing, with the programs taking on various formats, from narrowly targeted social-emotional learning (SEL) interventions to the comprehensive early childhood education (ECE)[1] curricula that embed SEL support in multiple materials and activities.

1 In this paper, we are going to use ECE as one of the two most commonly used terms describing this field. In some of the literature reviewed the term ECEC is used, which stands for Early Childhood Education and Care.

To assist educators in the adoption process, several review papers have been published that compare various SEL programs in their effectiveness in strengthening children's social and emotional competencies and preventing challenging behaviors (Joseph and Strain, 2003; Dunlap and Powell, 2009; O'Conner et al., 2017; Blewitt et al., 2018; Yang et al., 2019; Murano et al., 2020). Authors identified several characteristics shared by successful programs, such as well-defined and systematically addressed social and emotional target skills, teacher professional development and ongoing support, and the continuity in supporting social and emotional skills between school, family, and community (O'Conner et al., 2017). All these characteristics do not seem to be unique to ECE and could be applied to the interventions implemented at any grade level.

These reviews leave us with a question: If we assume a qualitative difference between different periods in child development, would not this imply that there are some unique features of young children's learning and development and that these unique features should be reflected in the design and implementation of the SEL interventions? We suggest that one of these unique features of early childhood is children's engagement in play as a freely chosen and intrinsically motivated activity controlled by the players. While humans engage in various forms of play way beyond early childhood (Van Vleet and Feeney, 2015), it is the early years when the impact of play on development is the greatest. The acknowledgment of the critical importance of play for young children is based on evidence from diverse fields, from evolutionary psychology (Greve and Thomsen, 2016) to child development (Vygotsky, 1967; Hewes, 2014) to pediatrics (Ginsburg, Committee on Communications, and Committee on Psychological Aspects of Child and Family Health, 2007; Yogman et al., 2018) and is reflected in the UN Convention on the Rights of the Child (Hodgkin and Newell, 2007). It, therefore, seems logical to include play as a context for SEL development in early childhood.

The words 'early childhood' and 'play' have been almost synonymous for so long that very few scholars question this connection. The pioneering work of Mildred Parten (1932) made the connection between play and children's social development universally recognized. This tradition of associating changes in children's play with their social development, however, stands in contrast to the fact that play is rarely mentioned in the reviews of the research on SEL programs and is never discussed as one of the active ingredients of these programs. In addressing child development from the perspective of the 'whole child,' this omission is problematic. In this paper, we discuss commonly used definitions of play, explore the different ways play is used in SEL interventions, examine reasons why play might be omitted or underused as a specific SEL strategy, and address ways in which developers of the SEL interventions might look at play in relation to the different aspects of their interventions.

The topic of play in early childhood has recently re-surfaced in the context of the increasing academic pressure experienced by preschool and kindergarten teachers leading to the virtual disappearance of play from ECE classrooms and the entire culture of childhood (Gray, 2011; Belknap and Hazler, 2014; Wohlwend and Peppler, 2015; Barblett et al., 2016; Bassok et al., 2016; Whitebread, 2017; Wasmuth and Nitecki, 2020; Digennaro, 2021). Given that many of today's children spend a significant portion of their waking hours in some kind of a school or center environment and that their opportunities to engage in play outside of this environment are diminishing (Singer et al., 2009), it is imperative to make sure that play becomes a critical part of any early childhood intervention and especially an SEL intervention.

2. Play: What's in the word

One of the reasons that play is not prominently featured in interventions that attempt to promote children's social and emotional development is the elusive nature of play: while everyone seems to have an intuitive understanding of how play is different from non-play, it is hard to convert this understanding into well-defined characteristics of play that can be reliably measured and manipulated. While the educational field has not come to a single definition of play, there seems to exist some agreement about the features of an activity that qualifies as play: it must be pleasurable, process-oriented, intrinsically motivated, meaningful, iterative, and controlled by a child (Canning, 2012; Zosh et al., 2017, 2022). The activity possessing these qualities is frequently described as *spontaneous*[2] or *free* play (Hewes, 2014) to distinguish it from other activities that retain some degree of playfulness but are not entirely intrinsically motivated or child controlled.

Those other 'playful' activities are given such names as *guided play* or *purposeful play* to emphasize the fact that this kind of play is controlled (at least partially) not by a child but by an adult (Hirsh-Pasek and Golinkoff, 2008; Weisberg et al., 2016; Yu et al., 2018; Allee-Herndon and Roberts, 2021). The adult-initiated play category also includes 'serious games' (Zosh et al., 2018), sometimes called *structured play* (Healey and Healey, 2019). The division between free play and adult-involved play is not static: an adult may intervene in children's play to infuse it with the academic content without completely taking it over. Zosh et al. (2018) attempted to capture the heterogeneity of children's playful experiences by conceptualizing play as a spectrum, with these experiences differing in terms of adult or child initiation and direction of play and the presence of a learning goal.

Authors sometimes combine free play, guided play, and games in a more general category of *playful* or *play-based learning* (Danniels and Pyle, 2018; Zosh et al., 2022). This kind of learning is argued to be preferable for young children as compared to learning in more 'schoolified' settings (Zosh et al., 2022). At the same time, when examining the role of play in social-emotional learning, it seems essential to unpack the concept of 'playful learning' and to identify the exact characteristics of a specific playful experience, such as the degree of adult-directedness or child agency. In our paper, we will reference these and other characteristics of play when discussing the use of play in SEL interventions.

In addition to the activities explicitly labeled 'playful,' many activities for young children in the SEL interventions are designed to have some play elements. Examples include children's role-playing that follows a script of a social situation (Wee et al., 2022) or using teddy

2 The term *spontaneous* is used in this context to mean *child-initiated* or *child-controlled*, which does not imply that play *spontaneously emerges* once a child reaches a certain age. In fact, if and when a specific kind of play emerges in children varies significantly among cultures and historical periods (Elkonin, 2005; Gaskins, 2014; Wood, 2014).

bears to help children express their feelings (Koplow, 2008). Unfortunately, based only on the researchers' accounts, it is hard to determine whether these activities were considered 'playful' by children. Even young children can detect when adults offer them a learning activity under the guise of play, and children perceive an activity as 'play' when there is "an element of choice and sharing of control" (Jensen et al., 2021, 493).

3. Locating play in the space of SEL interventions

Different kinds of play and other playful activities can be a component of an SEL intervention. To maximize the role of these activities, it is important to identify them correctly and examine their relationship with specific social and emotional skills. Combining disparate programs under an umbrella of 'play-based' makes it difficult to unpack these programs' effects on social and emotional development. We identified four ways play and the activities described as 'playful' are (or are not) included in the SEL interventions: (1) play is not included in the design of the intervention, thus making it 'invisible' for the intervention developers and researchers; (2) play is used in the intervention as the primary vehicle to promote SEL; (3) the intervention focuses on improving the quality of play, and (4) SEL is one of the areas targeted by a comprehensive play-based curriculum. For each of these four categories, we see different challenges in the design, implementation, and evaluation of the intervention associated with the way children engage in play.

3.1. Challenges in making 'invisible' play visible

Some interventions consist of a series of lessons, each teaching a specific skill, such as recognizing and labeling one's own and each other's emotions or inhibiting impulsive reactions (e.g., Domitrovich et al., 2007; Webster-Stratton and Reid, 2008). Teachers usually deliver these lessons in a large group setting (e.g., during circle time), and children then practice newly learned skills throughout their daily activities (Blewitt et al., 2018). In typical ECE classrooms, this practice would most likely occur during center time or other free play periods. Although play is not explicitly listed as a component of practice activities and therefore stays 'invisible' to the developers of the intervention, it is likely that it still figures in some way in what children are doing. As play within these classroom activities remains 'invisible,' it is unclear which of the activity's quantitative and qualitative characteristics allow children to practice newly learned social and emotional skills. The quantitative characteristics include but are not limited to the overall duration of activity as well as the duration of uninterrupted activity, the number of children in the same center at a given time, the number of children entering and exiting a center during a specific time period, etc.

The choice of qualitative characteristics to examine depends on the specific aspect of social-emotional learning. For example, qualitatively different activities such as joint block-building and social pretend play provide different opportunities for developing communication, cooperation, and perspective-taking skills. If play is part of this context where young children's social-emotional learning happens, it is crucial to make play 'visible' and take into account the elements that would make it more effective in supporting social-emotional skill development, such as the degree of children's control over the flow of play, the existence of rules and the opportunities to establish new rules, etc. (Burdette and Whitaker, 2005; Hewes, 2014; Jarvis et al., 2014; Nicolopoulou and Smith, 2022).

While the literature on these interventions does not provide a description of play in the treatment classrooms, it often mentions the curriculum used (e.g., HighScope or Creative Curriculum) or the kind of setting (Head Start classroom, public school pre-kindergarten, etc.). A closer look at the classrooms may show us significant variations in the implementation of the same curriculum or in following the same program guidelines. For example, most preschool classrooms have a substantial portion of time on their schedule described as 'free choice time' or 'center time.' This time block is usually when most of the indoor play takes place. Whether or not children have ample opportunity to practice social and emotional skills in play depends to some degree on how this free choice time is managed.

There are at least two variables that, in our opinion, should be considered in planning an intervention regarding play that happens in the activity centers during the free choice time that serves as a context for SEL. The first variable is the time children spend in one center or one activity. In some early childhood programs, children spend the entire time engaged in play as several centers get integrated into a general play theme. In the others, children rotate from one center to the next, which leaves them with less than 20 min to spend in each center (Paulick, 2019), although it has been known for a long time that children need at least 30 min to engage in high-level play (Christie and Wardle, 1992). In their paper, Christie and Wardle make a compelling case for allocating more time to uninterrupted play as they demonstrate the complexity of children's behaviors in the preparatory stages of high-level sociodramatic or constructive play. These behaviors are necessary for play to reach this level. The authors also address the issue of children needing access to all activity centers by suggesting closing centers on a rotating basis for children to have "several long play periods per week rather than short daily ones" (p. 30).

Many kindergarten classrooms have activity centers that do not necessarily allow for sociodramatic or constructive play. Instead, centers are where children practice academic skills (Bassok et al., 2016). The kind of play most common in kindergarten is board games, but sometimes the access to the games is contingent on children completing the academic assignments, so the actual time spent playing may vary.

Another variable to consider is the degree of teacher-directedness of the activities available in the centers during 'play time.' Contrary to its name, in many programs, the 'free choice time' provides children with limited choices as some centers get converted to small group teacher-directed activities (Paulick, 2019). While teacher-directed activities are associated with children's gains in academics (De Haan et al., 2014; Goble and Pianta, 2017), it is the child-managed activities that contribute to the development of such social-emotional skills as inhibitory control, attention, and resilience to stress (Burdette and Whitaker, 2005; Hewes, 2014; Goble and Pianta, 2017).

Making play (or the absence of play) in ECE classrooms 'visible' would help the developers of SEL interventions in the planning stages as they decide if their intervention has a good chance of producing its desired outcomes when implemented in these classrooms. It might

also provide an additional lens through which they can evaluate the implementation results.

3.2. Challenges in using play as an instrument to promote SEL

The interventions that use play to promote social and emotional competencies vary in the type of play used and the target skills. The typical features of these interventions are their relatively short time span (usually several weeks), short duration of play activities (20–30 min), and a relatively high degree of adult-directedness of the play activities offered to children. Examples of these activities include movement games with increasingly more complex rules (McClelland et al., 2019), listening to and acting out the stories (Joseph and Strain, 2003; Mondi et al., 2021), and playing group games (Barrow et al., 2015). The outcomes of these interventions include emotion control, self-regulatory skills, and a decrease in behavior problems (Healey and Healey, 2019). Children's social and emotional skills progress is typically measured immediately after the intervention ends and sometimes several months later on various measures, including teacher reports and standardized tests. Most authors report that assessing the transfer of skills targeted by their intervention to children's free play and other contexts falls outside the scope of their current research while at the same time acknowledging the importance of studying this transfer in the future.

The issue of the transfer from mostly adult-directed play used in these interventions to other classroom activities is especially important for such skills as self-regulation and emotion control. Are children able to apply newly learned skills in an activity that is completely child-initiated and child-controlled? Is there a gradual transition from adult regulation to self-regulation, and what are the activities that work best in facilitating this transition? With self-regulation and emotion control being a target of an SEL intervention, these questions need to be answered not only to determine the practical benefits of this intervention but also to inform the development of its new versions.

3.3. Challenges in improving the quality of play as a means to promote SEL

Over the past decades, evidence has been accumulating, indicating that children's pretend play is experiencing a decline not just in its quantity but in its quality as well (Jarvis et al., 2014; Lewis, 2017; Smirnova and Gudareva, 2017). Today's four-and five-year-olds are playing in a way more typical for younger children: their pretend scenarios are stereotypical, the use of props is non-imaginative, and they cannot sustain play for prolonged periods (Lemay et al., 2022). Researchers refer to this kind of play as 'low-level' or 'immature,' contrasting it with 'high-level' or 'mature' play that involves "elaborate group dramatizations and complex construction projects" (Christie and Wardle, 1992, p. 28). Immature play is associated with lower levels of various social and emotional skills, including self-regulation and cooperative behaviors (Slot et al., 2017). It is possible that this decline in the quality of play is one of the reasons that play-based interventions do not always produce expected results and that some scholars have started voicing their doubts about the validity of assigning play the central place in promoting health, learning, and development of young children (Lillard et al., 2013).

This observable proliferation of less developed or immature play prompted the researchers to design interventions to promote the quality of child-controlled play. The interventions promoting mature play were able to elevate levels of play and improve children's emotion control, executive functions, and self-regulation (Diamond et al., 2007; Blair and Raver, 2014; Perren et al., 2019; Richard et al., 2021; Adam et al., 2022). This group of interventions faces additional challenges associated with the complex nature of child-controlled play.

The need to maintain a delicate balance between teacher support and children's independence presents a significant challenge for assessing play, as many of the skills contributing to mature play, such as background knowledge or problem-solving, seem to belong to the unconstrained category: they cannot be taught directly but instead develop gradually through varied experiences (McCormick et al., 2021) Assessing unconstrained skills presents a challenge not only for the play researchers but for the entire field of child development and learning (Dowd and Thomsen, 2021). In addition, using some of the existing play measures does not provide an accurate picture of the status of play: often, what is being observed and measured is not yet actual child-controlled play, but most likely, it is still adult-guided play involving substantial teacher support. Using more fine-grain measures, such as teacher-child interactions in play, may establish the relationships between these interactions and children's social-emotional learning, such as their use of regulation-related skills (Moreno et al., 2017). Yet another approach is to assume that in an ECE classroom, children's play always reflects both: the adult support and the children's ability to benefit from this support. This approach yielded measures of play that combine children-level variables with teacher-level variables (Leong and Bodrova, 2012; Germeroth et al., 2019). These instruments provide an overall measure of play maturity in the context of scaffolded teacher-child interactions.

Another challenge lies in the timing of the evaluation. It takes a long time for children's play skills to fully develop and solidify, especially if these children initially have immature play skills. As a result, evaluations scheduled too early can yield lower-than-expected outcomes. Providing play support for an entire year or even longer, as well as monitoring levels of play maturity, will hopefully allow researchers and curriculum developers to collect necessary evidence linking the development of mature play to the growth of children's social and emotional skills.

3.4. Challenges in using a play-based curriculum to promote SEL

The variations in the use of the term *play* result in the variability of what counts as a play-based curriculum. Many such curricula self-define as play-based by contrast, i.e., implying that they are not using didactic modes of instruction but instead engage children in play and games without specifying what these play and games are (Reynolds et al., 2011). Some curricula implement play-based methods in teaching specific subject matters (mostly literacy and math), while others generalize the play-based approach to all areas of child development. There are several theoretical models underlying play-based curricula, including the idea of combining playful activities with varying degrees of adult-directedness (Zosh et al., 2022), infusing

academic objectives in promoting children's pretend play (Fleer et al., 2017), and co-constructing play as Developmental Education (Van Oers and Duijkers, 2013).

At the same time, many commercially available and teacher-designed curricula self-identify as 'play-based' by merely describing the classroom setup (activity centers and not desks). As we discussed in the 'invisible play' section, the presence of centers does not necessarily translate into the quantity or quality of play in these centers. Knowing not only the *intended* curriculum but also the *enacted* one (Porter et al., 2001) might help determine the optimal fit between the SEL intervention and the context where it gets implemented.

Regarding the evaluation, a play-based curriculum faces all challenges discussed in the previous sections. In addition, these curricula are currently expected to deliver results not only in the general areas of child development but also in foundational academic competencies. With the academic (mostly 'constrained') skills more amenable to change short-term (Casbergue, 2010; McCormick et al., 2021), play-based curricula often find themselves at a disadvantage compared to skill-based curricula when children get assessed on discrete academic skills only. Unfortunately, it is becoming harder and harder to use previous studies, including the classic Perry Preschool study (Schweinhart, 2019), in defense of play-based programs because of the changes in school readiness expectations. Conducting new longitudinal studies with children repeatedly assessed on both cognitive and noncognitive (Heckman, 2011) skills may help determine not only the immediate but also long-term effects of play-based instruction during early childhood. These effects may be latent, manifesting many years later, or they may interact with other educational or parenting factors resulting in a cumulative impact on child development (Maggi et al., 2010).

Additional evaluation challenges are associated with quality rating systems and other classroom environmental rating scales used to evaluate early childhood programs. Although these instruments position themselves as 'curriculum independent,' they often value classroom practices that promote only one specific type of play at the expense of other playful activities. For example, a widely used Early Childhood Environmental Rating Scale (Harms et al., 2014) distinguishes between a 'play area' defined as a space with pretend play materials, and 'interest centers' such as blocks or art. Such distinction implies that make-believe play in the classroom is limited to one area only and that children are discouraged from using materials from other centers as props in their play. A more holistic approach to evaluation that includes parents' and teachers' views might contribute to a more favorable opinion of play-based curricula among school administrators and policymakers.

4. Not just child's play

All SEL interventions seem to share some common concerns in regard to using play. One is an apparent contradiction between the child-controlled character of play and the adult role in supporting and enhancing play. With children coming to preschool and already playing at a mature level, the role of adults was limited until recently to providing the conditions (time, space, and materials) for play to happen; but it is no longer the case. Even when provided time, space, and materials, children playing with no adult support demonstrate a lower quality of play compared to children who receive some play tutoring from an adult delivered in the form of prompting, verbalization, and modeling (Kalkusch et al., 2021). Sometimes children even regress in their quality of play (Farran and Son-Yarbrough, 2001). At the same time, too much adult intervention in play destroys its voluntary and intrinsically motivated flow and thus potentially diminishes its potential benefits (Gmitrová and Gmitrov, 2003; Nome, 2015).

Also, play-based interventions cannot be completely formalized and manualized (Murphy and Gutman, 2012). While adults can and must create conditions for play, they cannot completely predict or control the result that emerges. Therefore, implementation fidelity cannot be reduced to teachers faithfully following the steps of the activities specified in the manual.

The analysis of play-based SEL interventions also highlights some of the challenges facing most of the interventions designed to be implemented in a classroom or any educational context: on the one hand, teachers should be given the freedom to adjust, modify, and individualize, but on the other hand, too much freedom makes it difficult to compare different classrooms and generalize the effect of the intervention. In addition, assigning teachers to a treatment group at random, as it is common in the evaluations of SEL interventions, may result in a poor fit between these teachers' prior experiences and educational philosophy on the one hand and the nature of the intervention on the other. This might lead to these teachers' low 'commitment to implement' (Cramer et al., 2021) and, in turn, to disappointing outcomes of the intervention. Although many of the education clearinghouses still use the results of the RCTs to select promising interventions, more and more educational researchers are now voicing their concerns about the preferential treatment of one particular research design and even about the appropriateness of the use of this 'gold standard' in education (Sullivan, 2011; Thomas, 2016).

In this paper, we examined some relationships between children's play and SEL interventions. Given the multifaceted nature of play and the multitude of interventions designed to promote children's social and emotional development, there could be many more relationships waiting to be examined. We propose that researchers pay attention to the quantity and quality of play while planning, implementing, or evaluating SEL interventions targeting young children. We expect that this might not only increase the effectiveness of these interventions but also contribute to our growing understanding of children's development and learning.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Adam, N., Blaye, A., Gulbinaite, R., Chabé-Ferret, S., and Farrer, C. (2022). A multidimensional evaluation of the benefits of an ecologically realistic training based on pretend play for preschoolers' cognitive control and self-regulation: from behavior to the underlying theta neuro-oscillatory activity. *J. Exp. Child Psychol.* 216:105348. doi: 10.1016/j.jecp.2021.105348

Allee-Herndon, K. A., and Roberts, S. K. (2021). The power of purposeful play in primary grades: adjusting pedagogy for children's needs and academic gains. *J. Educ.* 201, 54–63. doi: 10.1177/0022057420903272

Barblett, L., Knaus, M., and Barratt-Pugh, C. (2016). The pushes and pulls of pedagogy in the early years: competing knowledges and the erosion of play-based learning. *Australas. J. Early Childhood* 41, 36–43. doi: 10.1177/183693911604100405

Barrow, M. I., Jaques, R. L., Ponischil, K., Lengua, L. J., and Bunge, S. A. (2015) Improved cognitive flexibility after a structured play intervention with a high-risk sample of preschoolers. Unpublished Manuscript. Available from: https://www.academia.edu/25461921/Improved_cognitive_flexibility_after_a_structured_play_intervention_with_a_high_risk_sample_of_preschoolers

Bassok, D., Latham, S., and Rorem, A. (2016). Is kindergarten the new first grade? *AERA Open* 2:233285841561635. doi: 10.1177/2332858415616358

Belknap, E., and Hazler, R. (2014). Empty playgrounds and anxious children. *J. Creat. Ment. Health* 9, 210–231. doi: 10.1080/15401383.2013.864962

Blair, C., and Raver, C. C. (2014). Closing the achievement gap through modification of neurocognitive and neuroendocrine function: results from a cluster randomized controlled trial of an innovative approach to the education of children in kindergarten. *PLoS One* 9:e112393. doi: 10.1371/journal.pone.0112393

Blewitt, C., Fuller-Tyszkiewicz, M., Nolan, A., Bergmeier, H., Vicary, D., Huang, T., et al. (2018). Social and emotional learning associated with universal curriculum-based interventions in early childhood education and care centers: a systematic review and meta-analysis. *JAMA Netw. Open* 1:e185727. doi: 10.1001/jamanetworkopen.2018.5727

Burdette, H. L., and Whitaker, R. C. (2005). Resurrecting free play in young children: looking beyond fitness and fatness to attention, affiliation, and affect. *Arch. Pediatr. Adolesc. Med.* 159, 46–50. doi: 10.1001/archpedi.159.1.46

Canning, N. (2012). "Exploring the concept of quality play," in *Implementing Qual. Improvement Change Early Years*, eds. M. Reed and N. Canning (SAGE Publications) 75–91. doi: 10.4135/9781446250747.n6

Casbergue, R. M. (2010). Assessment and instruction in early childhood education: early literacy as a microcosm of shifting perspectives. *J. Educ.* 190, 13–20. doi: 10.1177/0022057410190001-204

Christie, J. F., and Wardle, F. (1992). How much time is needed for play? *Young Child.* 47, 28–33. https://www.jstor.org/stable/42728239

Cramer, T., Ganimian, A., Morris, P., and Cappella, E. (2021). The role of teachers' commitment to implement in delivering evidence-based social-emotional learning programs. *J. Sch. Psychol.* 88, 85–100. doi: 10.1016/j.jsp.2021.08.003

Danniels, E., and Pyle, A. (2018). Defining play-based learning," in *Centre of Excellence for Early Childhood Development*. I. E. Tremblay, M. Boivin, and R. D. V. Peters 1–5. Retrieved from: http://www.child-encyclopedia.com/playbased-learning/according-experts/defining-play-based-learning

De Haan, A. K., Elbers, E., and Leseman, P. P. (2014). Teacher-and child-managed academic activities in preschool and kindergarten and their influence on children's gains in emergent academic skills. *J. Res. Child. Educ.* 28, 43–58. doi: 10.1080/02568543.2013.851750

Diamond, A., Barnett, W. S., Thomas, J., and Munro, S. (2007). Preschool program improves cognitive control. *Science* 318, 1387–1388. doi: 10.1126/science.1151148

Digennaro, S. (2021). Decline of free play as a form of educational poverty. *J. Phys. Educ. Sport* 21, 657–660.

Domitrovich, C. E., Cortes, R. C., and Greenberg, M. T. (2007). Improving young children's social and emotional competence: a randomized trial of the preschool PATHS curriculum. *J. Prim. Prev.* 28, 67–91. doi: 10.1007/s10935-007-0081-0

Dowd, A. J., and Thomsen, B. S. (2021). *Learning Through Play: Increasing Impact, Reducing Inequality*. Billund, DK: LEGO Foundation.

Dunlap, G., and Powell, D. (2009). Promoting social behavior of young children in group settings: a summary of research. Roadmap to effective intervention practices # 3. *Tech. Assistance Center Soc. Emotional Intervention Young Children*

Elkonin, D. B. (2005). Chapter 2: on the historical origin of role play. *J. Russ. East Eur. Psychol.* 43, 49–89. doi: 10.1080/10610405.2005.11059243

Farran, D. C., and Son-Yarbrough, W. (2001). Title I funded preschools as a developmental context for children's play and verbal behaviors. *Early Child. Res. Q.* 16, 245–262. doi: 10.1016/S0885-2006(01)00100-4

Fleer, M., Veresov, N., Harrison, L., and Walker, S. (2017). Working with teachers' pedagogical strengths: the design of executive function activities for play-based programs. *Australas. J. Early Childhood* 42, 47–55. doi: 10.23965/AJEC.42.4.06

Gaskins, S. (2014). "Children's play as cultural activity" in *The SAGE Handbook of Play and Learning in Early Childhood*. eds. L. Brooker, M. Blaise and S. Edwards (Los Angeles, CA: SAGE Publications Ltd), 31–42.

Germeroth, C., Bodrova, E., Day-Hess, C., Barker, J., Sarama, J., Clements, D. H., et al. (2019). Play it high, play it low: examining the reliability and validity of a new observation tool to measure children's make-believe play. *Am. J. Play* 11, 183–221.

Ginsburg, K. R.Committee on Communications, & Committee on Psychosocial Aspects of Child and Family Health (2007). The importance of play in promoting healthy child development and maintaining strong parent-child bonds. *Pediatrics* 119, 182–191. doi: 10.1542/peds.2006-2697

Gmitrová, V., and Gmitrov, J. (2003). The impact of teacher-directed and child-directed pretend play on cognitive competence in kindergarten children. *Early Childhood Educ. J.* 30, 241–246. doi: 10.1023/A:1023339724780

Goble, P., and Pianta, R. C. (2017). Teacher–child interactions in free choice and teacher-directed activity settings: prediction to school readiness. *Early Educ. Dev.* 28, 1035–1051. doi: 10.1080/10409289.2017.1322449

Gray, P. (2011). The decline of play and the rise of psychopathology in children and adolescents. *Am. J. Play* 3, 443–463.

Greve, W., and Thomsen, T. (2016). Evolutionary advantages of free play during childhood. *Evol. Psychol.* 14:147470491667534. doi: 10.1177/1474704916675349

Harms, T., Clifford, R., and Cryer, D. (2014). *Early Childhood Environment Rating Scale (ECERS-3)*. New York, NY: Teacher College Press.

Healey, D., and Healey, M. (2019). Randomized controlled trial comparing the effectiveness of structured-play (ENGAGE) and behavior management (TRIPLE P) in reducing problem behaviors in preschoolers. *Sci. Rep.* 9, 1–9. doi: 10.1038/s41598-019-40234-0

Heckman, J. J. (2011). "Effective child development strategies" in *The Pre-K Debates: Current Controversies and Issues*. eds. E. Zigler, W. S. Gilliam and W. S. Barnett (Baltimore, MD: Paul H. Brooks Publishing Co), 2–8.

Hewes, J. (2014). Seeking balance in motion: the role of spontaneous free play in promoting social and emotional health in early childhood care and education. *Child Aust.* 1, 280–301. doi: 10.3390/children1030280

Hirsh-Pasek, K., and Golinkoff, R. M. (2008). Why play= learning. *Encycl. Early Childhood Dev.* 1, 1–7.

Hodgkin, R., and Newell, P. (2007). *Implementation Handbook for the Convention on the Rights of the Child*; United Nations Children's Fund (UNICEF): Geneva, Switzerland

Jarvis, P., Newman, S., and Swiniarski, L. (2014). On 'Becoming Social': the importance of collaborative free play in childhood. *Int. J. Play* 3, 53–68. doi: 10.1080/21594937.2013.863440

Jensen, H., Pyle, A., Alaca, B., and Fesseha, E. (2021). Playing with a goal in mind: exploring the enactment of guided play in Canadian and south African early years classrooms. *Early Years* 41, 491–505. doi: 10.1080/09575146.2019.1619670

Joseph, G. E., and Strain, P. S. (2003). Comprehensive evidence-based social-emotional curricula for young children: an analysis of efficacious adoption potential. *Top. Early Child. Spec. Educ.* 23, 62–73. doi: 10.1177/02711214030230020201

Kalkusch, I., Jaggy, A.-K., Bossi, C. B., Weiss, B., Sticca, F., and Perren, S. (2021). Promoting social pretend play in preschool age: is providing role-play material enough? *Early Educ. Dev.* 32, 1136–1152. doi: 10.1080/10409289.2020.1830248

Koplow, L. (2008). *Bears, Bears Everywhere!: Supporting Children's Emotional Health in the Classroom*. New York, NY: Teachers College Press.

Lemay, L., Parent, A.-S., Bouchard, C., Duval, S., and Montminy-Sanschagrin, N. (2022). "A closer look at the maturity of make-believe play in Quebec's kindergarten to inform the design of a professional development program" in *Teacher Education and Play Pedagogy: International Perspectives*. eds. E. Loizou and J. Trawick-Smith (London, UK: Routledge), 9–26.

Leong, D. J., and Bodrova, E. (2012). Assessing and scaffolding make-believe play. *Young Child.* 29, 28–34.

Lewis, P. J. (2017). The erosion of play. *Int. J. Play* 6, 10–23. doi: 10.1080/21594937.2017.1288391

Lillard, A. S., Lerner, M. D., Hopkins, E. J., Dore, R. A., Smith, E. D., and Palmquist, C. M. (2013). The impact of pretend play on children's development: a review of the evidence. *Psychol. Bull.* 139, 1–34. doi: 10.1037/a0029321

Maggi, S., Irwin, L. J., Siddiqi, A., and Hertzman, C. (2010). The social determinants of early child development: an overview. *J. Paediatr. Child Health* 46, 627–635. doi: 10.1111/j.1440-1754.2010.01817.x

McClelland, M. M., Tominey, S. L., Schmitt, S. A., Hatfield, B. E., Purpura, D. J., Gonzales, C. R., et al. (2019). Red light, purple light! Results of an intervention to promote school readiness for children from low-income backgrounds. *Front. Psychol.* 10:2365. doi: 10.3389/fpsyg.2019.02365

McCormick, M., Weiland, C., Hsueh, J., Pralica, M., Weissman, A. K., Moffett, L., et al. (2021). Is skill type the key to the preK fadeout puzzle? Differential associations between enrollment in preK and constrained and unconstrained skills across kindergarten. *Child Dev.* 92, e599–e620. doi: 10.1111/cdev.13520

Mondi, C. F., Giovanelli, A., and Reynolds, A. J. (2021). Fostering socio-emotional learning through early childhood intervention. *Int. J. Child Care Educ. Policy* 15, 1–43. doi: 10.1186/s40723-021-00084-8

Moreno, A. J., Shwayder, I., and Friedman, I. D. (2017). The function of executive function: everyday manifestations of regulated thinking in preschool settings. *Early Childhood Educ. J.* 45, 143–153. doi: 10.1007/s10643-016-0777-y

Murano, D., Sawyer, J. E., and Lipnevich, A. A. (2020). A meta-analytic review of preschool social and emotional learning interventions. *Rev. Educ. Res.* 90, 227–263. doi: 10.3102/0034654320914743

Murphy, S. L., and Gutman, S. A. (2012). Intervention Fidelity: a necessary aspect of intervention effectiveness studies. *Am. J. Occup. Ther.* 66, 387–388. doi: 10.5014/ajot.2010.005405

Nicolopoulou, A., and Smith, P. K. (2022). "Social play and social development" in *The Wiley-Blackwell Handbook of Childhood Social Development*. eds. P. K. Smith and C. H. Hart (Hoboken, NJ: Wiley Blackwell), 538–554.

Nome, D. (2015). "Kindergartens—schools without recess: the consequence of an instrumentalist notion of play" in *Instrumentalism in Education-Where is Bildung Left?* eds. S. Hillen and C. Apres (Münster, DE: Waxmann), 15–27.

O'Conner, R., De Feyter, J., Carr, A., Luo, J. L., and Romm, H. (2017). *A Review of the Literature on Social and Emotional Learning for Students ages 3–8: Characteristics of Effective Social and Emotional Learning Programs (Part 1 of 4)*. Washington DC: Regional Educational Laboratory Mid-Atlantic.

Parten, M. B. (1932). Social participation among preschool children. *J. Abnorm. Soc. Psychol.* 27, 243–269. doi: 10.1037/h0074524

Paulick, J. H. (2019). Articulating high quality free choice time in head start preschools: a framework to support professional development and classroom observations. *HS Dialog: Res. Pract. Early Childhood Field* 22, 1–26.

Perren, S., Sticca, F., Weiss-Hanselmann, B., and Bossi, C. B. (2019). Let us play together! Can play tutoring stimulate children's social pretend play level? *J. Early Child. Res.* 17, 205–219. doi: 10.1177/1476718X19849248

Porter, A. C., John, L., and Smithson, J. L. (2001). Defining, developing, and using curriculum indicators. CPRE Research report series. Consortium for Policy Research in Education.

Reynolds, E., Stagnitti, K., and Kidd, E. (2011). Play, language and social skills of children attending a play-based curriculum school and a traditionally structured classroom curriculum school in low socioeconomic areas. *Australas. J. Early Childhood* 36, 120–130. doi: 10.1177/183693911103600416

Richard, S., Baud-Bovy, G., Clerc-Georgy, A., and Gentaz, E. (2021). The effects of a 'pretend play-based training' designed to promote the development of emotion comprehension, emotion regulation, and prosocial behaviour in 5-to 6-year-old Swiss children. *Br. J. Psychol.* 112, 690–719. doi: 10.1111/bjop.12484

Schweinhart, L. J. (2019). "Lessons on sustaining early gains from the life-course study of Perry preschool" in *Sustaining Early Childhood Learning Gains: Program, School, and Family Influences*. eds. A. J. Reynolds and J. A. Temple (Cambridge, UK: Cambridge University Press), 254–267.

Singer, D. G., Singer, J. L., D'Agnostino, H., and Raeka DeLong, R. (2009). Children's pastimes and play in sixteen nations: is free-play declining? *Am. J. Play* 1, 283–312.

Slot, P. L., Mulder, H., Verhagen, J., and Leseman, P. P. M. (2017). Preschoolers' cognitive and emotional self-regulation in pretend play: relations with executive functions and quality of play. *Infant Child Dev.* 26:e2038. doi: 10.1002/icd.2038

Smirnova, E. O., and Gudareva, O. V. (2017). The state of play activity among today's preschoolers. *J. Russ. East Eur. Psychol.* 54, 252–270. doi: 10.1080/10610405.2017.1393291

Sullivan, G. M. (2011). Getting off the 'gold standard': randomized controlled trials and education research. *J. Grad. Med. Educ.* 3, 285–289. doi: 10.4300/JGME-D-11-00147.1

Thomas, G. (2016). After the gold rush: questioning the 'gold standard' and reappraising the status of experiment and randomized controlled trials in education. *Harv. Educ. Rev.* 86, 390–411. doi: 10.17763/1943-5045-86.3.390

Van Oers, B., and Duijkers, D. (2013). Teaching in a play-based curriculum: theory, practice and evidence of developmental education for young children. *J. Curric. Stud.* 45, 511–534. doi: 10.1080/00220272.2011.637182

Van Vleet, M., and Feeney, B. C. (2015). Play behavior and playfulness in adulthood. *Soc. Personal. Psychol. Compass* 9, 630–643. doi: 10.1111/spc3.12205

Vygotsky, L. S. (1967). Play and its role in the mental development of the child. *Sov. Psychol.* 5, 6–18. doi: 10.2753/RPO1061-040505036

Wasmuth, H., and Nitecki, E. (2020). (un) intended consequences in current ECEC policies: revealing and examining hidden agendas. *Policy Futures Educ.* 18, 686–699. doi: 10.1177/1478210320949063

Webster-Stratton, C., and Reid, M. (2008). Adapting the incredible years child dinosaur social, emotional, and problem-solving intervention to address comorbid diagnoses. *J. Children's Serv.* 3, 17–30. doi: 10.1108/17466660200800016

Wee, S.-J., Kim, S. J., Chung, K., and Kim, M. (2022). Development of children's perspective-taking and empathy through bullying-themed books and role-playing. *J. Res. Child. Educ.* 36, 96–111. doi: 10.1080/02568543.2020.1864523

Weisberg, D. S., Hirsh-Pasek, K., Golinkoff, R. M., Kittredge, A. K., and David Klahr, D. (2016). Guided play: principles and practices. *Curr. Dir. Psychol. Sci.* 25, 177–182. doi: 10.1177/0963721416645512

Whitebread, D. (2017). Free play and children's mental health. *Lancet Child Adolesc. Health* 1, 167–169. doi: 10.1016/S2352-4642(17)30092-5

Wohlwend, K., and Peppler, K. (2015). All rigor and no play is no way to improve learning. *Phi Delta Kappan* 96, 22–26. doi: 10.1177/0031721715583957

Wood, E. A. (2014). Free choice and free play in early childhood education: troubling the discourse. *Int. J. Early Years Educ.* 22, 4–18. doi: 10.1080/09669760.2013.830562

Yang, W., Datu, J. A. D., Lin, X., Lau, M. M., and Li, H. (2019). Can early childhood curriculum enhance social-emotional competence in low-income children? A meta-analysis of the educational effects. *Early Educ. Dev.* 30, 36–59. doi: 10.1080/10409289.2018.1539557

Yogman, M., Garner, A., Hutchinson, J., Hirsh-Pasek, K., Golinkoff, R. M., Baum, R., et al. (2018). The power of play: a pediatric role in enhancing development in young children. *Pediatrics* 142, 1–17. doi: 10.1542/peds.2018-2058

Yu, Y., Shafto, P., Bonawitz, E., Yang, S. C.-H., Golinkoff, R. M., Corriveau, K. H., et al. (2018). The theoretical and methodological opportunities afforded by guided play with young children. *Front. Psychol.* 9:1152. doi: 10.3389/fpsyg.2018.01152

Zosh, J. N., Gaudreau, C., Golinkoff, R. M., and Hirsh-Pasek, K. (2022). The power of playful learning in the early childhood setting. *Young Child* 77, 6–13.

Zosh, J. M., Hirsh-Pasek, K., Hopkins, E. J., Jensen, H., Liu, C., Neale, D., et al. (2018). Accessing the inaccessible: redefining play as a spectrum. *Front. Psychol.* 9:1124. doi: 10.3389/fpsyg.2018.01124

Zosh, J. M., Hopkins, E. J., Jensen, H., Liu, C., Neale, D., Hirsh-Pasek, K., et al. (2017). *Learning Through Play: A Review of the Evidence*. Billund, DK: LEGO Foundation.

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 27 April 2023
DOI 10.3389/feduc.2023.1031516

OPEN ACCESS

EDITED BY
Andres Molano,
University of Los Andes, Colombia

REVIEWED BY
Garry Squires,
The University of Manchester, United Kingdom
Lin Lin,
Shanghai University of Finance and Economics, China
Juliette Berg,
American Institutes for Research, United States
Rachel Abenavoli,
New York University, United States

*CORRESPONDENCE
Hui Zhao
✉ hzz23@psu.edu

RECEIVED 30 August 2022
ACCEPTED 05 April 2023
PUBLISHED 27 April 2023

CITATION
Zhao H, Lei P-W, Hart SC, DiPerna JC and Li X (2023) When is universal SEL effective under authentic conditions? Using LPA to examine program implementation in elementary classrooms.
Front. Educ. 8:1031516.
doi: 10.3389/feduc.2023.1031516

When is universal SEL effective under authentic conditions? Using LPA to examine program implementation in elementary classrooms

Hui Zhao[1]*, Pui-Wa Lei[1], Susan Crandall Hart[1], James Clyde DiPerna[1] and Xinyue Li[2]

[1]Department of Educational Psychology, Counseling, and Special Education, The Pennsylvania State University, University Park, PA, United States, [2]Cambium Assessment, Washington, DC, United States

As universal social-emotional learning (SEL) programs have become more common in K-12 schools, implementation practices have been found to affect program quality. However, research examining how multiple facets of program implementation interrelate and impact student outcomes, especially under routine conditions in schools, is still limited. As such, we used latent profile analysis (LPA) to examine implementation of a brief universal SEL program (Social Skills Improvement System SEL Classwide Intervention Program) in primary classrooms. Three latent profiles of implementation were identified based on dosage, adherence, quality of delivery, student engagement and teachers' impression of lessons. Although results suggested that classrooms with moderate- and high-level implementation practices generally showed higher gains in student outcomes than those with low-level implementation, these differences did not reach statistical significance except for academic motivation. Implications for school-based universal SEL program planning, implementation, and evaluation are discussed.

KEYWORDS

implementation fidelity, social-emotional learning (SEL), contextual factors, latent profile analysis, program quality, elementary classrooms, authentic conditions

1. Introduction

School-based universal social-emotional learning (SEL) aims to teach foundational social-emotional skills at the classroom or school level as a public health approach to improving student outcomes through prevention and promotion programming (Greenberg et al., 2017). Although studies demonstrate that SEL programs can increase student prosocial skills and attitudes, lower problem behaviors, and enhance academic performance (Durlak et al., 2011; Taylor et al., 2017), implementation is an important determinant for the success of such programs in practice (Durlak and DuPre, 2008; Low et al., 2016). When programs with prior evidence of efficacy are not implemented as intended in the real world, schools may lose time, money, and ultimately, the desired benefits for students (Sanetti and Collier-Meek, 2019).

Through implementation science, dimensions of program implementation have been identified, which can serve as indicators of how evidence-based interventions translate into real-world delivery systems and capture variation across teachers and contexts (Proctor et al.,

2011). Measuring and understanding these dimensions is a critical part of effectiveness trials in education, which allows decision-makers to evaluate the feasibility of intervention programs and the degree to which implementation practices affect program outcomes (O'Donnell, 2008). To date, there have been very few effectiveness trials of social-behavioral programs in authentic education settings (Chhin et al., 2018); therefore, little is known about how universal SEL programs are typically implemented under these conditions, including what contextual factors are associated with variation in implementation and what combination of approaches yield the most benefit to students in the real world. In addition to aiding educators in planning for resource allocation, training, and support for their local context, such information could also be helpful to intervention developers hoping to increase the feasibility and transportability of programs.

There is growing awareness that implementation is critical to understanding not just *if* universal SEL works for students and schools, but *how* and *under what conditions* it works (Jones et al., 2019). As implementation is widely considered a key influence on outcomes of prevention and intervention programs (Durlak and DuPre, 2008; Domitrovich et al., 2010; Durlak et al., 2011), SEL researchers are similarly moving away from examining *whether* implementation matters to focus on *what aspects* of implementation are most salient (Low et al., 2016). While implementation is multi-dimensional, implementation fidelity, which refers to the degree to which a program is delivered as intended by developers, has been studied most frequently in education science (Proctor et al., 2011; Durlak, 2016). Five aspects of implementation fidelity (i.e., dosage—amount delivered; adherence—components delivered as planned; implementation quality—competence in delivery; participant responsiveness—engagement/enthusiasm; and program differentiation—critical features that distinguish the program) have been identified as critical for achieving implementation that influences student outcomes (Dane and Schneider, 1998; Durlak and DuPre, 2008; Durlak, 2016).

A systematic review of 41 school-based mental health intervention studies found that aspects of implementation fidelity were positively associated with student outcomes 36% of the time; participant responsiveness most commonly (58%) related to intervention effects, with dosage (39%), adherence (28%), and quality (23%) observed less often (Rojas-Andrade and Bahamondes, 2019). These aspects are also thought to interact with each other within the context of systemic barriers and supports to promote benefits for students. For example, Carroll et al. (2007) proposed a conceptual model where interventions improve outcomes via adherence, with intervention complexity, implementation support, quality of delivery, and participant responsiveness moderating this relationship.

Individual efficacy studies of universal SEL programs have sought to identify and untangle the most important aspects of implementation for optimizing program outcomes, but findings have been mixed. For example, in one study of a preschool readiness program, student engagement and implementation quality were related to improvement in student outcomes, but program dosage was less salient (Domitrovich et al., 2010). Low et al. (2014) found that, while adherence did not influence program effects, higher classroom engagement in *Steps to Respect: A Bullying Prevention Program* (Committee for Children, 2005) was associated with more desirable student outcomes (e.g., student attitudes toward bullying, student climate and support).

Reyes et al. (2012) found no main effects for teacher training, dosage, and implementation quality on student outcomes of *RULER Approach* (Brackett et al., 2011), however, they did find an interaction effect. A combination of high attendance during training and high dosage had a positive impact on student social-emotional skills, but only in conjunction with above-average implementation quality; when implementation quality was low, high training attendance and dosage actually predicted negative student outcomes. In Humphrey et al.'s (2018) study of *Promoting Alternate Thinking Strategies* (PATHS; Greenberg et al., 1995), program fidelity and reach did not predict changes in student behaviors, but higher implementation quality and participant responsiveness were associated with lower reports of student externalizing behaviors. Surprisingly, however, higher dosage was associated with lower SEL skills and prosocial behaviors (Humphrey et al., 2018). Overall, these studies suggest a complex relationship between implementation and program outcomes that appears to differ by aspects of implementation, interventions, and outcomes assessed.

To date, most studies of universal SEL programs have examined independent variable-centered relationships between aspects of implementation and program outcomes (Durlak and DuPre, 2008; Low et al., 2016). As Durlak (2016) noted, though, implementation components can interact with one another, and despite evidence suggesting better implementation is associated with more promising program outcomes, there is a lack of systematic investigation of how patterns of implementation factors relate to different program outcomes. While traditional variable-centered analyses (e.g., Analysis of Variance, correlation, and regression) have been widely used in many implementation studies, these methods may be insufficient for examining inter-relationships among multiple implementation measures (Laursen and Hoff, 2006; Hennessey and Humphrey, 2020). Furthermore, variable-centered approaches assume that the population is homogeneous in terms of how predictors impact outcomes (Laursen and Hoff, 2006; Cheng et al., 2021). As such, studies that employ variable-centered analytic approaches may fail to account for the heterogenous inter-relationships between implementation aspects and program outcomes within a population. Variable-centered approaches also might not be able to detect heterogeneity of small subpopulations when sample size is small (Muthén and Muthén, 2000). In contrast, person-centered analytic approaches, such as latent profile analysis (LPA) and latent class analysis (LCA), can be used to identify unobserved subgroups that share similar characteristics to examine population heterogeneity (Laursen and Hoff, 2006) without making assumptions (e.g., linear and homogenous relationship) that traditional variable-centered approach would (e.g., Magidson and Vermunt, 2005). Person-centered models are particularly well-suited to address questions regarding whether combination of multiple implementation variables have differential effects on program outcomes. Specifically, they are helpful when exploring combined effects of many predictors on outcomes given traditional moderation analyses have limited capability to test all possible interaction terms (Spurk et al., 2020).

Acknowledging the multi-faceted and multiplicative nature of implementation, the likely dynamic interplay between implementation aspects, and the lack of clarity about what matters most for school-based SEL, recent studies have identified person-centered profiles of implementation using multiple indicators and investigated their association with student outcomes. We located two studies that

explored how the dynamic interplay of multiple implementation measures affect outcomes of universal SEL programs. Low et al. (2016) examined how patterns of dosage, adherence, student engagement, and program differentiation of *Second Step* (Committee for Children, 2012) influenced student outcomes and found three latent classes: high-quality, low-engagement, and low-adherence. The analysis sample consisted of 160 teachers who implemented *Second Step* in kindergarten, first, or second grade. Low et al. (2016) used LPA to identify subgroups of implementation classrooms. All the implementation measures were based on teacher's weekly self-reports. Multilevel prediction models (i.e., students nested within classrooms) were adopted to examine the effect of profile membership on student gains in SEL competencies, functioning, and disruptive behavior while controlling for proactive classroom management and percentage of English Language Learners in the classrooms. Because only the low-engagement group showed associations with poorer student outcomes, Low and colleagues identified student engagement as the most pivotal aspect of implementation, though they acknowledged it was a necessary but not sufficient requirement for program success.

Using hierarchical cluster analyses of observational data, Hennessey and Humphrey (2020) also identified four profiles of implementation for PATHS (Greenberg and Kusche, 1993) based on adherence, quality, dosage, reach, and engagement. The analysis sample was composed of 45 schools that implemented the program for children 9–11 years of age. Multilevel linear models (i.e., students nested within classrooms) were used to examine the association between clusters and student outcomes. Hennessey and Humphrey did not find evidence linking the PATHS implementation profiles and student academic outcomes (e.g., reading, writing, and math) controlling for student gender, percentage of free/reduced-price lunch eligibility, and academic outcomes at baseline, however. The authors noted that their profiles were primarily differentiated by dosage, as the other aspects of implementation were relatively high and stable across their sample.

Given the limited number of studies that have utilized person-centered methods to examine universal SEL implementation patterns, it is difficult to synthesize their results. The two aforementioned studies (Low et al., 2016; Hennessey and Humphrey, 2020) differed in the methods of assessing implementation (teacher-report vs. observer-report), outcomes of the implementation (SEL competency vs. academic achievement), and covariates included in the prediction models. These factors may at least partially explain why the two studies yielded different profiles of implementation and patterns of association between profile membership and student outcomes. Both studies also treated implementation clusters or profiles as known groups and did not account for measurement error in classifications in their analysis of implementation outcomes. Notably, both studies were efficacy trials in which external (researcher-provided) support mechanisms such as training, coaching, and/or monitoring were in place to support implementation. To our knowledge, there currently are no published universal SEL studies that have used person-centered methods to examine implementation profiles using data from effectiveness trials (i.e., program conducted under routine conditions where implementation efforts are driven by schools and in accordance with their typical practices and available resources).

It is important to note that contextual factors at community, school and teacher levels also have been found to influence SEL program implementation (Durlak and DuPre, 2008; Anyon et al., 2016; Durlak, 2016). Domitrovich et al. (2008) proposed a conceptual framework that synthesized macro-level (e.g., policies and funding), school-level (e.g., resources, school climate and culture), and individual-level (e.g., teachers' buy-in, confidence in delivery, professional training) determinants of school-based implementation. Similarly, studies have indicated a consistent positive association between teacher buy-in and aspects of implementation. For example, Beets et al. (2008) found that school climate affected teachers' beliefs and attitudes toward a school-based prevention program, *Positive Actio*n (Flay et al., 2001), which in turn impacted implementation dosage and adherence. Anyon et al. (2016) similarly found that teachers' lack of buy-in, affected by their principal's commitment or time pressure for academic-focused instruction, was a barrier to program delivery. Finally, Domitrovich et al. (2019) observed that teachers with positive attitudes toward the program tended to deliver SEL lessons more frequently, and their perceptions of SEL culture predicted material usage and quality of delivery.

Understanding the individual, school, and community factors associated with profiles of school-based universal SEL implementation can provide insights into what may facilitate or impede teacher practices. In Low et al.'s (2016) study, profile membership did not differ by individual-level factors such as gender, race, ethnicity, education, or grade taught; however, a larger number of older/more-experienced teachers were in the low-adherence class, while younger/less-experienced teachers tended to be in the low-engagement class. Nevertheless, their study did not examine other potential influences at the classroom- or school-level. Similarly, Hennessey and Humphrey (2020) did not examine any relationships between profile membership and contextual factors.

Implementation appears to play a critical role in accounting for the variability of school-based SEL program outcomes (Reyes et al., 2012; Low et al., 2014, 2016); however, few studies have attempted to examine patterns of implementation that are associated with program outcomes and contextual factors under typical conditions in schools. The goal of the current study was to examine the implementation practices (e.g., adherence, dosage) of teachers implementing the Social Skills Improvement System SEL Classwide Intervention Program (SSIS SEL CIP; Elliott and Gresham, 2017), a universal program designed to improve students' prosocial skills and reduce problem behaviors (DiPerna et al., 2015). Using data from an effectiveness trial conducted under routine conditions, the primary aims of the study were to (a) examine if there are different patterns of program implementation and (b) determine if observed implementation patterns are associated with contextual factors and student outcomes. Our specific research questions included:

(1) Are there different profiles of SSIS SEL CIP implementation classrooms based on dosage, adherence, quality of delivery, student engagement, and teachers' impression of lessons?
(2) What are the contextual characteristics of the profiles of implementation?
(3) Are profiles of implementation associated with student outcomes?

Results of this study can broaden our understanding of how evidence-based programs realistically translate into schools and potentially provide insight into "what works best" for developing and delivering school-based SEL programming.

2. Materials and methods

2.1. Participants

The analysis sample for this study consisted of 41 first- and second-grade classrooms from 13 elementary schools in the South Atlantic, East North and West North Central regions of the U.S. The number of participating classrooms within an individual school ranged from 1 to 6 (*median* = 3). The racial composition of the analyzed student sample ($N = 354$) was approximately 44.9% white, 30.2% Black/African American, 22.0% Hispanic/Latine, 4.5% other, 3.1% Asian, <1% American Indian/Alaskan Native, and < 1% Native Hawaiian or Other Pacific Islander.[1] Most students (95.5%) spoke English as their primary language. At the time of baseline data collection, about 6.5% of students were receiving special education services, and 23.4% were receiving supplemental services (e.g., Title 1, reading support, tutoring). The analyzed teachers ($N = 41$) were predominantly female (i.e., 90.2%), white (i.e., 78%), and native English speakers (i.e., 90.2%). About 9.8% of teachers were Hispanic/Latine, 4.9% were Black/African American, and 2.4% were Asian, other, or unknown. Approximately half (46.3%) were teaching Grade 1, with 63.4% of participating teachers having a Bachelor's degree and 36.6% having a Master's degree. The sample reported 14.39 years of teaching experience on average, and 34.3% had specialized teaching certificates in addition to regular education.

2.2. Measures

2.2.1. Aspects of implementation

To assess teachers' program delivery, data were collected regarding five aspects of implementation: dosage, adherence, quality of delivery, student engagement, and teacher's weekly impression of lessons. When possible, observer report was used (i.e., adherence, quality, student engagement); dosage, adherence, student engagement, and lesson impressions were assessed with teacher report.

Dosage was assessed via weekly and end-of-year survey responses. Specifically, teachers indicated the lesson(s) they taught each week and reported the completion of lessons again at the end of the implementation period (which corresponded with the end of the school year). These data sources were cross-referenced to create two dosage indicators: the number of lessons taught from the "core" SSIS SEL CIP units (out of 30 lessons across 10 core units) and the total number of lessons implemented across all units (out of 69 total lessons across 23 units).[2] The average number of core lessons implemented across classrooms was 23.23 (SD = 5.90, range = 7–30), and the average number of total lessons implemented was 26.20 (SD = 8.50, range = 7–46).

Adherence includes four indicators. *Self-reported adherence* was teachers' rating of their own adherence to the lesson scripts on a 5-point scale (1 = *Not at all* to 5 = *Completely*). The composite score was the average across weeks. *Observed steps* measured the completion of steps described in the curriculum (i.e., tell, show/do, practice) via classroom observation. Each step was scored as 0 (*Non-occurrence*) or 1 (*Occurrence*), and the percentage of steps completed was averaged across lesson observations. *Observed adherence* reflected observers' ratings of the degree to which teachers followed the verbal script of the SSIS SEL CIP lessons using a 5-point scale (1 = *0–20%* to 5 = *81–100%*). *Observed level of implementation* assessed the extent to which teachers implemented the primary sections of each lesson (i.e., tell, show/do, practice, monitor progress, generalize) using a 5-point scale (1 = *Not at all* to 5 = *Completely*), and a composite score was averaged across sections. Interrater agreement was 92.15% for *observed steps*, 76.7% (90% with 1-point tolerance for disagreement) for *observed adherence*, and 62.92% (87.9% with 1-point tolerance) for *observed level of implementation*.

Observed quality of delivery was assessed during independent lesson observations. Specifically, observers completed five items that measured preparedness, enthusiasm, responsiveness to questions, clarity of presentation, and skill in facilitating activities using a 5-point scale (1 = *Not at all* to 5 = *Completely*). A composite score was computed for each lesson and then averaged across weeks. Interrater agreement was 72% (96.5% with 1-point tolerance).

Student engagement assesses students' active engagement, enthusiasm/interest, and understanding of lesson using teachers' and observers' report. Teachers responded to three questions regarding student engagement during weekly lessons using a 5-point scale (1 = *Very low* to 5 = *Very high*). Observers rated the same items after each lesson observation, and the interrater agreement on lesson observations conducted by two independent observers was 65% (95% with 1-point tolerance).

Weekly impression of lessons was measured via a single question using a 5-point scale (1 = *Poor*, 2 = *Fair*, 3 = *Acceptable*, 4 = *Very good*, 5 = *Excellent*) that teachers completed weekly throughout the implementation period. The item was "Overall, how would you rate the lessons you taught during the current week?" The composite score was averaged across weeks. This item was used to assess teachers' overall judgment of SSIS SEL CIP lessons that were taught during that week.

2.2.2. Student outcomes

Social skills and problem behaviors were rated by teachers using the social skills improvement system rating scales-teacher form (SSIS-RST; Gresham and Elliott, 2008). The *Social Skills* scale measures communication, cooperation, assertion, responsibility, empathy, social engagement, and self-control, whereas the *Problem Behavior* scale assesses externalizing, bullying, hyperactive-inattentive, internalizing, and autistic behaviors. Items in both scales use a 4-point format (0 = *Never* to 3 = *Almost always*). The SSIS-RST demonstrated sound psychometric evidence (e.g., $\alpha = 0.88–0.98$; DiPerna et al., 2018).

Approaches to learning were measured using teacher's ratings on the Academic Competence Evaluation Scales (ACES; DiPerna and Elliott, 2000). The *Motivation* scale assesses students' learning attitudes, persistence, and interest. The *Engagement* scale measures students' attention and participation in academic activities. Items on both scales were rated using a 5-point format (1 = *Never* to 5 = *Almost always*). Both subscales showed strong psychometric evidence (e.g., $\alpha = 0.95–0.98$; DiPerna et al., 2018). Composite scores for each scale were converted to item response theory (IRT) scores to ensure the equivalence of assessment at the pre-test and post-test.

1 Participants were allowed to endorse more than one category of racial/ethnical group.

2 See the section of procedures for a detailed description of the core units.

2.2.3. Multilevel contextual factors

School-, teacher-, and class-level demographic information were collected during the year of the study. School information included percent of children eligible for free/reduced-price lunch, percent of racial/ethnical minority children, school size, and location. Teachers provided information about their own demographic characteristics including gender, race/ethnicity, certification, educational level, primary language, and years of teaching experience.

The classroom assessment scoring system: kindergarten-third grade scale (CLASS K-3; Pianta et al., 2008) was used to evaluate the instructional climate of participating classrooms. Research staff observed the implementation classrooms and completed ratings in regard to emotional support, instructional support and classroom organization. Each item was rated by two observers on a 7-point scale (1 = *low* to 7 = *high*). Intraclass correlations for paired CLASS observations have been shown to be acceptable (0.65–0.76; DiPerna et al., 2018).

The *Teacher SEL Beliefs* scale includes four items reflecting teachers' comfort with teaching SEL, 4 items assessing commitment to SEL, and four items measuring perceived school culture relative to SEL (Brackett et al., 2011). The internal consistency for each subscale was good ($\alpha \geq .80$). The assumptions supporting social-emotional teaching (*ASSET)* scale assesses teachers' beliefs about the degree to which SEL skills are malleable, compatible, and influential (Hart, 2021). The internal consistency for *ASSET* subscales and total score were satisfactory ($\alpha \geq .87$).

Lastly, teachers were asked on an end-of-year questionnaire to indicate their opinion about the percent of school time should be allocated to facilitate academic or SEL skills. The question was "What percentage of early elementary students' (Grades 1–2) school time should be focused on each domain?" and included response options for academic subject areas (reading, math, etc.) as well as SEL. The total percentage across all the domains were required to sum to 100%.

2.3. Procedure

The larger effectiveness trial was approved by the university's Institutional Review Board. Consistent with the goal of testing the effectiveness of the SSIS SEL CIP, districts that were already considering adopting a universal SEL program in the early grades as part of their typical practice were recruited to participate. With the goal of reaching geographically diverse school sites, information about the trial was distributed through online platforms and national professional networks. School representatives who requested more details were provided with additional information through individual conversations. Prior to enrolling a school into the study, we sought and received permission to conduct the research with administration according to district guidelines. In addition, active teacher and parent consent was obtained prior to data collection. Schools were randomly assigned to treatment condition such that, within each school, the SSIS CIP SEL was taught in either first- or second-grade classrooms while the other grade levels maintained business-as-usual practices. Data from classrooms assigned to the treatment condition were used for this study.

The SSIS SEL CIP includes 10 core instructional units and 13 advanced units (3 lessons per unit) that focus on social and emotional skills that a national sample of U.S. teachers identified as important to student success (e.g., listening to others, paying attention to your work, asking questions). Each SSIS SEL CIP lesson requires approximately 25–30 min and features multiple instructional phases (i.e., telling, showing, doing, practicing, monitoring progress, and generalizing) to promote skill acquisition and generalization. Materials include a teacher guide with scripted lesson plans, brief video clips that demonstrate examples and non-examples of social behaviors, scenarios describing common classroom scenarios for role plays, and cue card with emotion emojis. Because the goal of the larger project was to examine student outcomes under typical conditions (levels of support) and practices, teachers and schools decided how much, how often, and in what way to plan for and deliver SSIS SEL CIP units. Most teachers (79.5%) reported that they planned for implementation by preparing on their own, 38.5% reported planning with colleagues, and only about 12.8% attended a training conducted by their school or district.

Teachers' self-report and independent direct observations were completed to measure implementation fidelity of the SSIS SEL CIP. Teachers completed weekly surveys via online questionnaires to report the number of lessons completed, and rate the degree of adherence to the curriculum, quality of delivery, and student engagement. Teachers were compensated for their time spent completing questionnaires. Trained research assistants completed direct observations of implementation classrooms and rated teacher's adherence, quality of delivery, and student engagement. The average number of lessons observed per teacher was 5.40, with a minimum of three and a maximum of seven. Eighteen percent of observations were completed by paired observers. Student measures were completed immediately before and after program implementation.

The study duration was one school year. The SSIS SEL CIP materials were provided to the implementing grade levels prior to the baseline data collection window and then shared with the control classrooms at the end of post-test data collection.

2.4. Statistical analysis

The first step was to conduct a latent profile analysis (LPA) on implementation measures (i.e., dosage, adherence, quality, student engagement, and weekly impression of lessons) to explore if there were different patterns of classroom implementation of the SSIS SEL CIP. LPA is a statistical approach for identifying latent subgroups based on a set of observed indicators (Ferguson et al., 2020). LPA models can produce estimates of membership probability for each participant so that individuals sharing the same pattern of indicators are categorized into the same underlying class (Spurk et al., 2020). In this study, LPA was used as an exploratory tool to identify profiles in which classrooms shared similar patterns of implementation.

Multiple models were fitted to generate 1 to 4 latent profiles. Each model was compared against the previous one [i.e., (k-1) profiles] according to Akaike's Information Criterion (AIC), Bayesian Information Criterion (BIC), Sample-size Adjusted BIC (SABIC), entropy, LMR (Lo-Mendell-Rubin likelihood ratio test), and VLMR (Vuong-Lo-Mendell-Rubin likelihood ratio test). A smaller value of AIC, BIC, SABIC and a higher entropy indicate that a model fits better to the data (Ferguson et al., 2020). LMR and VLMR were used to test whether the model with [k] profiles fits better than the one with [k-1]

profiles. A non-significant test result suggests that the more parsimonious model is better fitting (Smith et al., 2021). We selected the final model based on all fit indices and interpretability of the profiles.

After the number of latent profiles were determined, the second step was to examine the association between the profile membership and contextual variables to identify ones that differed across profiles. Profile membership was determined based on the probability estimates produced by the solution of the chosen number of profiles (i.e., classrooms were assigned to the profile to which they had the highest probability of belonging). Contextual variables included school-, teacher-, and classroom-level demographic characteristics. Chi-square tests were conducted to examine whether profiles were associated with categorical factors (e.g., school location, teacher's certification, educational level). One-way ANOVA was used to test mean differences across profile for continuous variables (e.g., classroom environment, SEL belief, years of teaching). Due to the large number of contextual variables and the small classroom-level sample size by profile, judicial selection of contextual factors was necessary. We used this approach to identify statistically significant factors that should be considered for the next step of distal outcome analyses.

The third step was to explore the associations between implementation profile membership and outcomes of social skills, academic motivation, engagement, and problem behaviors controlling for contextual factors. We used auxiliary regression models for this purpose because they allowed auxiliary variables (contextual factors) to predict both profile membership and distal outcomes. Specifically, the manual BCH (Bolck et al., 2004) approach in Mplus was used because it could preserve profile membership and account for measurement error (Asparouhov and Muthén, 2014). Clustering by schools was accounted for by adjusting standard errors in Mplus using the cluster and "analysis = complex" commands. Robust maximum likelihood estimator (MLR) was used to estimate model parameters due to its robustness against violation of assumptions (e.g., normality) and ability to generate less biased estimates of parameters in comparison with the traditional maximum likelihood approach (Bakk and Vermunt, 2016).

A common-slope model was specified for each of the outcome variables to hold contextual factors constant across profiles. We fitted the model by constraining the regression slopes for covariates to be equal across profiles to estimate adjusted profile mean outcome differences (i.e., control for covariates). Contextual covariates included grade level, percentage of supplemental educational services, teaching experience, classroom organization, instructional support, and percentage of free/reduced-price lunch eligibility. These contextual variables were selected because they varied significantly across profiles and were included as covariates in prior implementation studies (Domitrovich et al., 2010; Reyes et al., 2012; Low et al., 2014, 2016; Cross et al., 2015; Humphrey et al., 2018). Even though percentage of students receiving supplemental services at classroom level was rarely examined in implementation studies, a previous efficacy trial of the SSIS-CIP (DiPerna et al., 2018) suggested that supplemental services had a negative effect on social skills and academic motivation. Therefore, it was included as a covariate in predicting outcome gains.

To account for baseline difference in student outcome variables, we calculated reliable change (RC) scores that reflect how much change occurred during the implementation period. The RC scores were computed by subtracting the pre-test scores from the post-test and then dividing the difference scores by the standard error of the difference (Jacobson and Truax, 1991). If the RC score is larger than critical values (e.g., $z = +/-1.96$ at α level of 0.05), the pre-post change is considered statistically reliable (Ferguson and Splaine, 2002). The sign of RC scores indicates the direction of change from pre-test to post-test. The RC score was used in the analysis because, first, we were interested in within-person change rather than relative change; second, it facilitated interpretation by accounting for standard error of measurement (i.e., whether the amount of pre-post change was reliable or due to random error) and was commonly used to decide clinical significance in mental and behavioral health (Ferguson and Splaine, 2002).

It is important to emphasize the exploratory nature of these analyses given the relatively small overall class-level sample size ($N = 41$) limits statistical power in detecting profile differences in both contextual factors and outcome gains. As such, the relevant results should be interpreted accordingly.

3. Results

3.1. Latent profiles of implementation classrooms

Descriptive statistics and intercorrelations for the implementation measures are presented in Table 1. There was a salient difference in variability of teacher's self-report and direct observation in terms of adherence and student engagement. It appears that teacher self-reported ratings of adherence were moderately associated with direct observations of adherence, with Pearson's r of 0.38. However, student engagement reported from teachers and their weekly impressions of lessons were weakly associated with all observer-reported implementation measures ($|\text{Pearson's } r| < 0.10$).

To address the first research question, models with 1–4 latent profiles were fitted. Table 2 presents the fit statistics for each of the models. The 4-profile solution had a non-positive definite matrix and was difficult to interpret, therefore it was eliminated from further consideration. The entropies of the 2- and 3-profile solutions were equivalent (entropy = 0.974), which indicated a high classification certainty for both models (Ferguson et al., 2020). Smaller values of AIC, BIC and SABIC suggested the 3-profile solution was fitting better than the 2-profile. The difference in BIC (i.e., $\Delta\text{BIC} > 10$) also provided strong evidence in support of the 3-profile model (Raferty, 1995). However, the non-significant results of VLMR and LMR tests suggested the 2-profile solution was better fitting ($p > 0.05$) than the 3-profile. Given the mixed results from multiple indices, we decided to retain the 3-profile solution because it provided more interpretive information.

Table 3 displays the descriptive data for the implementation measures by three latent profiles. Original scores from each measure were standardized to ease the interpretation of results across measures. One-way ANOVA and Tukey's HSD were conducted to compare the average variable scores across profile. Measures with asterisks indicate significant differences across profiles, and superscripts (i.e., a, b, and c) indicate significant pairwise differences. It suggested, for example, means of observer-reported adherence, quality of delivery and student engagement for Profile 1 were significantly lower than the other two

TABLE 1 Summary of descriptive statistics and intercorrelations for the implementation measures.

Measure	*M* (SD)	Min	Max	1	2	3	4	5	6	7	8	9
Direct observation												
1. Steps	0.74 (0.20)	0.16	0.99									
2. Adherence	3.89 (1.00)	1.00	5.00	0.90***	–							
3. Level of implementation	3.80 (0.88)	1.58	5.00	0.96***	0.86***	–						
4. Quality of delivery	4.26 (0.57)	2.25	5.00	0.82***	0.68***	0.87***	–					
5. Student engagement	4.29 (0.50)	3.22	5.00	0.58***	0.52***	0.65***	0.80***	–				
Teacher report												
6. Number of core lessons	23.27 (5.83)	7	30	0.45**	0.36*	0.45**	0.47**	0.40*	–			
7. Number of total lessons	26.20 (8.39)	7	46	0.35*	0.25	0.36*	0.42**	0.35*	0.77***	–		
8. Adherence	3.87 (0.42)	3.00	4.89	0.34*	0.38*	0.32*	0.18	0.02	0.19	0.30	–	
9. Student engagement	3.66 (0.40)	2.93	4.47	−0.02	−0.07	−0.03	−0.08	0.00	−0.16	−0.17	0.30	–
10. Impression of lessons	3.57 (0.45)	2.73	4.38	0.03	0.06	−0.05	−0.09	−0.01	−0.27	−0.18	0.19	0.72***

$N = 41$. Scores of the implementation measures were on the original scale.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

TABLE 2 Latent profile analysis (LPA) model fit summary and profile proportions.

Model	LL	AIC	BIC	SABIC	Entropy	VLMR	LMR	Profile proportions
						p-value	*p*-value	
1	−576.70	1193.41	1227.68	1165.07				
2	−502.83	1067.65	1120.77	1023.73	0.974	0.0925	0.0971	0.29, 0.71
3	−462.92	1009.84	1081.81	950.33	0.974	0.3586	0.3664	0.12, 0.37, 0.51
4	−445.06	996.12	1086.94	921.03	0.968	0.7932	0.7975	0.24, 0.12, 0.17, 0.46

LL, log likelihood; VLMR, The Vuong-Lo-Mendell-Rubin likelihood ratio test between [*k*-1] profiles and [*k*] profiles; LMR, The Lo-Mendell-Rubin adjusted likelihood ratio test between [*k*-1] profiles and [*k*] profiles.

profiles, and Profiles 2 and 3 differed by one standard deviation on most of the measures. As such, the 3-profile solution provided more information of implementation patterns compared to the 2-profile solution which only identified High and Low profiles.

It is noticeable that the profile sizes found in the 3-profile solution were small, especially for the smallest profile ($N = 5$). However, profile separation or the distance between profiles dictates the sample size required and hence power to detect the correct number of profiles in LPA (e.g., Tein et al., 2013; Ferguson et al., 2020). The high entropy of 0.974 for the 3-profile model suggests a high degree of profile separation certainty. Also as shown in Figure 1, the three profiles varied on the majority of the implementation measures, except for teachers' report of student engagement and weekly impression of lessons.

Based on the 3-profile solution, five classrooms (12%) were classified to the first profile in which teachers completed fewer lessons with less adherence and quality, and students showed lower level of engagement. Therefore, we characterized this group of classrooms as the "Low Implementation" Profile (Low IP). The second profile included 15 classrooms (37%) that reported lower dosage but moderate adherence, quality, and student engagement. Notably, the direct observation and teachers' report yielded somewhat inconsistent ratings of adherence and engagement. Due to the small variability of teacher-reported ratings, we focused on the observational ratings and labeled this group as the "Moderate Implementation" Profile (Moderate IP). Twenty-one classrooms (51%) in the third profile completed the most lessons and had the highest ratings on all observational assessments, thus we labeled them collectively as the "High Implementation" Profile (High IP).

3.2. Contextual characteristics of the three implementation profiles

To address the second research question, we examined the contextual characteristics of the three profiles at the school, class,

TABLE 3 Descriptive statistics of the implementation scores by profile.

Measure	Mean (SD)		
	Profile 1	Profile 2	Profile 3
	Low IP	Moderate IP	High IP
	(*N*=5)	(*N*=15)	(*N*=21)
Direct observation			
Steps*	−1.92 (0.62)	−0.45 (0.40)[a]	0.78 (0.33)[b]
Adherence*	−2.03 (0.81)	−0.19 (0.53)[a]	0.62 (0.46)[b]
Level of implementation*	−1.89 (0.40)	−0.47 (0.39)[a]	0.78 (0.39)[b]
Quality of delivery*	−1.55 (1.26)	−0.45 (0.52)[a]	0.69 (0.48)[b]
Student engagement*	−1.16 (0.94)	−0.36 (0.86)	0.53 (0.76)[b]
Teacher report			
Number of core lessons*	−0.53 (1.13)	−0.63 (1.04)	0.57 (0.52)[bc]
Number of total lessons*	−0.64 (0.81)	−0.41 (1.06)	0.45 (0.80)[b]
Adherence*	−1.14 (0.84)	0.15 (1.04)[a]	0.16 (0.86)[b]
Student engagement	0.06 (1.10)	−0.06 (1.10)	0.03 (0.95)
Impression of lessons	−0.10 (0.82)	0.05 (1.00)	−0.01 (1.08)

N = 41. Standard deviations are presented in the parenthesis. Scores of the implementation measures were on a *z*-score scale. Cell means on the same row were compared using One-way ANOVA and Tukey's HSD. [a]Significant difference was found between Profile 1 (Low IP) and Profile 2 (Moderate IP).
[b]Significant difference was found between Profile 1 (Low IP) and Profile 3 (High IP).
[c]Significant difference was found between Profile 2 (Moderate IP) and Profile 3 (High IP).
*$p < 0.05$.

and teacher levels (Table 4). Results suggested that the 5 Low IP classrooms were all in the second grade and from two schools located in suburban districts (in the South Atlantic and East North Central regions of the United States). They had the smallest percentage of students receiving free/reduced lunch (37.80%), students of color (53.40% at school level; 38.25% at class level) and students receiving supplemental services (14.35%). Nevertheless, they had the highest percentage of students receiving special educational services (12%). Three teachers (60%) were female, two teachers (40%) received both regular and special education certificates, and one teacher (20%) was racial/ethnical minority. Teachers had the greatest amount of prior teaching experience ($M = 20.60$, SD = 9.34), and they tended to believe that more school time should be allocated to foster academic skills (55% of the school day) than social-emotional skills (11.25%).

The 15 Moderate IP classrooms were from eight schools distributed across rural, suburban, and urban districts. These classrooms were located in schools serving a significantly higher percent of students (86.79%) receiving free/reduced-price lunch and a marginally significantly higher percent of students of color (73.36% at school level; 72.24% at class level). These classrooms also were observed to have the lowest levels of instructional support and classroom organization. Only one teacher (7%) was male, four teachers (27%) received regular and other credential, and four teachers (27%) were racial/ethnical minority.

The 21 High IP classrooms were from 10 schools with diverse student populations (53.76% Students of color at school level; 51.64% at class level) where nearly 6 in 10 students who were eligible for free/reduced lunch (58.71%). One teacher (5%) was male, three teachers (14%) were racial/ethnical minority and eight teachers (38%) received other credential along with regular educational certificate. The average years of teaching experience was 13.86, which was close to the Moderate IP but much lower than the Low IP classrooms. These classrooms demonstrated significantly higher levels of instructional support and classroom organization relative to the Moderate IP classrooms.

Significant associations were found between profiles and grade ($\chi^2 = 9.29, df = 2, p = .01$), school location ($\chi^2 = 18.31, df = 6, p = .01$) and teachers' primary language ($\chi^2 = 7.68, df = 2, p = .02$). However, results must be interpreted with caution because there were a few zero counts in the frequency tables. Significant mean differences were only detected for percentage of students receiving free/reduced-price lunch, classroom organization, and instructional support. We also examined the distribution of classrooms by schools across profiles and found that each profile was represented by more than one school (2, 8, and 10 schools in Profiles 1, 2, and 3, respectively). Selected covariates were included to predict profile membership using the BCH approach. Profile classification generated from the BCH approach was consistent with the results from one-step LPA (i.e., 3-profile solution). Auxiliary regression analysis was adopted to examine the association between covariates and profile. Results indicated that only percentage of free/reduced-price lunch eligibility was marginally associated with profile membership. Classrooms with higher percentage of free/reduced-price lunch eligibility were more likely to be assigned to Moderate or High IPs compared to Low IP.

3.3. Association between profile membership and student gains

We used the probability of profile membership to predict student gains in social skills, problem behaviors, academic motivation, and engagement after controlling for grade level, teaching experience, percentage of students receiving supplemental services, percentage of free/reduced-price lunch eligibility, classroom organization, and instructional support. School-, teacher- and class-level covariates were included in the auxiliary regression models to account for the nonequivalence across profiles and their potential influence on student gains suggested by previous studies (e.g., Domitrovich et al., 2010; Cross et al., 2015).

The auxiliary regression model results are shown in Table 5 by outcome variables. Intercepts 1, 2, and 3 represent the adjusted mean of gains for Profiles 1, 2, and 3, respectively, when covariates were kept constant. In examining the direction and relative magnitude of gains shown in Table 5, the Moderate IP classroom mean gains appeared highest in social skills, academic motivation, and academic engagement, followed by the High IP classrooms. The Low IP classrooms appeared to show reductions in social skills, academic motivation, and academic engagement over time. High IP classrooms demonstrated a reduction in problem behaviors, while both Moderate

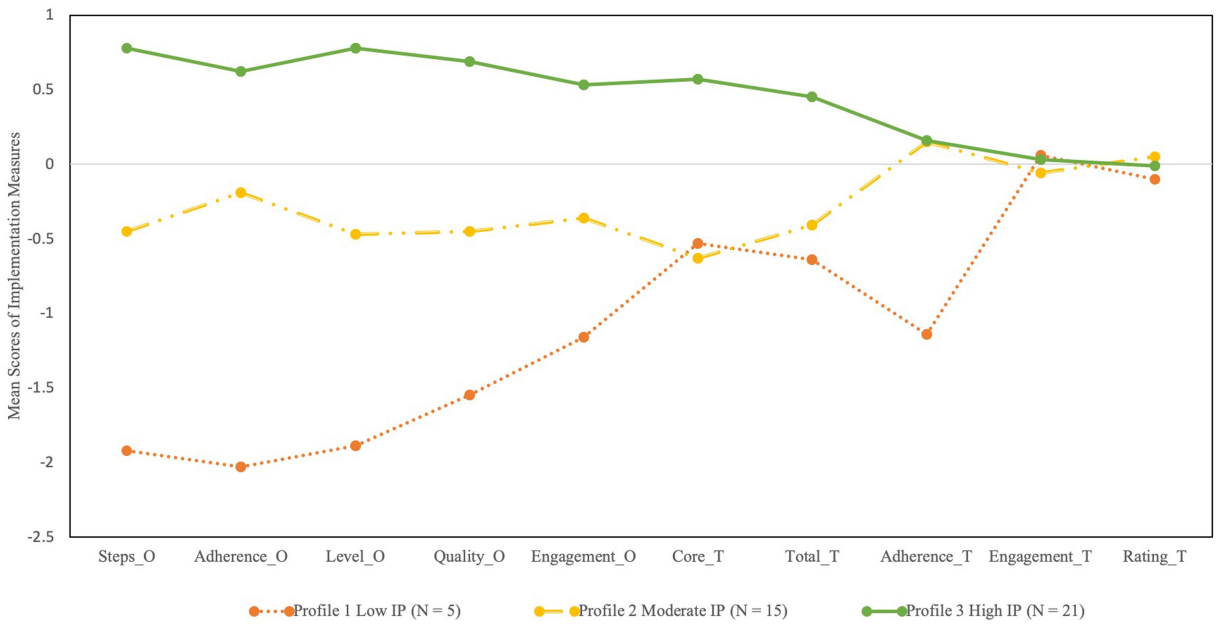


FIGURE 1
Mean scores of implementation measures by profile. Steps_O, Observed steps; Adherence_O, Observed adherence; Level_O, Observed level of implementation; Engagement_O, Observed student engagement; Core_T, Teacher reported number of core lessons taught; Total_T, Teacher reported number of total lessons taught; Adherence_T, Teacher reported adherence; Engagement_T, Teacher reported student engagement; Rating_T, Teacher's weekly impression rating of lessons.

and Low IP classrooms showed an increase. It is important to note that these observations about the means are descriptive; we discuss statistically significant differences among profiles using confidence intervals in the next section.

The estimates of the mean gains for each outcome with 95% confidence intervals are shown in Figure 2. Intervals for adjusted means that did not overlap indicated significant difference between the adjusted means. There were no statistically significant differences in adjusted gains of social skills, problem behaviors, or academic engagement across profiles. However, the Moderate [95% CI = (0.20, 0.82)] and High IP classrooms [95% CI = (−0.25, 0.53)] showed significantly higher gains in academic motivation compared to Low IP [95% CI = (−1.55, −0.41)], even though no significant difference was found between the Moderate and High IP classrooms. Across profiles, second grade classrooms had a significantly higher gain in social skills ($b = 0.38$, $SE = 0.13$, $p = 0.003$), and classrooms with higher classroom organization tended to gain more in social skills ($b = 0.35$, $SE = 0.11$, $p = 0.001$). Percentage of free/reduced-price lunch eligibility appeared to be negatively associated with gains in academic motivation ($b = -0.86$, $SE = 0.37$, $p = 0.022$) and engagement ($b = -0.89$, $SE = 0.39$, $p = 0.021$). Results also suggested that instructional support was negatively associated with student gains in academic engagement ($b = -0.32$, $SE = 0.12$, $p = 0.009$).

4. Discussion

The study identified three latent profiles of implementation of the SSIS SEL CIP under routine implementation conditions based on dosage, adherence, quality of delivery, student engagement, and weekly impression of lessons. The Low IP included about 12% of the classrooms and was characterized by low dosage, low fidelity to instructional scripts, less quality of implementation, and lower student engagement in the program lessons. The Moderate IP was comprised of about 37% of the classrooms; these classrooms demonstrated lower dosage but average adherence, quality of delivery, and engagement. About half of the classrooms fell into the High IP in which teachers delivered a greater number of lessons with high quality, the program was implemented with high adherence, and students appeared to be engaged to a high degree in the lessons.

It is important to note that the three profiles were labeled primarily based on observational data instead of teachers' self-report. Data collected directly from teachers (e.g., student engagement, impression rating) demonstrated limited variability. Social desirability may have played a role in the limited variability; however, one additional explanation is that the question used to solicit teachers' overall reflection on the lessons taught during that week was fairly broad and could have been interpreted differently across teachers, with some potentially rating the quality of curriculum and others potentially rating the quality of delivery (or some combination of the two). Student engagement reported by teachers and their impression rating of the lessons were also weakly associated with observer-reported implementation measures. A possible explanation could be that teachers and observers had different perceptions of implementation, particularly how students were engaged in SEL class activities. Teachers provided their perceptions of implementation practice retrospectively at the end of each week, whereas observers provided their ratings on individual lessons they observed in real-time. In their seminal meta-analysis, Durlak et al. (2011) noted that observational data are more objective and appear to be more correlated with program outcomes than teacher report data, and therefore recommended to use direct observation for implementation analysis

10.3389/feduc.2023.1031516

TABLE 4 Contextual characteristics of the three latent profiles.

Measure	Profile 1 Low IP (N =5)	Profile 2 Moderate IP (N =15)	Profile 3 High IP (N =21)
School-level demographic (%)			
Free/reduced-price lunch*	37.80 (39)	86.79 (19)[a]	58.71 (30)[b]
Ethnic/racial minority	53.40 (36)	73.36 (30)	53.76 (34)
School location (%)**			
Rural (mid-size)	0	20.00	33.33
Suburban (large)	100.00	26.67	52.38
Urban (mid-size)	0	13.33	14.29
Urban (large)	0	40.00	0
Geographic region (%)			
South Atlantic	60.00	53.33	28.57
East North Central	40.00	26.67	38.10
West North Central	0	20.00	33.33
Teacher-level demographic (%)			
Female	60.00	93.30	95.20
Ethnic/racial minority	20.00	26.67	19.00
Bachelor's degree	80.00	73.33	52.38
Master's degree	20.00	26.67	47.62
English as primary language*	100.00	73.33	100.00
General education teacher	100.00	80.00	95.24
General & Special education teacher	0	13.33	0
Teaching experience (yrs.)	20.60 (9.34)	13.07 (10.69)	13.86 (9.09)
Classroom-level demographic (%)			
Grade 1*	0	73.33	38.10
Grade 2*	100	26.67	61.90
Special educational services	12.00 (15.54)	9.36 (9.33)	8.25 (9.17)
Supplemental educational services	14.35 (22.04)	39.81 (39.25)	18.13 (15.60)
English language learners	6.33 (11.30)	6.02 (9.44)	11.41 (13.25)
Ethnic/Racial minority	38.25 (36.66)	72.24 (34.01)	51.64 (32.66)
Classroom environment			
Emotional support	5.40 (0.66)	5.23 (0.98)	5.33 (1.21)
Class organization*	5.10 (0.64)	4.82 (1.02)	5.72 (1.05)[b]
Instructional support*	2.00 (0.90)	1.73 (0.75)	2.85 (1.11)[b]
Teacher's belief in SEL			
SEL comfort	4.05 (0.60)	4.05 (0.54)	3.89 (0.72)
SEL commitment	4.15 (0.49)	4.37 (0.55)	4.20 (0.56)
SEL culture	3.67 (0.94)	4.00 (0.65)	3.76 (0.86)
SEL malleable	4.46 (0.41)	4.32 (0.37)	4.49 (0.36)
SEL compatible	3.88 (0.97)	4.26 (0.67)	4.19 (0.53)
SEL influential	4.50 (0.46)	4.50 (0.47)	4.54 (0.49)
School time allocated (%)			
Social-emotional skills	11.25 (2.50)	21.79 (7.49)	18.33 (8.99)
Academic skills	55.00 (16.83)	45.00 (10.19)	45.24 (15.85)

N = 41. Standard deviations are presented in parenthesis. Cell means on the same row are compared using One-way ANOVA and Tukey's HSD.
[a]Significant difference was found between Profiles 1 (Low IP) and Profile 2 (Moderate IP).
[b]Significant difference was found between Profiles 2 (Moderate IP) and Profile 3 (High IP).
*$p < 0.05$; **$p < 0.01$.

TABLE 5 Common-slope model estimates by outcome variables.

Effect	Social skills			Problem behaviors			Motivation			Engagement		
	B	SE	95% CI	*B*	SE	95% CI	*B*	SE	95% CI	*B*	SE	95% CI
Intercept 1	−0.58	0.56	[−1.67, 0.52]	0.13	0.72	[−1.28, 1.54]	−0.98	0.29	[−1.55, −0.41]	−0.56	0.47	[−1.48, 0.36]
Intercept 2	0.57	0.12	[0.33, 0.81]	0.27	0.26	[−0.24, 0.78]	0.51	0.16	[0.20, 0.82]	0.33	0.08	[0.17, 0.49]
Intercept 3	0.19	0.19	[−0.18, 0.56]	−0.18	0.14	[−0.45, 0.09]	0.14	0.20	[−0.25, 0.53]	0.22	0.17	[−0.11, 0.55]
Grade	0.38**	0.13		0.49	0.26		0.25	0.24		0.06	0.13	
Experience	0.00	0.01		0.01	0.02		0.01	0.01		−0.004	0.01	
% SUPED	−0.20	0.16		0.65	0.39		0.02	0.22		−0.45	0.24	
CO	0.35**	0.11		0.03	0.20		0.09	0.09		0.15	0.12	
IS	−0.13	0.24		−0.08	0.20		−0.17	0.09		−0.32**	0.12	
% SFR	−0.02	0.81		−0.36	0.86		−0.86*	0.37		−0.89*	0.39	
AIC	161.28			169.99			151.03			153.31		
BIC	205.19			213.91			194.94			197.22		
SABIC	123.83			132.54			113.58			115.86		

$N = 40$. Intercept 1 = adjusted means of outcomes for the Low IP classrooms; Intercept 2 = adjusted means of outcomes for the Moderate IP classrooms; Intercept 3 = adjusted means of outcomes for the High IP classrooms. Mean outcomes are adjusted for grade, teaching experience (yrs), percentage of students eligible for free/reduced-price lunch (% SFR), classroom organization (CO), instructional support (IS), and percentage of students receiving supplemental education service (% SUPED).
*$p < 0.05$; **$p < 0.01$.

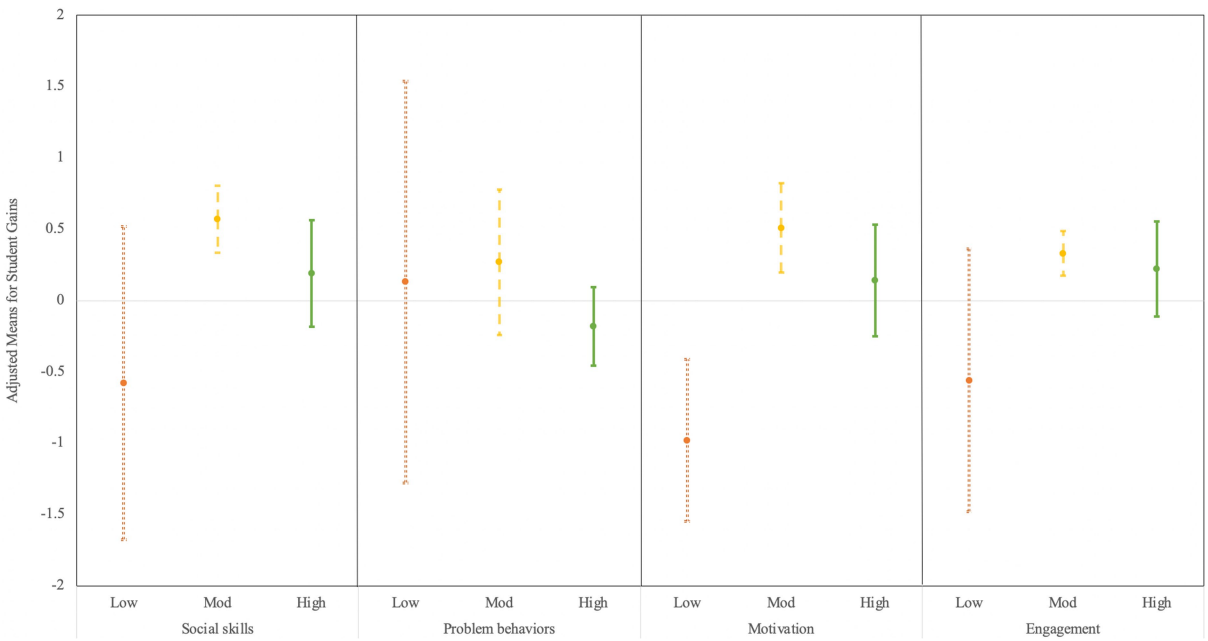


FIGURE 2
Auxiliary regression estimates of student gains with 95% CIs by Profile. Low, Low implementation profile; Mod, moderate implementation profile; High, high implementation profile. Mean outcomes are adjusted for grade, teaching experience (yrs), percentage of students eligible for free/reduced-price lunch (% SFR), classroom organization (CO), instructional support (IS), and percentage of students receiving supplemental education service (% SUPED).

if it is possible. One noticeable difference between our study and prior profile research is that we used different methods to assess implementation practices. Low et al. (2016) used teacher-reported rating, and their profile classification gave more weight to engagement and adherence than dosage and generalization. Hennessey and Humphrey (2020) adopted observational data, and the clusters they identified were primarily based on dosage. Our study used both teacher- and observer-reported rating of implementation and results

suggested that adherence, quality of delivery, and student engagement may play an integral role in differentiating profiles. Given the interpretation of program outcomes relies upon accurate assessment of implementation, there is a need for future research that incorporates and evaluates multiple methods to measure program implementation.

Regarding contextual factors, we found that the Moderate IP classrooms were comprised of a significantly higher percentage of students with free/reduced-price lunch eligibility. In addition, these classrooms were observed to have significantly lower instructional support and classroom organization than the High IP classrooms. Teachers also reported putting more daily instructional emphasis on social-emotional skills (21.79%) in the Moderate profile than the others. This suggests that teachers' emphasis on social-emotional skills across and throughout the school day and their interactions with students may facilitate the development of these skills even in the context of lower dosage of a formal SEL program. The Low IP teachers had the greatest amount of teaching experience and tended to place more daily instructional emphasis on academic skills (55%) than social-emotional skills (11.25%). Prioritizing academic instruction is perhaps the most likely reason why Low IP teachers completed fewer lessons with fidelity. This finding also supports the argument that time pressure for academic instruction may be a barrier to SEL program delivery in some schools and classrooms (Anyon et al., 2016). Even though prior literature suggests school climate and teacher's SEL beliefs may affect implementation (Durlak and DuPre, 2008; Durlak, 2016), our current study did not yield evidence to support this hypothesis. There are at least two possible explanations. First, the small sample size and resulting lower levels of statistical power may have been insufficient to detect significant differences; second, teacher's self-reported ratings may have been affected by social desirability, resulting in response patterns that were more positive than reality (Holden and Passey, 2009).

In investigating the association between profile membership and student gains, we found only one statistically significant difference. Moderate and High IP classrooms demonstrated higher gains in students' academic motivation compared to Low IP classrooms and controlling for contextual factors. Although not reaching a threshold of statistical significance, based on the adjusted means, the Moderate and High IP classrooms appeared to have more positive gains in social skills and academic engagement, with Moderate IP showing the most benefit. The Low IP classrooms, however, had negative gains in all of the outcomes. One noteworthy consideration is that increasing dosage beyond a certain threshold may not necessarily relate to improved student outcomes resulting from implementing universal SEL, which is consistent with and extends prior efficacy research (Domitrovich et al., 2010; Reyes et al., 2012; Humphrey et al., 2018). For example, Hennessey and Humphrey (2020) found their profiles, which were primarily differentiated by dosage, did not appear related to students' academic outcomes and Low et al. (2016) suggested that aspects of implementation delivery that are harder to manualize (e.g., implementer competency and student engagement) are necessary for maximal program benefit. In this sample, with the same amount of lessons completed, the Moderate IP classrooms showed higher gains in academic motivation compared to the Low IP classrooms; however, teachers in the Moderate IP classrooms delivered the lessons with higher adherence, quality, and student engagement.

The difference between student outcomes in the Moderate and High IP classrooms was not statistically significant in any of the outcomes, suggesting that there may be a certain threshold of covered content and implementation practices that result in the greatest benefit. Identifying such "core components" (Lawson et al., 2019; Wigelsworth et al., 2021) of SSIS SEL CIP implementation is an important direction for future research. In general, classrooms where lessons were delivered with higher levels of adherence and quality, as well as engaged students in class activities more often, were associated with higher gains in prosocial behavior, academic motivation, and engagement. This finding is consistent with results from a previous efficacy trial of an earlier edition of the SSIS CIP program (e.g., DiPerna et al., 2015, 2016, 2018), in which the program, implemented with high levels of fidelity, yielded positive effects for students. To date, the existing research on universal SEL implementation in schools has largely come from efficacy trials, during which a research team provides implementation supports in order to enhance internal validity when evaluating and isolating the causal impact of programs. However, studies conducted under authentic or routine conditions, in which educators independently determine implementation based on their needs, resources, and capacities, are few and far between. Such effectiveness research is critical for extending the external validity of efficacy trials—that is, results from studies conducted with minimal researcher oversight may better represent and generalize to the real-world context of typical schools.

As implementation strategies facilitated by school psychologists and other support personnel have been shown to improve delivery of evidence-based practices in schools (Merle et al., 2022), identifying teacher implementation profiles may have practical implications for those supporting educators delivering universal SEL programs in real-world conditions. Collier-Meek et al. (2017) discussed the need for feasible and flexible implementation supports to promote integrity of universal program delivery across teachers with varying levels of need. They found brief daily emails containing tips and reminders were sufficient for improving observer-rated adherence and quality for some, but not all, teachers, suggesting that teachers may need to be matched with differentiated implementation support much like student needs are matched with interventions of varying intensity in multi-tiered service delivery systems. For the profiles that emerged in our sample, teachers in the Low IP group may require more individualized support regarding program implementation (e.g., emailed performance feedback, in-person coaching, and/or modeling) than those in the other two groups). Rather than support focused on program implementation, teachers in the High IP group may benefit from guidance in how SEL skill development can be prioritized, integrated, and generalized into their instructional time and interactions with students. Knowledge of differing implementation profiles can ensure that scarce resources like time are used strategically and optimally to support teachers with effective program delivery (Fallon et al., 2018), and that teachers receive targeted support that meets their needs.

5. Limitations and conclusion

There are several limitations to the current study. First, even though there appeared to be high degree of separation among the three identified profiles, the small sample size did not provide sufficient power to detect smaller group differences or examine the differential effects of contextual factors on student outcomes for profiles of classrooms. As such, findings of this study should be viewed as preliminary and interpreted with caution. Additional

future studies with larger and more diverse samples are necessary to verify the number and patterns of implementation profiles in real-world conditions. Second, we only assessed implementation in terms of dosage, adherence, quality, student engagement and teacher's impression of lessons. Based on anecdotal data, some teachers made modifications to the lessons due to a variety of factors (e.g., lack of time, perceived student needs); however, these adaptations are not accounted for in the study. While adaptations may decrease adherence, thoughtful and intentional changes informed by accurate student data to better meet student needs may actually improve student outcomes within the context of a universal program, and this is an important area for future research (Hunter et al., 2022; Neugebauer et al., 2023). Third, as noted previously some of the questions on the teachers' weekly survey may have (unintentionally) been ambiguous or susceptible to social desirability. It is also important to note that the current study was correlational in nature. Besides implementation and contextual factors at the class and school levels, there could be other factors contributing to student outcomes (e.g., student-level demographic and behavioral characteristics) that we did not measure or control. As such, no causal inference should be made without further investigation.

Nonetheless, examining typical practices of teachers when delivering universal SEL and associated contextual factors provides insight regarding the role of aspects of implementation in facilitating program outcomes. Specifically, we examined the role of multiple facets of implementation and how they potentially associated with student's gains (or lack thereof) from the SSIS SEL CIP when implemented under routine conditions in elementary classrooms. Results suggest that considering a single component of implementation (e.g., dosage) is potentially insufficient to account for the overall quality of implementation. Also, given the small variability of teachers' self-report scores, direct observation may provide a more accurate evaluation of implementation quality. Findings also suggest the need to evaluate implementation via multiple dimensions and measures. In addition, results suggest that the relationship between implementation factors and student outcomes may be more nuanced than prior studies featuring individual indicators of program implementation.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: The datasets presented in this article are not readily available because the grant period for the larger effectiveness trial is still ongoing. Requests to access the datasets should be directed to PI of the larger effectiveness trial (jdiperna@psu.edu).

Ethics statement

This study was reviewed and approved by the Institutional Review Board of the Pennsylvania State University. Written informed consent was provided by participating teachers and the parents/legal guardians of participating students. Verbal assent also was provided by participating students.

Author contributions

HZ contributed to the conceptual framework, conducted the statistical analysis, and wrote the draft of the manuscript. P-WL contributed to the conceptual framework and conducted the statistical analysis. SH contributed to the introduction and discussion sections. JD contributed to the conception and revised the manuscript. XL organized and prepared the data set. All authors contributed to the article and approved the submitted version.

Funding

This research was supported, in whole or in part, by the Institute of Education Sciences, US Department of Education, through Grant R305A170047 to The Pennsylvania State University.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Author disclaimer

The opinions expressed are those of the authors and do not represent the views of the Institute or the US Department of Education.

References

Anyon, Y., Nicotera, N., and Veeh, C. A. (2016). Contextual influences on the implementation of a schoolwide intervention to promote students' social, emotional, and academic learning. *Child. Sch.* 38, 81–88. doi: 10.1093/cs/cdw008

Asparouhov, T., and Muthén, B. (2014). Auxiliary variables in mixture modeling: three-step approaches using Mplus. *Struct. Equ. Model.* 21, 329–341. doi: 10.1080/10705511.2014.915181

Bakk, Z., and Vermunt, J. K. (2016). Robustness of stepwise latent class modeling with continuous distal outcomes. *Struct. Equ. Model.* 23, 20–31. doi: 10.1080/10705511.2014.955104

Beets, M. W., Flay, B. R., Vuchinich, S., Acock, A. C., Li, K. K., and Allred, C. (2008). School climate and teachers' beliefs and attitudes associated with implementation of the positive action program: a diffusion of innovations model. *Prev. Sci.* 9, 264–275. doi: 10.1007/s11121-008-0100-2

Bolck, A., Croon, M., and Hagenaars, J. (2004). Estimating latent structure models with categorical variables: one-step versus three-step estimators. *Polit. Anal.* 12, 3–27. doi: 10.1093/pan/mph001

Brackett, M. A., Reyes, M. R., Rivers, S. E., Elbertson, N. A., and Salovey, P. (2011). Assessing teachers' beliefs about social and emotional learning. *J. Psychoeduc. Assess.* 30, 219–236. doi: 10.1177/0734282911424879

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., and Balain, S. (2007). A conceptual framework for implementation fidelity. *Implement. Sci.* 2, 1–9. doi: 10.1186/1748-5908-2-40

Cheng, W., Lei, P. W., and Diperna, J. C. (2021). Using a perseon-centered analysis to examine second grade outcomes of a universal SEL program. *Sch. Psychol. Forum*. doi: 10.1037/spq0000539

Chhin, C. S., Taylor, K. A., and Wei, W. S. (2018). Supporting a culture of replication: an examination of education and special education research grants funded by the Institute of Education Sciences. *Educ. Res.* 47, 594–605. doi: 10.3102/0013189X18788047

Collier-Meek, M. A., Fallon, L. M., and DeFouw, E. R. (2017). Toward feasible implementation support: E-mailed prompts to promote teachers' treatment integrity. *Sch. Psychol. Rev.* 46, 379–394. doi: 10.17105/SPR-2017-0028.V46-4

Committee for Children (2005). *Steps to Respect: A Bullying Prevention Program*. Seattle, WA.

Committee for Children. (2012). *Second Step Early Learning*. Seattle, WA.

Cross, W., West, J., Wyman, P. A., Schmeelk-Cone, K., Xia, Y., Tu, X., et al. (2015). Observational measures of implementer fidelity for a school-based preventive intervention: development, reliability, and validity. *Prev. Sci.* 16, 122–132. doi: 10.1007/s11121-014-0488-9

Dane, A. V., and Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clin. Psychol. Rev.* 18, 23–45. doi: 10.1016/S0272-7358(97)00043-3

DiPerna, J. C., and Elliott, S. N. (2000). *Academic Competence Evaluation Scales*. San Antonio, TX:Psychological Corporation.

DiPerna, J. C., Lei, P., Bellinger, J., and Cheng, W. (2015). Efficacy of the social skills improvement system Classwide intervention program (SSIS-CIP) primary version. *Sch. Psychol. Q.* 30, 123–141. doi: 10.1037/spq0000079

DiPerna, J. C., Lei, P., Bellinger, J., and Cheng, W. (2016). Effects of a universal classroom behavior program on student learning. *Psychol. Sch.* 53, 189–203. doi: 10.1002/pits.21891

DiPerna, J. C., Lei, P., Cheng, W., Hart, S. C., and Bellinger, J. (2018). A cluster randomized trial of the social skills improvement system-Classwide intervention program (SSIS-CIP) in first grade. *J. Educ. Psychol.* 110, 1–16. doi: 10.1037/edu0000191

Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., et al. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: a conceptual framework celene. *Adv. School Ment. Health Promot.* 1, 6–28. doi: 10.1080/1754730X.2008.9715730

Domitrovich, C. E., Gest, S. D., Jones, D., Gill, S., and DeRousie, R. M. S. (2010). Implementation quality: lessons learned in the context of the head start REDI trial. *Early Child. Res. Q.* 25, 284–298. doi: 10.1016/j.ecresq.2010.04.001

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

Durlak, J. A. (2016). Programme implementation in social and emotional learning: basic issues and research findings. *Camb. J. Educ.* 46, 333–345. doi: 10.1080/0305764X.2016.1142504

Durlak, J. A., and DuPre, E. P. (2008). Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am. J. Community Psychol.* 41, 327–350. doi: 10.1007/s10464-008-9165-0

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

Elliott, S. N., and Gresham, F. M. (2017). *SSIS SEL Edition Classwide Intervention Program Manual*. Bloomington, MN: Pearson.

Fallon, L. M., Collier-Meek, M. A., Kurtz, K. D., and DeFouw, E. R. (2018). Emailed implementation supports to promote treatment integrity: comparing the effectiveness and acceptability of prompts and performance feedback. *J. Sch. Psychol.* 68, 113–128. doi: 10.1016/j.jsp.2018.03.001

Ferguson, S. L., Moore, E. W., and Hull, D. M. (2020). Finding latent groups in observed data: a primer on latent profile analysis in Mplus for applied researchers. *Int. J. Behav. Dev.* 44, 458–468. doi: 10.1177/0165025419881721

Ferguson, R., and Splaine, M. (2002). Use of the reliable change index to evaluate clinical significance in SF-36 outcomes. *Qual. Life Res.* 11, 509–516. doi: 10.1023/A

Flay, B. R., Allred, C. G., and Ordway, N. (2001). Effects of the positive action program on achievement and discipline: two matched-control comparisons. *Prev. Sci.* 2, 71–89. doi: 10.1023/A:1011591613728

Greenberg, M. T., Domitrovich, C. E., Weissberg, R. P., and Durlak, J. A. (2017). Social and emotional learning as a public health approach to education. *Futur. Child.* 27, 13–32. doi: 10.1353/foc.2017.0001

Greenberg, M. T., and Kusche, C. A. (1993). *Promoting Social and Emotional Skills in Deaf Children: The PATHS Project*. Seattle, WA, US: University of Washington Press.

Greenberg, M. T., Kusche, C. A., Cook, E. T., and Quamma, J. P. (1995). Promoting emotional competence in school-aged children: the effects of the PATHS curriculum. *Dev. Psychopathol.* 7, 117–136. doi: 10.1017/S0954579400006374

Gresham, F., and Elliott, S. N. (2008). *Social Skills Improvement System Rating Scales*. Minneapolis, MN: Pearson Assessments.

Hart, S. (2021). Beliefs about social-emotional skills: Development of the assumptions supporting social-emotional teaching (ASSET) scale [Unpublished doctoral dissertation]. The Pennsylvania State University.

Hennessey, A., and Humphrey, N. (2020). Can social and emotional learning improve children's academic progress? Findings from a randomised controlled trial of the promoting alternative thinking strategies (PATHS) curriculum. *Eur. J. Psychol. Educ.* 35, 751–774. doi: 10.1007/s10212-019-00452-6

Holden, R. R., and Passey, J. (2009). "Social desirability" in *Handbook of individual differences in social behavior*. eds. M. R. Leary and R. H. Hoyle (New York: The Guilford Press), 441–454.

Humphrey, N., Barlow, A., and Lendrum, A. (2018). Quality matters: implementation moderates student outcomes in the PATHS curriculum. *Prev. Sci.* 19, 197–208. doi: 10.1007/s11121-017-0802-4

Hunter, L. J., DiPerna, J. C., Hart, S. C., Neugebauer, S., and Lei, P. (2022). Examining teacher approaches to implementation of a classwide SEL program. School. *Psychology* 37, 285–297. Advance online publication. doi: 10.1037/spq0000502

Jacobson, N. S., and Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J. Consult. Clin. Psychol.* 59, 12–19. doi: 10.1037/0022-006X.59.1.12

Jones, S. M., Farrington, C. A., Jagers, R., Brackett, M., and Kahn, J. (2019). National Commission on social, emotional, and academic development: a research agenda for the next generation. *The Aspen Institute*. Available online at: https://www.aspeninstitute.org/publications/a-research-agenda-for-the-next-generation/ (accessed March 13, 2023).

Laursen, B., and Hoff, E. (2006). Person-centered and variable-centered approaches to longitudinal data. *Merrill-Palmer Q.* 52, 377–389. doi: 10.1353/mpq.2006.0029

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

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

Low, S., Van Ryzin, M. J., Brown, E. C., Smith, B. H., and Haggerty, K. P. (2014). Engagement matters: lessons from assessing classroom implementation of steps to respect: a bullying prevention program over a one-year period. *Prev. Sci.* 15, 165–176. doi: 10.1007/s11121-012-0359-1

Magidson, J., and Vermunt, J. K. (2005). *Technical Guide for Latent GOLD 4.0: Basic and Advanced*. Belmont, MA: Statistical Innovations Inc.

Merle, J. L., Thayer, A. J., Larson, M. F., Pauling, S., Cook, C. R., Rios, J. A., et al. (2022). Investigating strategies to increase general education teachers' adherence to evidence-based social-emotional behavior practices: a meta-analysis of the single-case literature. *J. Sch. Psychol.* 91, 1–26. doi: 10.1016/j.jsp.2021.11.005

Muthén, B., and Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. *Alcohol. Clin. Exp. Res.* 24, 882–891. doi: 10.1111/j.1530-0277.2000.tb02070.x

Neugebauer, S. R., Sandilos, L. E., DiPerna, J. C., Hunter, L. J., and Hart, S. C. (2023). 41 teachers, 41 different ways: exploring teacher implementation of a universal social-emotional learning program under routine conditions. *Elem. Sch. J.*

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Rev. Educ. Res.* 78, 33–84. doi: 10.3102/0034654307313793

Pianta, R. C., La Paro, K. M., and Hamre, B. K. (2008). *Classroom Assessment Scoring System™: Manual K-3* Baltmore, MD: Paul H Brookes Publishing.

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., et al. (2011). Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm. Policy Ment. Health Ment. Health Serv. Res.* 38, 65–76. doi: 10.1007/s10488-010-0319-7

Raferty, A. E. (1995). Bayesian model selection in social research. *Sociol. Methodol.* 25, 111–163. doi: 10.2307/271063

Reyes, M. R., Brackett, M. A., Rivers, S. E., Elbertson, N. A., and Salovey, P. (2012). The interaction effects of program training, dosage, and implementation quality on targeted student outcomes for the RULER approach to social and emotional learning. *Sch. Psychol. Rev.* 41, 82–99. doi: 10.1080/02796015.2012.12087377

Rojas-Andrade, R., and Bahamondes, L. L. (2019). Is implementation fidelity important? A systematic review on school-based mental health programs. Contemporary school. *Psychology* 23, 339–350. doi: 10.1007/s40688-018-0175-0

Sanetti, L. M., and Collier-Meek, M. A. (2019). Increasing implementation science literacy to address the research-to-practice gap in school psychology. *J. Sch. Psychol.* 76, 33–47. doi: 10.1016/j.jsp.2019.07.008

Smith, P. E., Witherspoon, D. P., and Lei, P. W. (2021). The "haves, have some, and have Nots": a latent profile analysis of capacity, quality, and implementation in community-based afterschool programs. *Prev. Sci.* 22, 971–985. doi: 10.1007/s11121-021-01258-z

Spurk, D., Hirschi, A., Wang, M., Valero, D., and Kauffeld, S. (2020). Latent profile analysis: a review and "how to" guide of its application within vocational behavior research. *J. Vocat. Behav.* 120:103445. doi: 10.1016/j.jvb.2020.103445

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

Tein, J. Y., Coxe, S., and Cham, H. (2013). Statistical power to detect the correct number of classes in latent profile analysis. *Struct. Equ. Model.* 20, 640–657. doi: 10.1080/10705511.2013.824781

Wigelsworth, M., Verity, L., Mason, C., Qualter, P., and Humphrey, N. (2021). Social and emotional learning in primary schools: a review of the current state of evidence. *Br. J. Educ. Psychol.* 92, 898–924. doi: 10.1111/bjep.12480

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 27 April 2023
DOI 10.3389/fpsyg.2023.1065749

OPEN ACCESS

EDITED BY
Celene Domitrovich,
Georgetown University, United States

REVIEWED BY
Karen Bierman,
The Pennsylvania State University (PSU), United States
Spring Dawson-McClure,
Langone Medical Center, New York University, United States

*CORRESPONDENCE
John A. Gómez
✉ jgomezvaron@fordham.edu

SPECIALTY SECTION
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

RECEIVED 10 October 2022
ACCEPTED 30 March 2023
PUBLISHED 27 April 2023

CITATION
Gómez JA, Brown JL and Downer JT (2023) High quality implementation of 4Rs+MTP increases classroom emotional support and reduces absenteeism.
Front. Psychol. 14:1065749.
doi: 10.3389/fpsyg.2023.1065749

High quality implementation of 4Rs+MTP increases classroom emotional support and reduces absenteeism

John A. Gómez[1]*, Joshua L. Brown[1] and Jason T. Downer[2]

[1]Applied Developmental Psychology, Department of Psychology, Fordham University, Bronx, NY, United States, [2]Center for Advanced Study of Teaching and Learning, School of Education and Human Development, University of Virginia, Charlottesville, VA, United States

School-based social and emotional learning (SEL) programs are associated with improvements in children's SEL and academic outcomes, and the quality of classroom interactions. The magnitude of these effects increases at high levels of program implementation quality. This study aimed to (1) identify teachers' profiles of quality of implementation, (2) explore teachers and classroom characteristics contributing to their propensity to comply with high quality of implementation, and (3) examine the relations between school assignment to an SEL program, quality of classroom interactions, and child SEL and academic outcomes at different levels of teachers' compliance propensity. This study drew upon data from a cluster-randomized controlled trial evaluating the efficacy of 4Rs+MTP, a literacy-based SEL program, on third and fourth grade teachers (n=330) and their students (n=5,081) across 60 New York City public elementary schools. Latent profile analysis indicated that measures of teacher responsiveness and amount of exposure to implementation supports contributed to the differentiation of profiles of high and low quality of implementation. Random forest analysis showed that more experienced teachers with low levels of professional burnout had high propensity to comply with high quality of implementation. Multilevel moderated mediation analysis indicated that 4Rs+MTP teachers with high compliance propensity were associated with higher classroom emotional support and lower children's school absences than their counterparts in the control group. These findings may inform debates in policy research about the importance of providing the supports teachers need to implement SEL school programs with high quality.

KEYWORDS

quality of implementation, school program, social and emotional learning, classroom interactions, compliance propensity

1. Introduction

1.1. School-based social and emotional learning interventions and child development

School-based social and emotional learning (SEL) interventions encompass a series of intentional program efforts designed to promote children's learning and application of social, emotional and character skills required to succeed in school, workplace settings, relationships, and citizenship (Jones et al., 2019; Weissberg, 2019). Efficacy studies of such programs show

that, relative to children in control conditions, children participating in these programs show improvements in their social and emotional skills, positive attitudes toward self and others, positive social behavior, fewer conduct problems, lower emotional distress, and have higher academic achievement scores (Durlak et al., 2011; Grant et al., 2017; Taylor et al., 2017).

SEL interventions in schools typically have the objective of both teaching students specific social and emotional skills and creating caring and supportive classroom interactions where such skills flourish (Brown et al., 2010; Brackett et al., 2012). According to the Teaching through Interactions Framework (Hamre and Pianta, 2007; Hamre et al., 2013), quality of classroom interactions is organized into three domains: Emotional Support, Classroom Organization, and Instructional Support. High quality interactions within each of these domains are hypothesized to promote students' learning and social development (Hamre and Pianta, 2007; Pianta and Hamre, 2009). A randomized controlled trial of one school-based SEL program, Reading, Writing, Respect and Resolution (4Rs), showed positive effects on quality of third-grade classroom interactions as measured by independent observers (Brown et al., 2010). Moreover, children in schools implementing 4Rs showed lower hostile attribution biases and fewer depressive symptoms at the end of the first year compared to children in control schools (Jones et al., 2010), benefits that persisted and expanded to other outcome domains including teacher-reported attention skills, and aggressive and socially competent behaviors, following a second year of program implementation (Jones et al., 2011). These findings suggest that SEL programs improve teachers' support during classroom interactions, which may in turn lead to other benefits for children experiencing these higher quality classroom interactions. A developmental systems approach to the evaluation of school-based SEL programs affords comprehensive interpretations of changes in children's social and emotional functioning within contexts that provide rich and nurturing interactions (Roeser et al., 2000; Hamre and Pianta, 2005). However, no research to date has tested the quality of classroom interactions as a mediator of the effect of school-based SEL programs on children's developmental outcomes.

1.2. Quality of implementation

Research in program evaluation suggests that the positive effects of school-based interventions on children's outcomes depends largely on the quality of program implementation (Dane and Schneider, 1998; Fixsen et al., 2009; Lendrum and Humphrey, 2012; Durlak, 2016). For instance, in a meta-analytic study, Durlak et al. (2011) found that children in better-implemented SEL programs compared to poorly implemented programs showed greater gains in academic achievement and larger reductions in conduct problems and emotional distress. Similarly, in an elementary school mental health program, Dix et al. (2012) found that the difference between children in high- and low-implementation schools represented a difference in academic performance favoring children in high implementation schools equivalent to 6 months of schooling.

One of the limitations of most implementation studies is an over-emphasis on fidelity of implementation of program activities, that is, the extent to which the program was implemented as planned (Dane and Schneider, 1998). In this regard, measures such as the amount of program activities participants implemented (dosage) and participant enactment of program protocols (adherence) are considered to determine how well the program was implemented (Fixsen et al., 2009). Less attention has been paid to other factors associated with participants' responsiveness to program implementation and quality of delivery of program activities (Durlak and DuPre, 2008; Lendrum et al., 2016). In addition, implementation studies have overlook the importance of factors associated with the supports teachers need for effective program delivery (Domitrovich et al., 2010), such as on-going coaching, training, and modeling from program experts (Hadden and Pianta, 2006; Kraft et al., 2018). Over the past two decades, the operationalization of fidelity of implementation has been extended to incorporate the construct of quality of implementation, including various measures of the amount and quality of both program implementation and implementation supports (Domitrovich et al., 2010; Dix et al., 2012). Figure 1A illustrates the model we use to distinguish program implementation from implementation supports in this study.

Previous research on school-based SEL interventions has found associations between several measures of program implementation and implementation supports and child and teacher outcomes. For instance, high quality of program delivery by teachers and children's responsiveness to the intervention has been associated with reductions in children's aggressive behavior and improvements in academic outcomes (Domitrovich et al., 2010; Humphrey et al., 2016). High program dosage and adherence have also been associated with improvements in children's social competence (Rosenblatt and Elias, 2008). Similarly, teachers' adherence to SEL program practices has been linked to gains in teachers' quality of interactions with children in their classrooms (Abry et al., 2013). Teachers' access to on-going coaching has also been linked to improvements in teacher emotional and instructional support in their classroom interactions with children (Pianta et al., 2008).

Program implementation and implementation supports are also associated. For instance, teachers attending high numbers of coaching sessions are also more likely to implement a higher number of program activities than teachers with few coaching sessions (Pas et al., 2015), and the strength of the teacher-coach working alliance has been associated with higher teacher adherence to program protocols (Wehby et al., 2012). Given this dynamic interaction between program implementation and implementation supports, it is pertinent to examine these two factors simultaneously in models evaluating the effect of quality of implementation as a whole on targeted program outcomes (Durlak and DuPre, 2008; Domitrovich et al., 2010). However, including several interrelated measures of implementation and implementation supports in a single model increases the risk of multicollinearity and the probability of Type I error. In this scenario, a comprehensive methodological approach that integrates measures of program implementation and implementation supports into a global index of quality of implementation may be a reasonable alternative to investigate the effects of implementation quality on program outcomes (Dix et al., 2010).

One example of a comprehensive measure of implementation is provided by Dix et al. (2012) in the study of the effects of a two-year SEL program (KidsMatter) on the academic outcomes of children from 100 elementary schools. Guided by the framework proposed by Domitrovich et al. (2008), an overall index of school implementation was created to sort schools into high and low implementation groups. The index was created using Latent Class Analysis (LCA) with a total

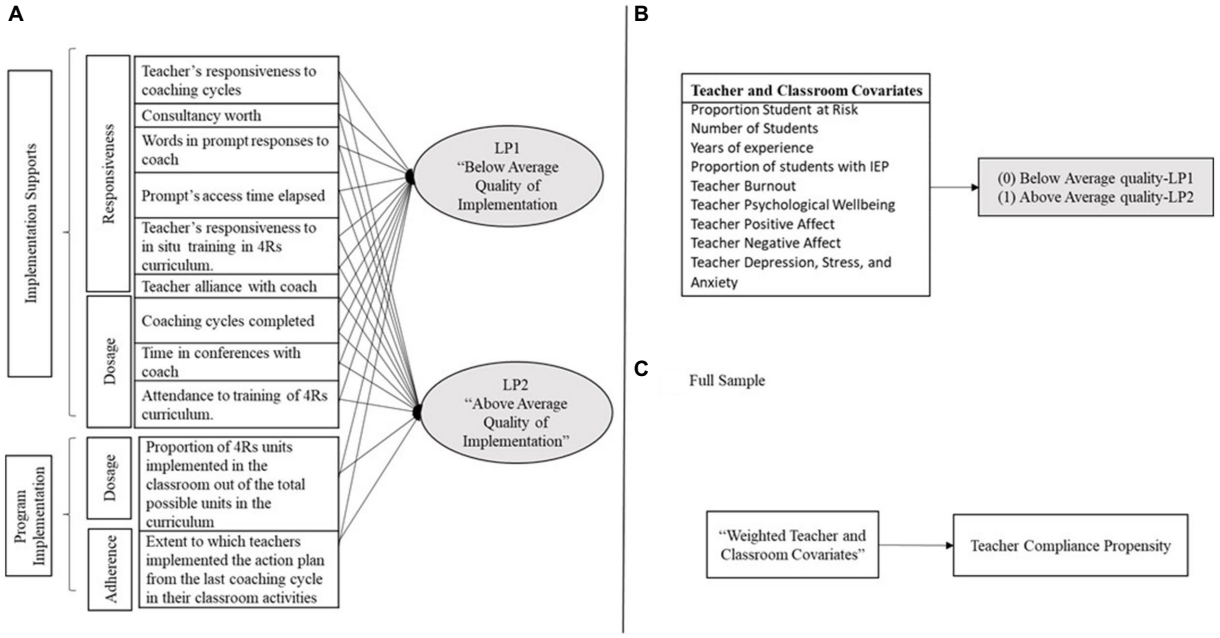


FIGURE 1
(A–C) Models guiding the estimation of teacher compliance propensity.

of 37 items measuring adherence, dosage, and quality of delivery of both program implementation and implementation supports. However, only 13 items were found to differentiate adequately between high and low implementing schools; these particular items corresponded to adherence and quality of delivery of implementation support, as well as dosage of program implementation (Slee et al., 2009). This index successfully predicted better academic outcomes (in Literature and Mathematics) for children in high, compared to low, implementing schools. The difference between children from low and high implementing schools was robust to school level SES and was equivalent to a difference in academic performance of up to 6 months of schooling (Dix et al., 2012).

Research testing the effects of quality of implementation for core program components on program outcomes, or using multidimensional measures of implementation (e.g., dosage, adherence, responsiveness, quality of delivery), often restrict their study to the treated sample. Thus, the generalization of results is restricted to individuals sharing characteristics with one subsample of the study. Information about the quality of implementation is supposed to yield better insights into program related causes of the treatment effects; however, access to limited information about implementation and only in the treated sample is not sufficient for making better causal inferences about the treatment effects in the population.

1.3. Implementation and program effects

Recently, some studies are using a method to estimate the causal effect of complying with implementation of SEL interventions on child and teacher outcomes (Berg et al., 2017; Panayiotou et al., 2019), called Compliers Average Causal Effect (CACE). CACE is a method developed originally by Imbens and Rubin (1997) to identify differences in medical program outcomes between those who received the treatment and those who did not in both treatment and control groups. This approach requires clear cut-offs regarding participants who will be considered compliers, in order to make effective comparisons between participants in the treatment condition who did not comply with treatment and participants in the control condition who did not receive treatment (Follmann, 2000). In studies of SEL programs using CACE, compliers are often defined with a single measure of dosage (e.g., those who implemented one standard deviation above the median number of program activities; Panayiotou et al., 2019). In education research, this approach is limited. In addition to the term "complier" applied to an active and autonomous agent such as teachers implementing a program in their own field, it assumes those teachers who implemented less than the cut-off score (e.g., 1/3 of the curricular activities) and their children who received less program dosage are comparable with those in the control group who never implemented/received the intervention (Sagarin et al., 2014). However, research in SEL implementation has found that even when program dosage is low, the quality of delivery may make a difference in terms of program outcomes (Durlak and DuPre, 2008; Humphrey et al., 2018). Therefore, using a single measure of dosage to estimate CACE is similar to an arbitrary decision.

Other methods have been proposed that do not require equating noncompliers in a treatment group with participants in a control group (Sagarin et al., 2014). A propensity score approach has been proposed by Follmann (2000) to estimate the probability of compliance using covariates. This method allows for the estimation of, for example, teacher compliance propensity using teacher and classroom characteristics (covariates) known to predict high quality implementation among teachers (Downer et al., 2009b; Domitrovich et al., 2019). For instance, teachers in classrooms with high proportions of at-risk students and high emotional exhaustion (burnout) have shown lower implementation of SEL activities in the classroom (Musci

et al., 2019). Therefore, using baseline covariates as predictors of compliance in the treatment group, it is possible to predict compliance propensity in the control group with similar baseline covariates. Translating this approach to a school-based SEL intervention context, compliance propensity could be estimated using baseline covariates as predictors of high and low quality of implementation in the treatment group and, using similar covariates, also predict compliance propensity in the control group. Since the propensity approach does not require equating noncompliers in the treatment group with select participants in the control group, teacher compliance propensity can be estimated using several indicators of high quality of implementation, in addition to dosage.

The magnitude of treatment effects at different levels of teacher compliance propensity may provide valuable information about how the high quality of implementation in SEL program may influence program effects on teachers and children. Particularly, this high quality of implementation may increase the effects of SEL programs in improving classroom interactions where children thrive with opportunities to learn and develop social and emotional skills needed to succeed in life. However, there is no research to date using compliance propensity in evaluations of SEL program implementation.

1.4. The current study

This study aims to understand the effects of an SEL program on classroom interactions and child SEL and academic outcomes as moderated by teachers' compliance propensity. Furthermore, this study examines the role of classroom quality of interactions as a mediator of the relationship between an SEL program and child outcomes at different levels of teachers' compliance propensity (see Figure 2). This study drew upon data collected as part of a cluster-randomized controlled trial evaluating the efficacy of a literacy-based SEL program (4Rs + MTP) implemented and tested in two consecutive cohorts of students within 60 New York City (NYC) public elementary schools during the 2015–2016 and 2016–2017 school years. The 4Rs + MTP program integrates two distinct and complementary evidence-based interventions: Reading Writing, Respect and Resolution (4Rs), a universal, school-based program integrating social and emotional competencies into the language arts curriculum for grades K-5 (Jones et al., 2011), and MyTeachingPartner (MTP), a coaching model that is based on providing teachers with personalized feedback and on-demand support of curriculum implementation, through web-based and *in situ* teacher-coach interaction (Pianta et al., 2008). The 4Rs program uses an ecological developmental approach (Bronfenbrenner and Morris, 1998) which posits that children develop negotiation strategies in interpersonal interactions within specific contexts. Accordingly, the 4Rs program includes social-cognitive processes associated with aggressive behaviors (e.g., hostile attribution bias), and classroom quality of interactions as proximal outcomes in their theory of change (Aber et al., 1998, 2011). Activities in the 4Rs program involve the selection of high-quality children's literature that invites children, with the guidance of their teachers, to learn how to handle anger and use skills like listening, cooperation, assertiveness, and negotiation during interpersonal conflicts in classroom (Aber et al., 2011). The MyTeachingPartner (MTP) coaching approach draws on attachment theory as instantiated within classroom interactions (Hamre et al., 2013), positing that the quality of teacher-student interactions is pivotal for student learning, with a particular focus on the support teachers provide to create caring and trusting relationships with their students. The integration of the 4Rs program and the MTP coaching approach extends the focus of 4Rs in promoting positive interpersonal relationships in classrooms, including activities to promote high quality child-teacher relationships along with the activities to promote positive interpersonal negotiation strategies with peers. The MTP coaching approach also includes ongoing coaching to support teachers' implementation of the 4Rs program curriculum. In sum, the 4Rs + MTP program aims to provide a systematic and comprehensive approach to in-person training on quality of classroom interactions and 8 cycles of video-based and web-mediated coaching focused on teachers' implementation of the 4Rs curricular units in the classroom (e.g., lessons, book talk) that target the development of social and emotional competencies within the language arts. The program is designed to promote teachers' psychological well-being, and high-quality classroom interactions that foster student learning, and the development of students' social and emotional competencies and academic functioning. Preliminary findings based on the intent-to-treat analyses from the efficacy trial of the 4Rs + MTP program demonstrate positive intervention effects after 1 year of 4Rs + MTP implementation on teacher anxiety and stress, observation-based ratings of emotionally supportive interactions in classrooms, and children's social competence, aggressive behavior and conduct problems as reported by teachers controlling for baseline scores on each outcome (Brown et al., 2019).

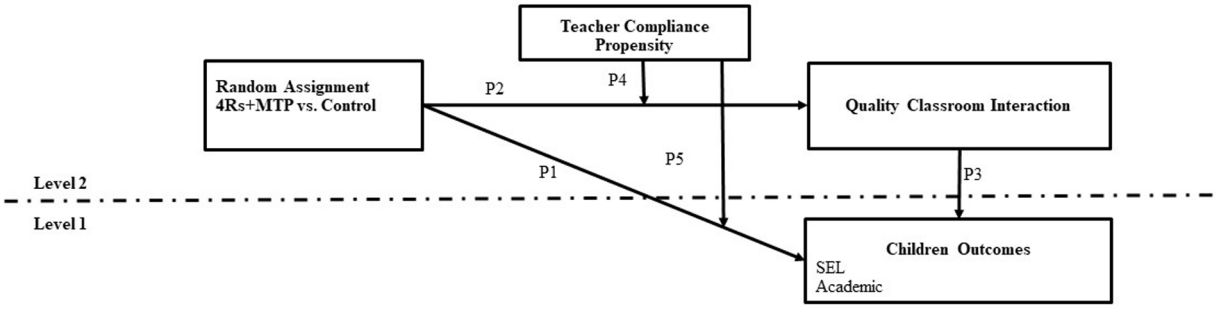


FIGURE 2
Multilevel moderated mediation model. Level 1 refers to variables at the child level, and level 2 refers to variables at the teacher and classroom level. P1–P5 refer to paths representing the relationships between variables.

The specific aims of this study were twofold. The first aim was to estimate teacher compliance propensity. Since there is no research to date estimating propensities of quality of implementation, the research questions for this objective are exploratory in nature, although partially informed by previous research on the quality of program implementation. Here we explore which components of program implementation and implementation supports would discriminate between profiles of quality of implementation for teachers in the treatment group (see Figure 1A). Previous work has found that both measures of implementation and implementation support successfully allow the classification of schools into high and low quality of implementation classes (Dix et al., 2010). We then explore teacher and classroom covariates that predict teacher compliance propensity for teachers in the treatment group (see Figure 1B) and the full sample (treatment and control groups; see Figure 1C). Previous research has found that teacher age, job burnout, and baseline quality of classroom interactions, and classroom characteristics such as percentage of behaviorally at risk children, are associated with quality of implementation (Downer et al., 2009b; Berg et al., 2017; Domitrovich et al., 2019). The second aim of this study was to examine teacher compliance propensity as a moderator of the relationship between school random assignment to 4Rs + MTP and (a) quality of classroom interaction (see Figure 2, paths P2 and P4), and (b) children's SEL and academic outcomes (see Figure 2, paths P1 and P5). And (c) children's SEL and academic outcomes as mediated by the quality of classroom interaction (see Figure 2, paths P2, P4, and P3)?

It was expected that teachers randomly assigned to 4Rs + MTP would have higher quality classroom interactions than teachers in the control group, when examined at an above average level of compliance propensity. Previous findings show that teachers who receive consultation and on-going support from coaches, compared with no personalized support, were able to provide better support to their students in the classroom (Pianta et al., 2008; Early et al., 2017). It was also expected that teachers with above average compliance propensity in schools randomly assigned to 4Rs + MTP would have children with higher social–emotional and academic outcomes than children of teachers in the control group. The literature on quality of implementation indicates that, in contrast to poorly implemented programs, well-implemented programs are associated with increases in children's prosocial behavior (Rimm-Kaufman et al., 2007) and academic achievement (Dix et al., 2012), and reductions in conduct problems and emotional distress (Durlak and DuPre, 2008; Durlak et al., 2011) as well as unexcused absences from school (Neace and Munoz, 2012). Lastly, it was expected that teachers with above average compliance propensity in schools randomly assigned to 4Rs + MTP would have children with higher social–emotional and academic outcomes as mediated by higher quality of classroom interactions than children of teachers in schools randomly assigned to the control group. Previous research has found that supportive and nurturing interactions in the classroom are associated with better child behavioral (Portnow et al., 2018; Rucinski et al., 2018) and cognitive outcomes (Curby et al., 2009; Reyes et al., 2012). These findings along with evidence regarding positive effects of SEL interventions and quality of implementation on classroom quality and child outcomes suggest that children may benefit from SEL interventions through the effects on the quality of classroom interactions when interventions are well implemented.

2. Materials and methods

2.1. Participants

Data for this study were collected across two cohorts (2015–2016; 2016–2017) as part of the randomized controlled trial (RCT) of the 4Rs + MTP program. Across both cohorts, the study sample is comprised of 5,081 third- and fourth-grade children (treatment $n = 2{,}326$; control $n = 2{,}755$) taught by 334 teachers (treatment $n = 151$; control $n = 183$) from 60 urban, high needs elementary schools (treatment $n = 31$, control $n = 29$). Teachers in the treatment group were each assigned one of seven dedicated 4Rs + MTP coaches. There was a similar proportion of teachers in third (45.5%) and fourth (44%) grades, and most were female (90.9%). The ethnic/racial composition of teachers were White (38.9%), Hispanic/Latina (27.8%) and Black or African American (21.9%). On average, each classroom had approximately 22 children, and on average, 15 children per classroom ($SD = 5.12$) participated in the study. Table 1 shows teacher and classroom demographic characteristics.

Children were 51.6% female with an average age of 8.8 years old (range: 5–12 years old), 65% Latine, 22% Black, 6% White, 5% Asian, and 2% other. Table 2 presents demographic characteristics of children in the sample.

2.2. Procedures

This randomized controlled study of the 4Rs + MTP program was carried out in two phases for each of two consecutive cohorts of participating schools. The first phase included school recruitment, school random assignment to treatment and control conditions, and consenting of teachers and children. The second phase included data collection and program implementation which took place during one school year. A timeline of activities during these two phases of the study is provided in Table 3.

Before the school random assignment process, third and fourth-grade teachers from each participating school were sent *via* email consent forms explaining the purpose of 4Rs + MTP and training and implementation procedures for the treatment group, and emphasizing that participation in the study was voluntary. Across cohorts there were a total of 444 eligible 3rd and 4th grade teachers out of which 336 consented to participate (76% overall, 76% in treatment group and 75% in control group). Trained research team members visited classrooms of all participating teachers and provided students with a brief, age-appropriate explanation of the study and the procedures for data collection. Students received consent forms in English and Spanish to take home to their caregivers/guardians. Students who returned a consent form signed by their parent/guardian indicating either consent or denial, were given a new grade-appropriate children's book. From the 7,708 eligible students across cohorts, 5,081 (66%) were consented (treatment $N = 2{,}326$; control $N = 2{,}755$). All study procedures were approved by the local Department of Education.

Participating teachers and children completed baseline measures during Fall/Winter (wave 1) and end of year measures during Spring (wave 2). At each wave, teachers completed a battery of assessment for each of the consented children in their classrooms, as well as their own self-assessments, including demographic information about themselves and their classrooms. Children completed self-assessments

TABLE 1 Teacher and classroom demographic characteristics (n=334).

		Treatment		Control		Total	
		n	%	*n*	%	*n*	%
Gender	Female	136	90.1	164	90	300	90.9
	Male	14	9.2	16	9	30	9.1
	Missing	1	0.7	3	2	4	1.2
Race	White	46	30.5	73	40	119	38.9
	Hispanic or Latina	40	26.5	45	25	85	27.8
	Black-African American	41	27.1	26	14	67	21.9
	Multi racial	7	4.6	15	8	22	7.2
	Asian	4	2.6	6	3	10	3.3
	Other	2	1.3	0	0	2	0.6
	Hispanic and black	1	0.6	0	0	1	0.3
	Missing	10	6.6	18	10	28	8.38
Teaching certificate	Regular or standard state	123	81	133	73	256	78.8
	Provisional or other type	3	1.9	11	6	14	4.3
	Probationary certificate	19	12.6	24	13	43	13.2
	Temporary certificate	3	1.9	2	1	5	1.5
	Other certificate	1	0.6	6	3	7	2.1
	Missing	2	1.3	7	4	9	2.69
Highest degree	Bachelor's degree	8	5.0	10	5	18	5.5
	Master's degree	138	91	164	90	302	92.3
	Specialists degree	2	1.0	2	1	4	1.2
	Doctorate degree	2	1.0	0	0	2	0.6
	Other degree	0	0.0	1	1	1	0.3
	Missing	1	1.0	6	3	7	2.1
Random assignment		151	100.0	183	100	334	100.0
Grade	3rd grade	67	44.0	80	44	147	45.5
	4th grade	64	42.0	78	43	142	44.0
	Mixed	17	11.0	17	9	34	10.5
	Missing	3	2.0	8	4	11	3.3
Classroom type	General education	90	60	102	56	192	59.8
	ICT/CTT/inclusion	37	25	46	25	83	25.9
	Self-contained special Ed	21	14	25	14	46	14.3
	Missing	3	2	10	5	13	3.9
Language program	None/English	132	87	147	80	279	86.4
	Dual language	5	3	10	5	15	4.6
	Bilingual	9	6	4	2	13	4.0
	ENL/ESL	2	1	14	8	16	5.0
	Missing	3	2	8	4	11	3.3
	Tx		**Ct**		**Total**		
	M	***SD***	***M***	***SD***	***M***	***SD***	**Min**	**Max**
Years of teaching experience	11.99	8.60	9.54	6.40	10.70	7.60	1	40
Years at current school	8.03	7.23	6.39	5.61	7.10	6.40	1	31
Class size	21.55	5.77	23.11	5.97	22.40	5.90	6	33
Proportion of girls	0.47	0.14	0.47	0.12	0.48	0.13	0	1
Proportion of IEP	0.13	0.23	0.18	0.25	0.16	0.23	0	1
Proportion of LEP	0.29	0.33	0.29	0.33	0.29	0.33	0	29

ICT, Integrated Co-Teaching; CTT, Collaborative Team Teaching; IEP, Individualized Education Plan; LEP, Limited English proficiency.

TABLE 2 Child demographic characteristics.

		Treatment		Control		Total	
		n	%	*n*	%	*n*	%
School random assignment		2,326	46	2,755	54	5,081	100.0
Missing							
Gender	Female	1,216	52.27	1,404	50.96	2,620	51.56
	Male	1,089	46.83	1,336	48.50	2,425	47.73
Race	Latine	1,501	64.53	1,816	65.92	3,317	65.28
	Black	613	26.35	514	18.66	1,127	22.18
	White	113	4.86	197	7.15	310	6.10
	Asian	50	2.15	182	6.61	232	4.57
	Multiracial	18	0.77	19	0.69	37	0.73
	Native American	10	0.43	12	0.44	22	0.43
Missing		21	0.90	15	0.54	36	0.71
		M	*SD*	*M*	*SD*	*M*	*SD*
Age		8.78	0.79	8.81	0.80	8.80	0.8
		Min	Max	Min	Max	Min	Max
		6	12	5	12	5	12

via classroom administration of paper and pencil surveys, and trained observers visited each classroom to observe and rate the quality of classroom interactions and teaching practices.

2.2.1. Implementation of the 4Rs+MTP program

The 4Rs + MTP program implementation includes professional development for teachers in support of their effective delivery of the 4Rs curriculum in their classrooms. Seven experienced former educators served as 4Rs + MTP coaches and provided six in-person training sessions followed by eight cycles of one-on-one coaching to teachers in 4Rs + MTP schools, emphasizing content and strategies aimed to improve the quality of interactions in the classroom *via* the Teaching through Interactions framework (Hamre et al., 2013). Below, procedures for each component of 4Rs + MTP program are described.

2.2.1.1. The 4Rs program

The 4Rs component of the program consists of a curriculum divided into seven units, each focused on promoting skills to understand and handle feelings, listening to others, establish nurturing relationships through cooperation, negotiation, and building community. In total, 4Rs offers 66 in-class activities for third grade and 70 for fourth-grade classrooms. 4Rs includes a professional development component for teachers, consisting of six in person training sessions, each 6 h long, aimed to equip teachers with the knowledge and techniques needed to implement the 4Rs curriculum effectively in classrooms. Teachers also received *in situ* support from coaches to model, co-teach, and provide feedback on teacher implementation of 4Rs curricular activities.

2.2.1.2. MyTeachingPartner coaching

The MyTeachingPartner (MTP) component of the program provided teachers with support and professional development focused on one-to-one video-based coaching and access to video exemplars through a web-based interactive platform. 4Rs + MTP included eight video-based teacher-coach cycles, each focused on coaches (a) observing video recordings from each of their teachers' classrooms to identify effective teacher-child interactions during 4Rs + MTP implementation; (b) providing each of their teachers with written prompts to focus teachers' attention and generate teacher reflection on their interactions with students during the program implementation; and (c) establishing a supportive and reliable coach-teacher alliance where teachers felt comfortable asking questions and reflecting on their challenges without feeling judged or evaluated (Hadden and Pianta, 2006). Each cycle lasted approximately 4 to 5 weeks and was repeated eight times during the year of 4Rs + MTP implementation.

Data were collected on teachers' implementation of program activities in the classroom and the implementation supports they received through training and coaching. Teachers in the treatment group used the online MTP web platform to keep weekly logs of their implementation of 4Rs + MTP lessons and units. Additionally, the MTP web platform registered information about teachers' logins and visits to the platform's webpages and the duration of these visits.

2.3. Measures

2.3.1. Classroom interactions and children's SEL and academic outcomes

2.3.1.1. Student academic competence and school attendance

Student academic competence was measured using the New York State standardized English Language Arts (ELA) and Math achievement tests. The ELA population-scale score mean was 599.79 ($SD = 20.22$), and 599.77 ($SD = 20.17$) for third and fourth grades, respectively. Math population-scale score means for third and fourth grade in 2018 was 599.48 ($SD = 20.19$), and 599.38 ($SD = 20.23$).

Children's school attendance was measured using direct reports of class attendance from the local Department of Education (DOE) records at the end of the year prior to the start of the study and again at the end of the main year of the study in each cohort. The DOE attendance data provides information on children's number of days absent and number of days present during the school year. Baseline measures of attendance correspond to the total days present in the year prior to intervention delivery which was 2014–2015 for cohort 1 and 2015–2016 for cohort 2, whereas the end of the year measures of attendance corresponds to data from 2015 to 2016 for cohort 1 and 2016–2017 for cohort 2, the end of the main year of the study in each cohort.

2.3.1.2. Anxious and depressive symptoms (child report)

Children's self-reported depressive and anxiety symptoms were measured through the depression subscale and anxiety subscale of the Behavioral Assessment System for Children (BASC; Kamphaus and Reynolds, 1998). The depression subscale is comprised of 13 true/false statements, such as "Nothing ever goes right for me." Children's anxious symptoms were measured using the self-report 13-item anxiety subscale of BASC. An example item includes, "I get so nervous I cannot breathe." Scale reliabilities ranged from $\alpha = 0.85$ in fall/winter to $\alpha = 0.85$ in spring.

TABLE 3 Study timeline.

School year	2014–2015		2015–2016				2016–2017			
	SP	SU	F	W	SP	SU	F	W	SP	SU
Cohort 1										
School recruitment	×									
Teacher consent		×								
Random assignment		×								
Student consent			×							
4Rs + MTP implementation and professional development			×	×	×					
Wave 1 data collection (G3&4)				×						
Wave 2 data collection (G3&4)					×					
Cohort 2										
School recruitment					×					
Teacher consent					×	×				
Random assignment					×					
Student consent							×			
4Rs + MTP implementation and professional development							×	×	×	
Wave 1 data collection (G3&4)							×	×		
Wave 2 data collection (G3&4)									×	

G3, Grade 3 teachers, students, and classrooms; G4, grade 4 teachers, students, and classrooms. SP, Spring; SU, Summer; F, Fall; W, Winter.

2.3.1.3. Aggressive behavior (child report)

Children reported on their own aggressive behaviors using the Aggression Scale (Orpinas and Frankowski, 2001). This scale is comprised of six items that ask children to report how many times they have engaged in specific aggressive behaviors over the past couple of weeks (0 = Never; 1 = Once or twice; 2 = A few times; and 3 = Many times). Examples of items are "I teased a kid at school" and "I pushed, shoved, or hit a kid at school." Scale reliabilities ranged from $\alpha = 0.82$ in fall/winter to $\alpha = 0.79$ in spring.

2.3.1.4. Aggressive behavior (teacher report)

Children's aggressive behaviors were assessed using teachers' reports on the aggression subscale of the Behavioral Assessment System for Children (BASC-AGG; Kamphaus and Reynolds, 1998). Teachers responded 14 questions regarding the frequency of a particular child behavior over the past 30 days on a 4-point scale (1 = Never; 2 = Sometimes; 3 = Often; 4 = Almost always). Example items included "Argues when denied own way" or "Is a sore loser." Cronbach's alphas for this study ranged from $\alpha = 0.92$ in fall/winter to $\alpha = 0.91$ in spring.

2.3.1.5. Student social competence (teacher report)

To assess student social competence, an average of 19-items from the teacher-reported Social Competence Scale (Conduct Problems Prevention Research Group, 1990) was used. Teachers rated 19 items regarding child behavior over the past 30 days on a 4-point scale (1 = Never; 2 = Sometimes; 3 = Often; 4 = Almost always). Sample items included "Expresses needs and feelings appropriately" and "Cooperates with peers without prompting." Internal consistency was high across the fall/winter and spring waves ($\alpha = 0.96$ and $\alpha = 0.98$, respectively).

2.3.1.6. Hostile attribution bias (child report)

Children's self-reported hostile attribution bias (HAB) was measured using a 6-item adaptation of the Home Interview (Dahlberg et al., 1998) developed initially by Dodge et al. (1986). In this version, children are presented with six visual and verbal representations (vignettes) of ambiguous but provocative social scenarios. Following the presentation of each vignette, children were presented with four possible causal attributions regarding the intent of the provocateur and were asked to select one causal attribution. Two attributions refer to the provocateur's intent as benign or accidental = 0 (e.g., The ball slipped and hit you), and two responses describe the provocateur's intent as hostile or purposeful = 1 (e.g., the student was being mean). Responses were coded as either 1 (hostile) or 0 (benign), and then averaged across items, with higher scores indicating greater hostile attribution bias. This measure had adequate internal consistency across both assessment waves (α's = 0.74 to 0.78).

2.3.1.7. Aggressive interpersonal negotiation strategies (child report)

Following the assessment of their attributions of intent with scenarios from the Home Interview described above (Dahlberg et al., 1998), children were asked what they would do next in each of the six scenarios and then were asked to select one from among four possible response strategies. Responses were coded as either 1 (aggressive; e.g., Break something that belongs to that child) or 0 (non-aggressive; e.g., Not play with the child again) and then averaged across items. The *Aggressive Interpersonal Negotiation Strategies* score is created averaging children's responses across items, with higher scores indicating greater tendencies to react aggressively. Internal consistencies ranged from $\alpha = 0.86$–0.87.

2.3.1.8. Quality of classroom interactions (observer rated)

The quality of classroom interactions was measured using the Classroom Assessment Scoring System-Upper Elementary Version (CLASS-UE; Pianta et al., 2012). CLASS-UE is an observational measure to evaluate three domains of teacher-student interactions: emotional support, classroom organization, and instructional support.

Emotional support refers to a teacher's skills and strategies in providing safe and supportive environments, where students feel secure and become more self-reliant in their explorations during problem-solving situations, feel positively related to others, and autonomous (Roeser et al., 2000; Hamre and Pianta, 2005). The domain of emotional support consists of three dimensions: positive climate, teacher sensitivity, and regard for student perspectives. Classroom organization refers to a teacher's competence in providing structured, organized, and sequenced practices that help develop children's self-regulatory skills (Blair, 2002). The domain of classroom organization consists of three dimensions: behavior management, productivity, and negative climate (reverse-coded). Instructional support refers to the pedagogical strategies a teacher uses to help children develop a sense of curiosity for learning, think about their learning and thinking processes (Baird, 1986), and in general, promote cognitive and language development in the classroom (Hamre et al., 2013). The instructional support domain consists of five dimensions: instructional learning formats, content understanding, analysis and inquiry, quality of feedback, and instructional dialog. Each dimension of CLASS is rated on a scale of 1 (low) to 7 (high).

A live classroom observation was conducted in each participating teacher's classroom. A team of 18 classroom observers who were trained to reliability and certified on the CLASS-UE conducted the observations, and a single observer rated each of the 11 dimensions for each classroom. Domain scores were then calculated by taking the average of the dimension scores within each domain.

As we have reported elsewhere (Doyle et al., 2022), interrater reliability (IRR) was calculated using the 50 observations (16%) that were double-coded at observation 1 and the 39 observations (12%) that were double-coded at observation 2. IRR was calculated using a one-way random intraclass correlation (ICC), which captures rater consistency across two measured constructs (Shrout and Fleiss, 1979). The ICC is a conservative measure of interrater reliability, as it includes both the variability within and across observers. ICCs can range from −1 to +1, with values less than 0.5 indicating poor reliability, values between 0.50 and 0.75 indicating moderate reliability, values between 0.75 and 0.90 indicating good reliability, and values greater than 0.90 indicating excellent reliability (Koo and Li, 2016). In the current study, ICCs were 0.62 and 0.74 for Emotional Support, 0.45 and 0.88 for Classroom Organization, and 0.59 and 0.72 for Instructional Support at observation 1 and observation 2, respectively.

2.3.2. 4Rs+MTP program implementation and supports

All measures of program implementation and implementation supports are presented in Table 4. Correlation analysis of the 11 measures of implementation quality shows that each measure of implementation or implementation supports is significantly correlated with at least one other measure at or above $r = +/-0.2$. In addition, none of these correlation coefficients are high (<0.6), which suggest that measures of the same construct (dosage, adherence, and responsiveness) explain some but not all variation among these constructs (see Supplementary Table 1).

2.3.2.1. Program implementation

Dosage or exposure to the seven units of the 4Rs curriculum was measured through teachers' report of the number of units they implemented in their classroom. Exposure to a curricular unit was defined as the implementation of an entire unit consisting of at least three activities (one read aloud of the target book, one lesson activity, and one additional activity). Units with less than three activities implemented were considered incomplete. Overall, 72% of teachers completed unit 1; 80%, unit 2; 73%, unit 3; 62%, unit 4; 41%, unit 5; 31% unit 6; and 18%, unit 7. The proportion of units implemented in the classroom out of the seven total possible units in the curriculum was considered an indicator of exposure to program units (see Supplementary Tables 2, 3).

Adherence to program implementation was measured through coaches' reports of teachers' adherence to program activities. Coaches watched videos of their teachers' implementation of activities in the classroom and evaluated teachers' implementation of the action plan agreed on during the coach-teacher conference by rating a one item-sentence: "There was evidence in this video that the teacher implemented the action plan from the last cycle." This item was rated using a 1 to 4 scale where 1 = no; 2 = Parts of the plan; 3 = Yes, the whole plan; 4 = Not Applicable. For cohort 1, coaches rated this item at the end of each of the eight cycles, whereas for cohort 2, coaches rated teachers' videos only following cycles two and six. To be consistent with the measure of teachers' adherence across cohorts, only ratings for cycle two ($M = 2.29$, $SD = 0.51$) and cycle six ($M = 2.45$, $SD = 0.53$) were considered in both cohorts and then aggregated to create one average score of Adherence to program implementation.

2.3.2.2. Implementation supports

Teacher's responsiveness to coaching cycles was assessed with a 4-item measure. Specifically, following coaching cycles two and six, coaches rated teachers on a 4-point scale (1 = *Strongly Disagree; 2 = Disagree*; 3 = *Agree; and* 4 = *Strongly Agree*) for the following items: "The teacher's responses to questions were in-depth and detailed" and "Based on the teacher's responses, she/he appears engaged in the prompt process." Cronbach's alphas for this scale were 0.83 for cycle two, 0.81 for cycle six, and 0.82 for the aggregate of cycles two and six. A composite score with the aggregates of the two cycles was used to represent a teacher's responsiveness to cycles.

Teachers' ratings of the worth of their coach's consultation (consultancy worth) were collected during cycles two and six using a 9-item scale that assessed their satisfaction with various components of the coaching process, including web resources, coaching consultation, and productivity. Using a 4-point response range (1 = *Strongly Disagree; 2 = Disagree*; 3 = *Agree;* and 4 = *Strongly Agree*), teachers rated items such as "This meeting helped me identify specific strategies that I can use in my classroom" and "These prompts focused on issues that were relevant to my practice." Internal consistency for cycles two and six were high ($\alpha = 0.90$ and 0.91, respectively). A single score of consultancy worth was created with the average of the nine items aggregated across cycles. The decision to measure teacher's responsiveness to coaching cycles and consultation worth in only two of the eight cycles was driven by efforts to minimize the burden teachers might experience during the evaluation of implementation supports.

TABLE 4 Descriptive statistics of implementation variables.

	n	*n* Missing	Missing%	Mean	*SD*	*Min*	*Max*	*ICC Coach*
Teacher's responsiveness to cycles	144	7	4.6	3.3	0.4	2.1	4	0.11
Consultancy worth	144	7	4.6	3.5	0.4	1.9	4	0.06
Words in prompt responses	145	6	4.0	64.6	30	20.9	210	0.09
Prompt access time elapsed	145	6	4.0	7.1	7.4	0.0	61	0.27
Teacher's responsiveness to training	147	4	2.6	4.6	0.4	3.2	5	0.10
Teacher alliance	147	4	2.6	3.4	0.6	0.9	4	0.26
Couching cycles completed	145	6	4.0	7.4	1.3	1.0	8	0.25
Time in conferences	145	6	4.0	30.9	8.3	16.2	54	0.62
Attendance to training	146	5	3.3	5.4	0.9	2.0	6	0.05
Amount of program activities implemented in classroom	142	9	6.0	28.9	13.4	3.0	87	0.22
Adherence to program activities	144	7	4.6	2.4	0.4	1.0	3.5	0.31

ICC Coach, Intraclass correlation of teachers nested within coaches.

The number of words teachers used in their responses to coach prompts was used as a proxy for the time and effort teachers invested in their own professional development through the 4Rs + MTP program (Downer et al., 2009b). The total "words in the prompt responses" were summed within cycles and then averaged across the cycles completed for each teacher. In addition, the time elapsed between the moment the coach posted the prompt and the moment the teacher accessed it was used as a proxy for a teacher's interest and engagement with the coaching process. This 'prompt access time elapsed' was averaged across cycles completed for each teacher.

Teachers' responsiveness to in-person training sessions was evaluated using teachers' ratings of 20 items across four domains: trainer's knowledge, training learning environment, organization and materials, and learning outcomes of the training session. Pearson correlations among domains ranged between 0.72 and 0.94 across the 6 days. Cronbach's alphas for the 20 items ranged between 0.86 and 0.98 across the six training days. Given the high correlation between domains and reliability of the 20 items in general, items were averaged within and then across the six training sessions to create an aggregate score reflecting teachers' responsiveness to training.

Teacher-coach working alliance was measured at the end of the coaching cycles using coach ratings of 34 items from the Measure of Coach and Teacher Alliance–Coach Report (Bradshaw et al., 2009). This scale measures five domains: Working Relationship, Coaching Process, Investment, Benefits of Coaching, and Barriers to Coaching. Scores were averaged across domains to create a global teacher-coach working alliance score with an overall Cronbach's alpha of 0.96.

Four measures of dosage or exposure to implementation supports were considered: (1) number of coaching cycles completed by teacher, (2) time teacher spent in conferences with their coach, (3) teacher attendance at the training, and (4) time teacher spent visiting the 4Rs + MTP intervention website. The number of coaching cycles completed by teacher ('Coaching cycles completed') during the intervention was determined by teachers responding to their coaches' prompts. Coaches and teachers also reported on their contacts using the web platform, noting the total time spent during each conference. Reports of time spent in the conference were averaged across completed cycles ($M_{coaches} = 28.54$ min, $SD = 7.77$; $M_{teachers} = 34.15$ min, $SD = 12.26$) and then one score with averages of coach and teacher reports was created as an indicator of "teachers" time spent in conference' with coaches. A measure of teacher attendance in training was computed as the number of days a teacher attended in-person sessions of training out of the 6 days total days of training that were provided. Finally, to evaluate teacher exposure to teaching resources available through the 4Rs + MTP intervention website, the web platform captured information about the amount of time spent by teachers visiting the library page with text and video examples of high-quality teacher practices and visiting their confidential consultancy page where teachers could watch their teaching practices as edited by their coach. From this web usage data, the total amount of "time spent visiting the website" was calculated as an indicator of web-based exposure to the intervention. A cut-off maximum of 15 mins per page visit was used to correct for the times teachers ended their web session but forgot to logout.

2.3.3. Covariates

2.3.3.1. Professional burnout

Teacher burnout was assessed using the emotional exhaustion and personal accomplishment subscales of the Maslach Burnout Inventory-Educator Survey (MBI-ES; Maslach et al., 1986, 2001). The depersonalization subscale was not included as it has shown poorer internal consistency when compared to the other two scales (Schaufeli et al., 2001); therefore, it was excluded from the survey to reduce survey length. The emotional exhaustion subscale included nine items (e.g., "I feel emotionally drained from my work") and the personal accomplishment subscale included eight items (e.g., "I feel exhilarated after working closely with my students"). Teachers were instructed to report the frequency with which they experienced the job-related stressors using a 7-point Likert scale ranging from 0 ("never") to 6 ("every day"). Emotional exhaustion and personal accomplishment both showed acceptable internal consistency with Cronbach's alpha values of 0.92 and 0.72, respectively.

2.3.3.2. Depression, anxiety, and stress

The Depression, Anxiety, and Stress Scale—Short Form (DASS-21; Lovibond and Lovibond, 1995), is a self-report measure that assesses symptomatology of depression, anxiety, and stress among adults. Each of the three subscales contains seven items. Teachers

rated the degree to which given statements applied to them over the past week on a 4-point Likert-type scale ranging from 1 (Did not apply to me at all) to 4 (Applied to me very much, or most of the time). Sample items include "I feel that I had nothing to look forward to" (depression), "I was worried about situations in which I might panic and make a fool of myself" (anxiety), and "I found it hard to wind down" (stress). In the current study, the three subscales were moderate to strongly correlated ($r = 0.47–0.67$) and there was strong internal consistency among the items (Cronbach's alpha = 0.90), therefore, the DASS-21 total score was used. Prior to conducting analyses, the DASS-21 total score was transformed, using the natural logarithm to base 10, to reduce the level of skewness and kurtosis. Before transformation, the mean DASS-21 score was 0.24 ($SD = 0.30$) and after transformation the mean DASS-21 score was 0.08 ($SD = 0.09$).

2.3.3.3. Teacher psychological wellbeing

The Psychological Well-Being Scale is a self-report measure that assesses teachers' autonomy, personal growth, and positive relations with others (Ryff and Keyes, 1995). Each scale contains 7 items. Teachers rated the degree to which they agree with personal statements on a 7-point Likert-type scale ranging from 1 (strongly agree) to 7 (strongly disagree). Sample items included: "I tend to be influenced by people with strong opinions" (autonomy), "I am not interested in activities that will expand my horizons" (personal growth), and "Most people see me as loving and affectionate" (positive relations with others). The internal consistency for the total scale was moderate to high (Cronbach's alpha = 0.72), therefore the total score based on an average of the 21 items was used for analysis in this study.

2.3.3.4. Positive and negative affect scale

The Positive and Negative Affect Scale (PANAS; Watson et al., 1988) consists of 10 words that describe positive and negative emotions. Teachers read each word and indicated the extent to which they had felt that way during the past few weeks on a 5-point Likert type scale ranging from 1 (Very Slightly or Not at All) to 5 (Extremely). Sample items of the subscales included: "Enthusiastic" (positive affect) and "Irritable" (negative affect). Internal consistencies were high for both positive affect (Cronbach's alpha = 0.89) and negative affect (Cronbach's alpha = 0.83).

Additional teacher-reported teacher and classroom demographic characteristics were considered as covariates, including teachers' race/ethnicity, teachers' years of experience, classroom type (e.g., special education, ICT/CTT), class size, proportion of students in the classroom with Individualized Education Plans, proportion of students in the classroom with Limited English Proficiency, proportion of female students in the classroom, and proportion of behaviorally at-risk students in classroom (i.e., students above norm cut-off scores on aggression and/or conduct problems were coded as 1). Cohort (Cohort 1 = 0; Cohort 2 = 1) was included as a covariate in all analyses.

3. Results

3.1. Missing data

Percentage of missing data on implementation variables was low, ranging between 2.6 and 6%. Table 4 shows descriptive statistics of quality of implementation variables and percentage of missing observations per variable. Several child level and teacher-classroom level variables showed more than or close to 10% of missing data. At the child level, Math (56%) and ELA score (57%) at wave 1 showed the highest proportion of missing data, whereas teacher burnout (9.7%) and teacher negative affect (9.4%), showed the highest proportion at the teacher/classroom level. Tables 5, 6 show descriptive information and proportion of missing observations of child outcomes and teacher-classroom variables, respectively. To examine whether data were missing completely at random (MCAR), tests for child level and teacher-classroom level variables at both wave 1 and wave 2 were performed using the function TestMCARNormality from the R package MissMech (Jamshidian et al., 2014). An additional test to detect missing patterns (Little, 1988) was also conducted on both datasets using the function LittleMCAR from the R package BaylorEdPsych (Beaujean and Beaujean, 2012).

Intraclass correlations of student variables ranged between 0.06 and 0.37, and between 0.02 and 0.19 when clustered by teacher ID and by school ID, respectively (see Intraclass Correlations, ICC, Table 5). Since shared variance of children's outcomes clustered by school were low (McCoach and Adelson, 2010), only teacher identification number was included as a cluster variable in the models for this study. The dataset including the full sample of teachers and students was imputed using the R package multiple imputation with multivariate imputation by chained equation (MICE; Zhang, 2016). Variables at level 2 (teacher/classroom) were imputed using the function 2only.pmm that aggregates level-1 predictors and imputes the level-2 variables using predictive mean matching (pmm; Kleinke, 2017). Variables at level 1 (child level) were imputed using random forest (Shah et al., 2014). Each imputation was performed separately for treatment and control groups and then combined into a single dataset. Twenty imputed datasets were used for analyses. Results of analysis of missing patterns and a description of multiple imputation procedures is presented in Supplementary Data Sheet 1.

3.2. Main analyses

3.2.1. Identifying profiles of quality of implementation among teachers

The first aim of this study was to identify teacher compliance propensity by exploring which components of program implementation and implementation supports would discriminate between profiles of quality of implementation for teachers in the treatment group (see Figure 1A). Latent Profile Analyses (LPA) was used to estimate the probability of teachers belonging to different profiles (clusters; Oberski, 2016) of quality of implementation. The LPA was estimated using the array of assessments of program implementation and implementation support (described in section Measures). Implementation variables were standardized. Intraclass correlations (ICC) of implementation variables suggested that part of the variances could be attributed to the nested structured of the data (i.e., teachers nested within coaches; *ICC*, $M = 0.21$, Min = 0.54, Max = 0.62; see Table 4). To account for the effect of coaches, dummy codes for coach were included in the subsequent LPA.

LPA analysis was performed in Mplus using maximum likelihood estimation to handle missing data (Muthén, 2004). The model of two profiles of implementation showed a good fit (*Loglikelihood* = −2077.53, *AIC* = 4237.07, *BIC* = 4359.67, *Entropy* = 0.99). Of the 147 teachers

TABLE 5 Descriptive statistics of child outcomes.

	n	Miss%	Mean	*SD*	*Min*	*Max*	*ICC* teacher	*ICC* school
Hostile attribution bias	4,092	19.5	0.37	0.33	0	1	0.06	0.02
T1 hostile attribution bias	4,435	12.7	0.36	0.31	0	1	0.06	0.02
Aggressive interpersonal strategies	4,089	19.5	0.21	0.32	0	1	0.09	0.03
T1 aggressive interpersonal strategies	4,432	12.8	0.17	0.29	0	1	0.08	0.02
Internalizing symptoms	4,055	20.2	0.42	0.22	0	1	0.06	0.02
T1 internalizing symptoms	4,407	13.3	0.43	0.21	0	1	0.06	0.02
Aggressive behavior child report	4,085	19.6	0.53	0.61	0	3	0.15	0.07
T1 aggressive behavior child report	4,427	12.9	0.47	0.58	0	3.3	0.12	0.05
Aggressive behavior teacher report	4,469	12.0	1.44	0.53	1	4	0.17	0.04
T1 aggressive behavior teacher report	4,718	7.1	1.41	0.53	1	4	0.18	0.06
Social competence	4,469	12.0	2.99	0.75	1	4	0.26	0.05
T1 social competence	4,718	7.1	2.92	0.74	1	4	0.24	0.07
ELA score	4,616	9.2	298.13	34.31	168	408	0.34	0.15
T1 ELA score	2,175	57.2	293.53	35.35	163	398	0.37	0.10
Math score	4,684	7.8	290.85	38.85	165	397	0.34	0.19
T1 math score	2,213	56.4	291.54	36.82	176	401	0.33	0.18
School absences	5,044	0.7	11.53	11.12	0	102	0.07	0.03
T1 school absences	4,854	4.5	12.31	11.63	0	97	0.08	0.05

T1, Time 1-baseline measure; ICC Teacher, Intraclass correlations with teacher and school ID as cluster variables.

included in the analysis (i.e., four teachers who dropped out at the beginning of the study were excluded from this analysis), 81(55%) were members of latent profile one (LP1, below average) and 66 (45%) were members of latent profile two (LP2, above average). The probability of being in LP1 was significantly predicted by two implementation variables: consultancy worth ($B = -0.89$, SE = 0.03, $p < 0.001$) and time in conferences ($B = -0.24$, SE = 0.11, $p < 0.025$); whereas four variables significantly predicted the probability of being in LP2: consultancy worth ($B = 1.08$, SE = 0.001, $p < 0.001$), time in conference ($B = 0.29$, SE = 0.120, $p = 0.015$), prompt access time elapsed ($B = -0.183$, SE = 0.079, $p = 0.020$), and teacher responsiveness to training ($B = 0.23$, SE = 0.11, $p = .040$). Figure 3 shows the standardized coefficients for each of the implementation constructs for each of the LPs.

Validation analyses of the LPs suggested differences in four of 11 implementation constructs for teachers classified in LP1 and LP2 groups, as indicated by the Welch Two Sample *t*-test (Welch, 1947; see Table 7). Specifically, compared to teachers in the LP2 (above average) group, teachers in LP1 (below average) had significantly lower responsiveness to training and ratings of consultancy worth, spent significantly less time in conferences with their coaches, and higher prompt access time elapsed (i.e., spent more time to access the prompts provided by their coaches). None of the differences between LP1 and LP2 on the other five variables of implementation support and the two variables from program implementation were significant.

Results of the LPA, validation analysis, and visual inspection of the graph of standardized coefficients by LPs suggest that LP1 reflects a profile of teachers with below average quality of implementation and LP2 reflects a profile of teachers with above average quality of implementation (see Figure 3). Differences between these two profiles can be grouped as differences in teacher responsiveness to implementation supports and the dosage of implementation support they received, with teachers in the below average profile being significantly less responsive and receiving less support than teachers in the above average profile. None of the program implementation variables significantly characterized the profiles of quality of implementation, but this result should be interpreted with caution as explained later in discussion and limitations. The four teachers who dropped the intervention and therefore did not implement the program or receive implementation support were manually assigned to LP1 (below average profile).

3.2.2. Estimating teacher compliance propensity using teacher and classroom covariates

To further address our first study aim, we then used Follmann's (2000) propensity score approach to explore teacher and classroom covariates that predicted teacher compliance propensity for teachers in the treatment group (see Figure 1B and the full sample treatment and control groups; see Figure 1C). Compliance propensity for the treated sample, that is, the propensity to be in the above average implementation latent profile (LP1 = 0; LP2 = 1) was estimated using nine baseline measures of classroom and teacher characteristics known to predict quality of implementation. These measures include proportion of students at behavioral risk, number of students in classroom, proportion of students with active IEPs, teachers' number of years of experience, teacher burnout, teacher psychological wellbeing, teacher positive affect, teacher negative affect, and teacher score of depression, anxiety, and stress (aggregated). Analysis was performed using Random Forest, a machine learning technique (Zhao et al., 2016) robust to non-normal data which performs well with multivariate data of different formats (continuous and categorical). Random Forest uses recursive partitioning, which applied to

TABLE 6 Descriptive statistics of classroom and teacher variables.

	n	Missing %	Mean	*SD*	*Min*	*Max*
Emotional support	314	4.8	4.30	0.91	2.0	7.0
T1 emotional support	320	3.0	4.52	0.82	2.2	6.5
Instructional support	314	4.8	3.32	0.87	1.0	5.5
T1 instructional support	320	3.0	3.55	0.80	1.4	6.25
Classroom organization	314	4.8	5.94	0.74	3.3	7.0
T1 classroom organization	320	3.0	5.89	0.70	2.8	7.0
Proportion student at risk	318	3.6	0.20	0.19	0.0	0.9
Number of students	321	2.7	22.36	5.92	6.0	33.0
Year of experience	322	2.4	10.71	7.56	1.0	40.0
Proportion of students with IEP	321	2.7	0.15	0.23	0.0	1.0
Teacher burnout	298	9.7	1.55	0.86	0.0	4.3
Teacher psychological wellbeing	301	8.8	5.98	0.55	3.9	7.0
Teacher positive affect	302	8.5	3.97	0.69	1.6	5.0
Teacher negative affect	299	9.4	1.68	0.62	1.0	4.4
Teacher depression, stress, and anxiety	301	8.8	0.25	0.29	0.0	1.9

T1, Time 1-baseline measure.

propensity score estimation in this study consists of recurrently splitting the data into nodes (i.e., groups) based on values of the categories of a categorical covariate or a cutoff applied to a continuous covariate that discriminates between quality of implementation profiles (LP1 and LP2). Propensity scores can be obtained as the proportion of cases in the above average quality of implementation profile at each terminal node. Random Forest was implemented using the cforest unbiased function from the R package "Party" (V1.3-5; Hothorn et al., 2006). Bias was controlled by running a large number of trees with bootstrapped samples of the same size as the original sample and combining results, options were set to 1,000 trees and a random sample of $m = 3$ predictors, chosen from all possible predictors $p = 9$ using the formula $m = \sqrt{p}$. Compliance propensity was estimated for each imputed dataset and then aggregated across datasets to create an average compliance propensity score (Mean compliance = 0.43; SD = 0.11; Skewness = 0.24; Kurtosis = −0.64). The distribution of compliance propensities was similar across datasets. Density plots of compliance propensity across 20 imputed datasets can be found in Supplementary Figure 1. Not all nine baseline classroom and teacher characteristics contributed equally to the estimation of compliance propensity as suggested by coefficients of mean decrease difference in accuracy (MDD). Covariate coefficients of mean decrease of accuracy greater than zero indicate that the absence of such covariates have an impact in decreasing the accuracy of the model (Louppe et al., 2013). MDD was computed using the function varimp and allowing association with covariates with a threshold of 0.2, which implies that the resulting predictor importance score is conditional on the importance of other predictors similar to beta coefficients in regression models (Strobl et al., 2008). According to this indicator, number of years of experience as a teacher was the most important covariate predicting compliance propensity, with the highest mean decrease in accuracy coefficient ($MDD = 0.015$), followed by teacher burnout (MDD = 0.052). The other seven covariates have MDD coefficients lower than 0, suggesting low to zero contribution to the model of classification (see Supplementary Table 4A). Further analysis in the treated sample show that teacher compliance propensity significantly predicted three variables related to implementation supports: teacher's ratings of consultancy worth ($B = 0.68$; $SE = 0.51$), average time in conference ($B = 0.23$; $SE = 0.70$), and teacher's responsiveness to training ($B = 0.24$; $SE = 0.69$) These findings suggest the resulting propensity scores estimated by years of experience and teacher's levels of burnout, are significantly associated with variables of implementation supports that characterized the above and below profiles of quality of implementation (see Supplementary Table 4B). Compliance propensity for teachers in the control group was predicted using the weights of the nine covariates based on the propensity estimation conducted for teachers in the treatment group. Visual inspection of box and whisker plots suggested an adequate area of common support, that is, the area of the distribution of compliance propensity includes values for teachers in the treatment and control group. To evaluate covariate balance across treatment and control conditions, analysis of covariance (ANCOVA) was used to examine differences on each covariate between treatment and control groups at above and below average quality of implementation after controlling for compliance propensity. Results showed that interactions between random assignment and levels of above and below average compliance propensity were not significant for any of the covariates used to predict profiles of quality of implementation, suggesting balance was achieved for all covariates across the randomly assigned conditions (see Supplementary Table 5).

3.2.3. Relations among school random assignment to 4Rs+MTP, teacher compliance propensity, quality of classroom interactions, and child outcomes

Our second aim was to examine teacher compliance propensity as a moderator of the relationship between school random assignment to 4Rs + MTP and (a) quality of classroom interaction (see Figure 2, paths P2 and P4), (b) children's academic and SEL outcomes (see Figure 2, paths P1 and P5), and (c) children's academic and SEL

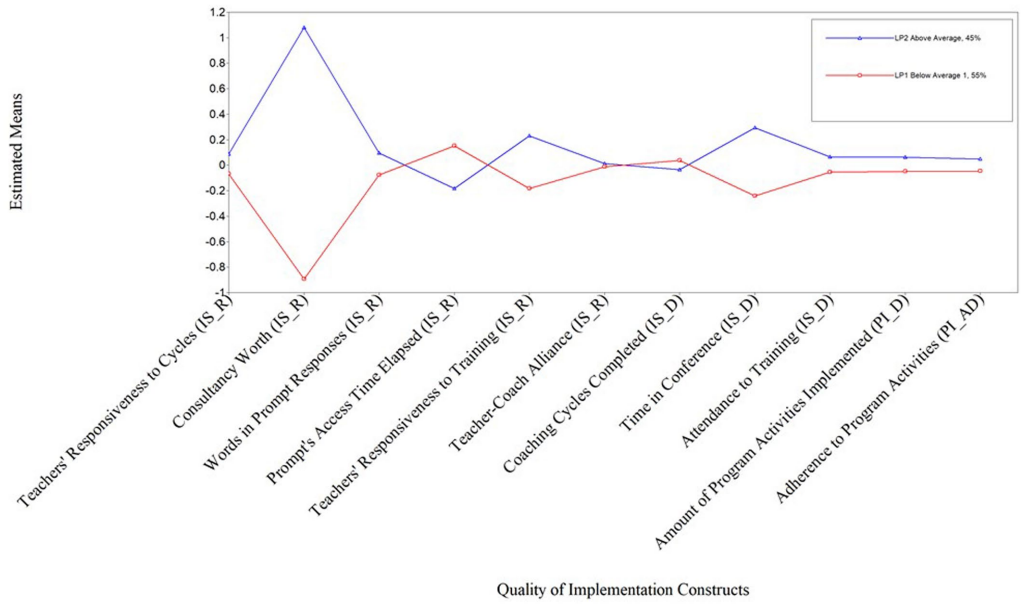


FIGURE 3
Estimated means of quality of implementation constructs by latent profile IS_R, Implementation Support Responsiveness; IS_D, Implementation Support Dosage; PI_D, Program Implementation Dosage; PI_AD, Program Implementation Adherence.

TABLE 7 Mean differences of quality of implementation constructs by latent profile.

	LP1 Below average	LP2 above average			
Implementation support	Mean (*SD*)	Mean (*SD*)	*t*	*df*	*p*
Teacher's responsiveness to cycles	3.27 (0.36)	3.34 (0.41)	−1.13	128.23	0.259
Consultancy worth	3.15 (0.22)	3.82 (0.17)	−20.86	141.66	0.001
Words in prompt responses	62.31 (30.33)	67.33 (29.56)	−1.01	139.61	0.316
Prompt's access time elapsed	8.24 (8.87)	5.77 (4.79)	2.13	123.82	0.035
Teacher's responsiveness to training	4.51 (0.35)	4.68 (0.34)	−2.84	140.44	0.005
Teacher alliance	3.34 (0.65)	3.39 (0.49)	−0.52	144.31	0.606
Couching cycles completed	7.47 (1.22)	7.33 (1.40)	0.61	130.10	0.540
Time in conferences	28.87 (7.93)	33.42 (8.07)	−3.41	137.56	0.001
Attendance to training	5.37 (1.03)	5.49 (0.81)	−0.80	143.96	0.425
Program implementation					
Amount of program activities implemented in classroom	28.26 (12.52)	29.75 (14.36)	−0.65	126.02	0.514
Adherence to program activities	2.34 (0.44)	2.38 (0.44)	−0.56	136.02	0.574

outcomes as mediated by the quality of classroom interactions (see Figure 2, paths P2, P4, and P3).

Here we used multilevel modeling to test the relations between random assignment (a level-2 predictor) and child academic and SEL outcomes (level-1 outcomes), mediated by quality of classroom interaction (a level 2 variable). This multilevel mediation is examined at different levels of teacher compliance propensity (a level-2 moderator), using an adaptation of the general path analytic framework proposed by Edwards and Lambert (2007) for testing direct, indirect, and total effects on an outcome at different levels of a moderator. Although Edwards and Lambert (2007) do not discuss the moderated mediation in a multilevel path, to account for the nested structure of students within teachers-classrooms in this study, a random intercept is also included in the regression equations (Tingley et al., 2014; Rockwood, 2017; Finch, 2022).

The multilevel moderated mediation model was fit using the SemTools R package (Jorgensen et al., 2019), which allows the estimation of multilevel analyses in multiple imputed datasets using the structural equation models (SEM.mi) function. Inspection of correlations among residuals (higher than 0.1) and modification indices suggest covariation between domains of quality of classroom interaction, therefore covariances among emotional support, classroom

organization, and instructional support were included in the model. Teacher ID was used as a cluster variable. The model converged in the 20 imputed datasets and results from the pooled fit measures showed an adequate fit x^2 (134) = 386.69 $p < 0.01$, CFI = 0.91, TLI = 0.84, RMSEA = 0.02. Rubin's (1987) rules were used to pool point and *SE* estimates across 20 imputed data sets, and to calculate degrees of freedom for each parameter's *z*-test and 95% confidence interval (*CI*). Coefficients were tested with an $\alpha = 0.05$, two-tailed level of significance. However, trends in the hypothesized direction are reported using one-tailed tests (90% *CI*). To test direct and indirect effects, robust confidence intervals were estimated using a Monte Carlo test of mediation (MacKinnon et al., 2004) with 1,000 random samples with population values equal to the coefficients and covariance of the sample (Preacher and Selig, 2012). Quality of classroom interaction and child outcomes were regressed on school treatment assignment (0 = control, 1 = 4Rs + MTP), compliance propensity, the interaction of treatment assignment and compliance, the three domains of quality of classroom interaction, cohort (0 = cohort 1; 1 = cohort 2), and values of the target outcome at wave 1 (grand mean centered for classroom outcomes, and group mean centered for child outcomes). Each path of the moderated mediation was tested at two levels of the moderator, teachers' compliance propensity-mean centered, namely, 1SD below average compliance and 1SD above average compliance. An example equation with a detailed explanation can be found in Supplementary Data Sheet 2.

3.2.3.1. Moderation of compliance propensity on the effects of 4Rs+MTP on quality of classroom interactions

Random assignment to 4Rs + MTP was associated at the trend level ($p < 0.10$) with positive effects on emotional support when moderated by above average compliance ($b = 0.25$, SE = 0.15, 95% CI [−0.035, 0.535]). Since the effect of treatment on emotional support followed the hypothesized direction, this direct path was tested at the 90% CI using the Monte Carlo method for mediation with 1,000 replications. Results show that the conditional effect of treatment on emotional support was significant at the 90% CI [0.0208, 0.4736]. By contrast, when evaluated at below average level of compliance, the effect of treatment on emotional support was not significant and close to zero ($b = 0.04$, SE = 0.12, 95% CI [−0.193, 0.278]). The effects of treatment on instructional support and classroom organization at below average compliance were also not significant. Table 8 shows parameters and test statistics of the direct effects of treatment on all three domains of classroom interactions, conditioned on different levels of the moderator compliance (path "a"). Parameter estimates and test statistics for each predictor in the model can be found in Supplementary Table 6.

3.2.3.2. Moderation of compliance propensity on the effects of 4Rs+MTP on children's academic and SEL outcomes

Teacher compliance propensity moderated the effect of treatment on children's school absences. Random assignment to 4Rs + MTP was associated with significantly fewer school absences than random assignment to the control group when moderated by above average teachers' compliance propensity ($b = -0.084$, SE = 0.04, 95% Monte Carlo CI: [−0.13, −0.012]). Association of random assignment to 4Rs + MTP and school absences was not significant when tested at levels below average compliance ($b = 0.03$, SE = 0.039, 95% CI [−0.045, 0.107]). Compliance propensity did not moderate the effects of treatment on the other seven children's outcomes. Coefficients and test statistics for each predictor can be found in Supplementary Table 7. See Supplementary Table 8 for parameters and test statistics of the direct effects of treatment on all child outcomes, conditioned on different levels of the moderator compliance (path "c").

3.2.3.3. Moderation of compliance propensity on the effects of 4Rs+MTP on children's academic and SEL outcomes, as mediated by quality of classroom interactions

Results of the effects of domains of quality of classroom interactions on children's academic and SEL outcomes showed a significant positive effect of classroom organization on children's academic test scores ($b = 4.61$, SE = 1.857, 95% CI: [0.013, 9.64]). In addition, there was also a trend level positive effect of instructional support on academic test scores ($b = 2.90$, SE = 1.69, 95% CI: [0.086, −0.411], $p = 0.08$). Although not significant, there were also trend level effects of quality of classroom interactions on child reported aggressive behavior. Higher emotional support was associated with lower aggressive behavior ($b = -0.03$, SE = 0.02, 95% CI: [−0.063, 0.004], $p = 0.08$). Likewise, higher classroom organization was associated with lower child reported aggressive behavior ($b = -0.03$, SE = 0.02, 95% CI: [−0.065, 0.005], $p = 0.09$). Finally, there were also trend level associations in the expected direction between classroom organization and children's aggressive interpersonal strategies ($b = -0.02$, SE = 0.01, 95% CI: [−0.032, 0.002], $p = 0.08$) and internalizing symptoms ($b = -0.01$, SE = 0.01, 95% CI: [−0.002, 0.002], $p = 0.09$).

None of the domains of quality of classroom interactions mediated the effect of treatment status on child outcomes. However, the total effect of treatment on school absences was significant when evaluated at levels of above average compliance ($b = -0.07$, SE = 0.036, 95% CI: [−0.1.44, −0.003]; see Figure 4; Supplementary Tables 7, 8).

4. Discussion

Evidence for evaluation of SEL school interventions shows that promoting children's social and emotional skills in schools can have important positive effects on children's outcomes, including academic achievement, reduced school absences, and reduced problem behaviors (Jones et al., 2019; Weissberg, 2019); and on classroom and school outcomes, such as improvements in school climate and more supportive relationships in classroom interactions (Pianta et al., 2008; Brown et al., 2010; Hamre et al., 2013). However, studies testing the role of classroom interactions as mediators of the effects of SEL programs on children's outcomes are scarce. In addition, evidence regarding the implementation of SEL programs suggests programs impacts are amplified when programs are implemented with high quality (Dane and Schneider, 1998; Durlak, 2016). Few studies of SEL programs to date include implementation variables as predictors in their program impact models (Derzon et al., 2005; Humphrey et al., 2018) or use holistic measures of quality of implementation instead of relying solely on measures of dosage (Durlak and DuPre, 2008; Domitrovich et al., 2010).

Accordingly, this study had two primary aims. First, we explored several indicators of teachers' adherence to and dosage of 4Rs + MTP program activities, and responsiveness and exposure to program implementation supports to identify teachers' profiles of quality of

TABLE 8 Effects of 4Rs+MTP on quality of classroom interactions conditioned on teacher compliance propensity (Path "a").

Paths	Teacher compliance propensity											
	Below average						Above average					
	b	*SE*	*B*	*Z*	95% CI		*b*	*SE*	*B*	*Z*	95% CI	
					Lower	Upper					Lower	Upper
TX → ES	0.04	0.12	0.01	0.36	−0.193	0.278	0.25	0.15	0.17	0.17[t]	−0.035	0.535
TX → IS	−0.10	0.13	−0.07	−0.77	−0.343	0.150	0.01	0.14	0.02	0.02	−0.273	0.278
TX → CO	−0.14	0.11	−0.11	−1.29	−0.346	0.071	−0.06	0.12	−0.03	−0.51	−0.288	0.170

TX, Random assignment to 4Rs + MTP (1) vs. Control (0); ES, Emotional Support; IS, Instructional Support; CO, Classroom Organization.

FIGURE 4
Mediation of quality of classroom interactions in the effect of random assignment to 4Rs+MTP on child school absences at different levels of teacher compliance propensity.

implementation. We then estimated teachers' propensity to comply with implementation based on teacher personal characteristics and classroom characteristics known to predict quality of implementation. Second, we examined the effects of the 4Rs + MTP program on classroom quality of interactions and children's academic and SEL outcomes at different levels of teachers' compliance propensity, including examining whether classroom quality of interactions mediated the effect of the 4Rs + MTP program on children's SEL and academic outcomes, when moderated by levels of teachers' compliance propensity.

This is the first study to date to take a compliance propensity approach to understanding whether and how quality of implementation moderates the effects of an SEL program on classroom and child outcomes, and as such it is somewhat exploratory in nature. It is important to state that the compliance propensity approach to examining quality of program implementation may lack precision relative to the use of direct indicators of treatment teachers' implementation quality. However, the strength of this approach is that it affords inclusion of teachers in the control group for whom there were no direct measures of implementation, thus reducing the corresponding bias when making causal inferences about treatment effects.

4.1. Profiles of teachers' quality of implementation and compliance propensity

Teachers' quality of implementation of the 4Rs + MTP program was represented by two profiles consisting of teachers with below and above average quality of implementation. Teachers in the above average quality of implementation profile were characterized by their high responsiveness and their high exposure to implementation supports; whereas teachers in the below average quality of implementation profile were characterized by their low responsiveness and low exposure to implementation supports.

Traditionally, research on implementation has distinguished between high and low quality of implementation using measures of program implementation, such as dosage and adherence to program implementation (Dusenbury et al., 2004). Findings in the current study suggest that the distinction between high and low quality of implementation might also be characterized by measures of teachers' responsiveness and dosage of implementation supports. These findings support claims from researchers in implementation science suggesting the importance of including measures of dosage and responsiveness to implementation supports, such as time spent with coaches, teacher's ratings of consultancy worth and teacher responsiveness to training, also documented in the literature of quality of implementation as central to program impact (Domitrovich et al., 2010; Wehby et al., 2012; Pas et al., 2015).

Profiles of quality of implementation were established based on distinguishable patterns in the implementation supports teachers received through the 4Rs + MTP program. However, measures of dosage and adherence to program implementation did not play a significant role in the characterization of teacher's profiles of quality of implementation in this study. These results may be seen as contradictory with past literature about the positive associations between the quality and quantity of coaching and training teachers received and teacher's dosage of and adherence to program implementation (Downer et al., 2009a,b; Wehby

et al., 2012; Patti et al., 2015; Ashworth et al., 2018). In accordance with this literature, it was expected that both measures of implementation supports and program implementation would significantly discriminate between above and below teacher's profiles of quality of implementation. However, it is worth noting that, as described in our methods, measures of dosage of program implementation were provided by teacher's self-reports of the number of activities they implemented in their classrooms. Previous findings show that teacher's self-reports of dosage are frequently high, which compromises the ability of these kind of measures to discriminate between high and low dosage of implementation (Domitrovich et al., 2010). Further research should include observed measures of dosage in program implementation, which have been found to be more reliable than self-reported measures (Durlak and DuPre, 2008; Domitrovich et al., 2010). Regarding adherence to implementation of curricular activities in the classroom, differences on this coach reported measure might have been obscured due to the decision to control for coach assignment during the estimation of profiles, and thus minimize the potential bias of coach's idiosyncratic judgments. Differences in teacher-coach alliance, also reported by coaches, and found to be a significant predictor of program implementation in previous research (e.g., Wehby et al., 2012) might have also been affected by the decision to control by coach assignment in this study.

4.2. Compliance propensity and teacher and classroom characteristics

Years of experience as a teacher and professional burnout were the sole contributors to the estimation of teacher compliance propensity. Findings suggest that more experienced teachers and teachers who reported lower levels of burnout were more likely to be high compliers in implementing the 4Rs + MTP program. Although exploratory, these results are consistent with prior evidence that shows links between variation in teacher professional experience and program implementation (Fixsen et al., 2009; Weiss et al., 2014), specifically findings that early career teachers report low quality of implementation (Domitrovich et al., 2019) and more experienced teachers spend more time in conferences with their coaches (Downer et al., 2009a). In addition, recent evidence from a randomized trial testing the integration of the Good Behavior Game and MTP among early career teachers found significant intervention effects on student behavior and achievement but only among those teachers exhibiting high baseline levels of distress and disruptive behavior in their classrooms (Tolan et al., 2020). Interestingly, in the current study, proportion of children in the classroom considered at behavioral risk was not significantly associated with teacher implementation despite prior evidence to the contrary (Musci et al., 2019).

4.3. Compliance propensity and relationships among 4Rs+MTP, classroom quality of interactions and children's outcomes

There was a trend level effect in the moderation of compliance propensity on the effects of 4Rs + MTP on classroom emotional support. When examined at above average levels of compliance, 4Rs + MTP had a positive effect on classroom emotional support. This effect, however, was negligible for teachers with below average compliance propensity. Previous work has found that teachers receiving support through on-going coaching and web-based material in the MyTeachingPartner (MTP) coaching program were better at proving emotional and instructional support to their students (Pianta et al., 2008). Further, as noted above, preliminary findings from the efficacy trial of the 4Rs + MTP program show positive main effects of overall exposure to 4Rs + MTP implementation on classroom emotional support (Brown et al., 2019). Although the aims of the current study did not include the evaluation of specific components of quality of implementation on classroom quality of interactions, findings suggest the effect of the 4Rs + MTP program on emotional support is amplified among teachers who have a higher propensity to comply with the implementation, including their propensity to benefit from the coaching support provided through MTP.

Students´ social and emotional skills are effectively taught and learned within caring and supportive environments (Collaborative for Academic, Social, and Emotional Learning, 2017), where students feel secure and positively related to others (Roeser et al., 2000; Hamre and Pianta, 2007). However, teachers also need social and emotional skills to build caring and supportive relationship with their students (Kingston and Wilensky, 2018). The support teachers receive from coaches during program implementation might improve the emotional resources teachers need to provide higher quality emotional support in their classrooms. For instance, other research has found that teachers receiving ongoing coaching have reported increased self-awareness, self-management, and improved relationships with students (Patti et al., 2015). Moreover, strong teacher-coach alliance buffered the negative effects of teacher burnout on teacher implementation of SEL activities with their students (Wehby et al., 2012).

By contrast, the effects of 4Rs + MTP on instructional support and classroom organization were not moderated by teachers' compliance propensity. More research is needed to understand the influence of variations in quality of implementation in the effect of 4Rs + MTP on instructional support and classroom organization.

Compliance propensity also moderated the effects of 4Rs + MTP on child school attendance. Specifically, when examined at above average compliance propensity, 4Rs + MTP was associated with fewer school absences. This effect, however, was negligible for children from teachers with below average compliance propensity. This study provides evidence about the role of SEL programs in improving children's school attendance when the program is well implemented, and in this case, when teachers have high propensity to implement the program with high quality.

This contribution is relevant in the context of SEL program implementation, considering the limitation of most SEL programs with regard to having significant impacts on academic attainment (less than 10% in the United States; Grant et al., 2017). In a prior quasi-experimental study of the 4Rs + MTP program, the integration of MTP coaching for teachers with the prior 4Rs intervention model also yielded positive effects on children's school attendance relative to the 4Rs program without MTP coaching (Doyle et al., under review).

Previous literature suggests that highly supportive classrooms are likely to encourage students' attendance (Barth, 1984; Baker et al., 2001; McCluskey et al., 2004). However, none of the domains of classroom quality of interactions mediated the effects of 4Rs + MTP in reducing school absences, suggesting that the mechanism through which 4Rs + MTP is associated with lowering child absences at the

above average levels of compliance propensity is not explained by overall higher quality classroom interactions. An alternative explanation is that teachers with high compliance propensity provide effective support to individual students, perhaps those at higher risk of truancy, without necessarily extending this support (or not extending it to the same degree) to all children in the classroom. Interventions targeting individuals or small groups at risk of truancy, instead of at the classroom as a whole, have been common practice in education (Teasley, 2004; Reid, 2013). Findings might suggest that these teachers in 4Rs + MTP would develop the skills needed to provide effective support to students at risk of truancy, resulting in their higher attendance.

Finally, domains of classroom interaction quality (i.e., emotional support, instructional support, and classroom organization) did not mediate the effects of 4Rs + MTP on any of the other child outcomes at levels of above average compliance propensity. It is possible these nonsignificant effects are a function of the mediating mechanism (i.e., domains of classroom interaction quality) not actually being evident when examined based on different levels of the moderator. Since teacher years of experience and burnout were the main predictors of compliance propensity, examining the effects of treatment at above average compliance propensity is virtually equivalent to examining the effects of treatment at high levels of teaching experience and low levels of burnout. It is possible, that experienced teachers with low levels of burnout might have developed strong skills in promoting effective classroom quality of interactions, such that treatment differences are noticeable only in one specific domain of classroom quality of interactions: emotional support.

Although, in this study there was not a significant mediation effect of emotional support on child outcomes, taken together, the evidence from previous studies and findings from the current study suggests that the quality of supports received by teachers with high compliance propensity in 4Rs + MTP would bolster or improve their emotional skills in ways that may in turn increase their ability to provide effective emotional support during their interactions with students. Thus, teachers would develop emotional skills needed to provide emotional support in their classrooms through highly supportive interactions with their coaches. More research is needed, nevertheless, to understand the effect of high-quality coaching in improving teachers' emotional skills, and the mediating role of these skills on the relationship between SEL programs such as 4Rs + MTP and the quality of teachers' emotional support in classrooms.

4.4. Limitations

One of the limitations in the current study was its sole reliance on teacher personal characteristics and classroom characteristics to estimate teacher compliance propensity with implementation, while excluding the potential effects of coach assignment to teachers. It is worth noting that coach assignment was controlled for during the identification of profiles of quality of implementation for the treated teachers. Accordingly, teachers' compliance propensity in this study should be interpreted as teachers' probability of complying with implementation given personal and classroom characteristics, and regardless of the particular influence of coaches in the implementation process. While the role coaches played in implementation might have influenced teachers´ program implementation, the decision to exclude coach assignment was the trade-off for being able to calculate compliance propensity for teachers in the control group for whom coach assignment was not available.

Some measures of child outcomes in 4Rs + MTP might be limited in terms of evaluating the effects of the program. For instance, academic measures rely only on test scores but are not sensitive to other indicators of academic performance and engagement such as the quality of a child's academic work and participation in their classroom. Although the CLASS observation measure assessed information on interaction quality at the classroom level, it did not provide information about child level interaction quality, which may be a distinctly sensitive indicator of children's classroom interaction experience. Including observational measures of child interactions with teacher and peers, such as inCLASS (Booren et al., 2012), might provide valuable information to evaluate the effects of SEL programs on child engagement during academic tasks and classroom interactions. Regarding social and emotional outcomes, while this study included validated scales and questionnaires that provide reliable results that allow comparisons with findings from other studies and that were largely proximal to the 4Rs + MTP program's theory of change (e.g., measures of social-cognitive processes associated with aggression), other potential child social–emotional outcomes relevant to the program's goals were not included in the current study (e.g., child empathy).

Finally, domains of classroom interaction quality (i.e., emotional support, instructional support, and classroom organization) did not mediate the effects of 4Rs + MTP on any of the child outcomes at above average levels of compliance. This study relied on data collected between winter and spring. Significant mediation effects may take more than the time elapsed during this period or even one full school year to manifest. Treatment differences in emotional support due to high compliance with treatment assignment, might not be sufficient to significantly mediate the effects of treatment on child outcomes when examined in this group of teachers. The possibilities for expanded effects to other domains of classroom interactions and a significant mediation effect of classroom emotional support on child outcomes will be examined in the subsample of third grade teachers who were followed and assessed in the fall and spring of the subsequent school year along with their new class of 3rd grade students.

4.5. Conclusion

Consistent with previous research, findings in this study show that more experienced teachers with low levels of burnout were more likely to comply with high quality implementation of the 4Rs + MTP program (Fixsen et al., 2009; Downer et al., 2009a; Wehby et al., 2012; Domitrovich et al., 2019; Musci et al., 2019; Weiss et al., 2021). These teachers showed better skills in providing emotional support and their students had fewer school absences than students of teachers with similar compliance propensity in the control group.

While this study only found significant moderation of high compliance propensity on the effect of 4Rs + MTP on child school attendance and a trend on teacher emotional support, it provides a steppingstone toward understanding the extent to which teachers' propensity to comply with implementation influences the effects of SEL programs on classroom and child outcomes. This research provides evidence that implementing SEL programs by teachers with high compliance propensity, may have positive impacts on classroom emotional support, increasing the opportunities for providing

nurturing and caring environments that promote children's development and learning, and increasing school attendance. This is consistent with a developmental cascades approach (Masten and Cicchetti, 2010), prioritizing intervention with teachers as part of a developmental system that can then facilitate positive developmental changes at the child level. This study contributes to the understanding of quality of implementation of SEL programs, particularly, a compliance propensity approach highlights the importance of providing the resources teachers need to increase their propensity to implement an SEL program with high quality, which in turn, increases the likelihood of desired program impacts.

The interest in research on implementation has grown in recent years, in part due to the potential answers such research can provide policymakers in determining adequate or minimum levels of implementation needed for a program to be effective (Meyers et al., 2012; Smith and Hutchinson, 2022). Particularly, findings from the current study suggest that the implementation supports teachers receive during program implementation is a key component in securing high levels of compliance.

Post hoc analyses of variance of teachers' quality of implementation in this study suggests that teacher quality of implementation likely varied systematically by coach assignment. Although measures of coach alliance and coaching worth provide valuable information about the quality of the working relationship between teachers and coaches, these measures rely on *de facto* ratings from the perspective of teachers and/or coaches, and are therefore susceptible to temporal (recall) bias. In this regard, research examining quality of implementation might benefit from an interpersonal perspective, such that the relationship between teachers and coaches becomes the focal unit of analysis. Research on implementation might benefit from observed measures of specific dimensions of teacher-coach quality of interaction that can be linked to improvements in teachers' practices in classrooms.

Quality of implementation might be considered a moderator of the effects of program on teacher outcomes, classroom quality of interactions, and child outcomes. As discussed above, coaching is pivotal in helping teachers to develop the social and emotional skills needed to provide emotional support in their classrooms, which in turn might contribute to improving children's academic and SEL outcomes. Path analysis might be a useful alternative for researchers interested in examining multiple mediating mechanisms by which programs affect child outcomes. Such complex analyses may illuminate how programs generate positive changes in teachers social and emotional skills when implemented with high quality, and enable effective and sustainable high quality interactions in classrooms and improvements in child academic and social and emotional functioning.

The compliance propensity approach could be extended to include factors pertaining to program implementation at the school and district level that might influence teachers' propensity for high quality program implementation. The support schools receive from districts to allocate resources needed to implement the program, the school climate and school level of preparedness to embark on structural changes, are factors that have been examined and shown to contribute to SEL program implementation and child outcomes (Kendziora and Osher, 2016; Oberle et al., 2016; Domitrovich et al., 2019). Using a multilevel propensity approach (Li et al., 2013; Leite et al., 2015; Li and Fraser, 2015; Fuentes et al., 2022) could contribute to the understanding of the effects of teacher propensity for high quality program implementation on classroom and child outcomes, while accounting for their transactions with and within broader levels of the context. This dynamic system perspective of program implementation is consistent with the idea of a holistic comprehension of quality of implementation proposed in this study.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the initial study Co-PIs, JB (cjobrown@fordham.edu) and JD (jd2fe@virginia.edu).

Ethics statement

The studies involving human participants were reviewed and approved by the Institutional Review Board (IRB) at Fordham University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Acknowledgments

We thank all of the fantastic children, teachers, and schools that took part in this research. This study was supported by an Institute of Education Sciences, U.S. Department of Education grant (#R305A140559) to Fordham University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1065749/full#supplementary-material

References

Aber, L., Brown, J. L., Jones, S. M., Berg, J., and Torrente, C. (2011). School-based strategies to prevent violence, trauma, and psychopathology: the challenges of going to scale. *Dev. Psychopathol.* 23, 411–421. doi: 10.1017/S0954579411000149

Aber, J. L., Jones, S. M., Brown, J. L., Chaudry, N., and Samples, F. (1998). Resolving conflict creatively: evaluating the developmental effects of a school-based violence prevention program in neighborhood and classroom context. *Dev. Psychopathol.* 10, 187–213. doi: 10.1017/S0954579498001576

Abry, T., Rimm-Kaufman, S. E., Larsen, R. A., and Brewer, A. J. (2013). The influence of fidelity of implementation on teacher–student interaction quality in the context of a randomized controlled trial of the responsive classroom approach. *J. Sch. Psychol.* 51, 437–453. doi: 10.1016/j.jsp.2013.03.001

Ashworth, E., Demkowicz, O., Lendrum, A., and Frearson, K. (2018). Coaching Models of School-Based Prevention and Promotion Programmes: A Qualitative Exploration of UK Teachers' Perceptions. *School Mental Health* 10, 287–300.

Baird, J. R. (1986). Improving learning through enhanced metacognition: a classroom study. *Eur. J. Sci. Educ.* 8, 263–282. doi: 10.1080/0140528860080303

Baker, M. L., Sigmon, J. N., and Nugent, E. M. (2001). *Truancy reduction: Keeping students in school*. Washington, DC: U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.

Barth, R. P. (1984). Reducing nonattendance in elementary schools. *Soc. Work. Educ.* 6, 151–166. doi: 10.1093/cs/6.3.151

Beaujean, A. A., and Beaujean, M. A. A. (2012). *Package 'BaylorEdPsych'. R Package Version 0.5. Waco, TX: Baylor University.[Google Scholar]*.

Berg, J. K., Bradshaw, C. P., Jo, B., and Ialongo, N. S. (2017). Using complier average causal effect estimation to determine the impacts of the good behavior game preventive intervention on teacher implementers. *Adm. Policy Ment. Health Ment. Health Serv. Res.* 44, 558–571. doi: 10.1007/s10488-016-0738-1

Blair, C. (2002). School readiness: integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *Am. Psychol.* 57, 111–127. doi: 10.1037/0003-066X.57.2.111

Booren, L. M., Downer, J. T., and Vitiello, V. E. (2012). Observations of children's interactions with teachers, peers, and tasks across preschool classroom activity settings. *Early Educ. Dev.* 23, 517–538. doi: 10.1080/10409289.2010.548767

Brackett, M. A., Rivers, S. E., Reyes, M. R., and Salovey, P. (2012). Enhancing academic performance and social and emotional competence with the RULER feeling words curriculum. *Learn. Individ. Differ.* 22, 218–224. doi: 10.1016/j.lindif.2010.10.002

Bradshaw, C. P., Pas, E. T., Domitrovich, C. E., Reinke, W. M., Herman, K., and Poduska, J. M. (2009). *Measure of coach and teacher Alliance-coach report*. Baltimore: Johns Hopkins University.

Bronfenbrenner, U., andMorris, P. A. (2010). The ecology of developmental processes, In *Handbook of child psychology: Vol. 1. Theoretical models of human development* eds. W. Damon and R. M. Lerner, 5th Ed. (New York: Wiley), 993–1029.

Brown, J. L., Jones, S. M., LaRusso, M. D., and Aber, J. L. (2010). Improving classroom quality: teacher influences and experimental impacts of the 4Rs program. *J. Educ. Psychol.* 102, 153–167. doi: 10.1037/a001816

Brown, J. L., Lowenstein, A. E., Sutton, E., Carlton, R., and Downer, J. T. (2019). Experimental impacts of the 4Rs+MTP program on teachers' Well-being, classroom interactions, and Children's social-emotional and academic development. In: *Paper presented at the Society for Research on educational effectiveness*, Washington, DC.

Collaborative for Academic, Social, and Emotional Learning. (2017). What is SEL? Collaborative for academic, social, and emotional learning. Available at: http://www.casel.org/what-is-sel/

Conduct Problems Prevention Research Group. (1990). Social competence scale—teacher version. Duke University, Fast Track, Durham, NC. Available at: http://www.fasttrackproject.org/techrept/s/sct/sct.pdf (Accessed 22 December 2009).

Curby, T. W., Rimm-Kaufman, S. E., and Ponitz, C. C. (2009). Teacher–child interactions and children's achievement trajectories across kindergarten and first grade. *J. Educ. Psychol.* 101, 912–925. doi: 10.1037/a0016647

Dahlberg, L. L., Toal, S. B., and Behrens, C. B. (1998). *Measuring violence-related attitudes, beliefs, and behaviors among youths: A compendium of assessment tools*. Atlanta, GA.: Division of Violence Prevention. National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.

Dane, A. V., and Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clin. Psychol. Rev.* 18, 23–45. doi: 10.1016/S0272-7358(97)00043-3

Derzon, J. H., Sale, E., Springer, J. F., and Brounstein, P. (2005). Estimating intervention effectiveness: synthetic projection of field evaluation results. *J. Prim. Prev.* 26, 321–343. doi: 10.1007/s10935-005-5391-5

Dix, K., Keeves, J. P., Slee, P. T., Lawson, M. J., Russell, A., Askell-Williams, H., et al. (2010). KidsMatter primary evaluation: technical report and user guide. Available at: https://research.acer.edu.au/cgi/viewcontent.cgi?article=1024&context=learning_processes (Accessed 26 October 2019)

Dix, K. L., Slee, P. T., Lawson, M. J., and Keeves, J. P. (2012). Implementation quality of whole-school mental health promotion and students' academic performance. *Child Adolesc. Mental Health* 17, 45–51. doi: 10.1111/j.1475-3588.2011.00608.x

Dodge, K. A., Pettit, G. S., McClaskey, C. L., and Brown, J. (1986). Social competence in children. *Monogr. Soc. Res. Child Dev.* 51:i. doi: 10.2307/1165906

Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., et al. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: a conceptual framework. *Adv. School Ment. Health Promot.* 1, 6–28. doi: 10.1080/1754730X.2008.9715730

Domitrovich, C. E., Gest, S. D., Jones, D., Gill, S., and DeRousie, R. M. S. (2010). Implementation quality: lessons learned in the context of the head start REDI trial. *Early Child Res. Q.* 25, 284–298. doi: 10.1016/j.ecresq.2010.04.001

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

Downer, J. T., Kraft-Sayre, M. E., and Pianta, R. C. (2009a). Ongoing, web mediated professional development focused on teacher–child interactions: Early childhood educators' usage rates and self-reported satisfaction. *Early Educ. Dev.* 20, 321–345. doi: 10.1080/10409280802595425

Downer, J. T., LoCasale-Crouch, J., Hamre, B. K., and Pianta, R. C. (2009b). Teacher characteristics associated with responsiveness and exposure to consultation and online professional development resources. *Early Educ. Dev.* 20, 431–455. doi: 10.1080/10409280802688626

Doyle, N. B., Downer, J. T., Brown, J. L., and Lowenstein, A. E. (2022). Understanding high quality teacher-student interactions in high needs elementary schools: an exploration of teacher, student, and relational contributors. *Sch. Ment. Heal.* 14, 997–1010. doi: 10.1007/s12310-022-09519-0

Doyle, N. B., Gomez Varon, J., Downer, J. T., and Brown, J. L. (under review). *Testing the integration of a teacher coaching model and a social-emotional learning and literacy intervention in urban elementary schools*.

Durlak, J. A. (2016). Programme implementation in social and emotional learning: basic issues and research findings. *Camb. J. Educ.* 46, 333–345. doi: 10.1080/0305764X.2016.1142504

Durlak, J. A., and DuPre, E. P. (2008). Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am. J. Community Psychol.* 41, 327–350. doi: 10.1007/s10464-008-9165-0

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

Dusenbury, L., Brannigan, R., Hansen, W. B., Walsh, J., and Falco, M. (2004). Quality of implementation: developing measures crucial to understanding the diffusion of preventive interventions. *Health Educ. Res.* 20, 308–313. doi: 10.1093/her/cyg134

Early, D. M., Maxwell, K. L., Ponder, B. D., and Pan, Y. (2017). Improving teacher-child interactions: a randomized controlled trial of making the Most of classroom interactions and my teaching partner professional development models. *Early Child Res. Q.* 38, 57–70. doi: 10.1016/j.ecresq.2016.08.005

Edwards, J. R., and Lambert, L. S. (2007). Methods for integrating moderation and mediation: a general analytical framework using moderated path analysis. *Psychol. Methods* 12, 1–22. doi: 10.1037/1082-989X.12.1.1

Finch, W. H. (2022). Multivariate analysis of variance for multilevel data: a simulation study comparing methods. *J. Exp. Educ.* 90, 173–190. doi: 10.1080/00220973.2020.1718058

Fixsen, D. L., Blase, K. A., Naoom, S. F., and Wallace, F. (2009). Core implementation components. *Res. Soc. Work. Pract.* 19, 531–540. doi: 10.1177/1049731509335549

Follmann, D. A. (2000). On the effect of treatment among would-be treatment compliers: an analysis of the multiple risk factor intervention trial. *J. Am. Stat. Assoc.* 95, 1101–1109. doi: 10.1080/01621459.2000.10474306

Fuentes, A., Lüdtke, O., and Robitzsch, A. (2022). Causal inference with multilevel data: a comparison of different propensity score weighting approaches. *Multivar. Behav. Res.*, 57, 916–939. doi: 10.1080/00273171.2021.1925521

Grant, S., Hamilton, L. S., Wrabel, S. L., Gomez, C. J., Whitaker, A., Leschitz, J. T., et al., (2017). *Social and emotional learning interventions under the every student succeeds act: Evidence review*. RAND Corporation. Santa Monica, CA.

Hadden, D. S., and Pianta, R. B. (2006). "Clinical consultation with teachers for improved preschool literacy instruction" in *Clinical approaches to emergent literacy intervention*, ed. Justice L. M. (San Diego, CA: Plural Publishing) 99–124.

Hamre, B. K., and Pianta, R. C. (2005). Can instructional and emotional support in the first grade classroom make a difference for children at risk of school failure? *Child Dev.* 76, 949–967. doi: 10.1111/j.1467-8624.2005.00889.x

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

Hamre, B. K., Pianta, R. C., Downer, J. T., DeCoster, J., Mashburn, A. J., Jones, S. M., et al. (2013). Teaching through interactions: testing a developmental framework of teacher effectiveness in over 4,000 classrooms. *Elem. Sch. J.* 113, 461–487. doi: 10.1086/669616

Hothorn, T., Hornik, K., and Zeileis, A. (2006). Unbiased recursive partitioning: a conditional inference framework. *J. Comput. Graph. Stat.* 15, 651–674. doi: 10.1198/106186006X133933

Humphrey, N., Barlow, A., and Lendrum, A. (2018). Quality matters: implementation moderates student outcomes in the PATHS curriculum. *Prev. Sci.* 19, 197–208. doi: 10.1007/s11121-017-0802-4

Humphrey, N., Barlow, A., Wigelsworth, M., Lendrum, A., Pert, K., Joyce, C., et al. (2016). A cluster randomized controlled trial of the promoting alternative thinking strategies (PATHS) curriculum. *J. Sch. Psychol.* 58, 73–89. doi: 10.1016/j.jsp.2016.07.002

Imbens, G. W., and Rubin, D. B. (1997). Bayesian inference for causal effects in randomized experiments with noncompliance. *Ann. Stat.* 25, 305–327. doi: 10.1214/aos/1034276631

Jamshidian, M., Jalal, S., and Jansen, C. (2014). MissMech: an R package for testing homoscedasticity, multivariate normality, and missing completely at random (MCAR). *J. Stat. Softw.* 56, 1–31. doi: 10.18637/jss.v056.i06

Jones, S., Bailey, R., and Kahn, J. (2019). The science and practice of social and emotional learning: Implications for state policymaking *State Education Standard*, 19, 18–24.

Jones, S. M., Brown, J. L., and Aber, J. (2011). Two-year impacts of a universal school-based social-emotional and literacy intervention: an experiment in translational developmental research. *Child Dev.* 82, 533–554. doi: 10.1111/j.1467-8624.2010.01560.x

Jones, S. M., Brown, J. L., Hoglund, W. L., and Aber, J. L. (2010). A school-randomized clinical trial of an integrated social–emotional learning and literacy intervention: impacts after 1 school year. *J. Consult. Clin. Psychol.* 78, 829–842. doi: 10.1037/a0021383

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., Rosseel, Y., Miller, P., Quick, C., et al. (2019). *semTools: Useful tools for structural equation modeling. R package version 0.5-2. Retrieved from* https://CRAN.R-project.org/package=semTools.

Kamphaus, R. W., and Reynolds, C. R. (1998). *BASC monitor for ADHD*. Circle Pines, MN: AGS Publishing.

Kendziora, K., and Osher, D. (2016). Promoting children's and adolescents' social and emotional development: district adaptations of a theory of action. *J. Clin. Child Adolesc. Psychol.* 45, 797–811. doi: 10.1080/15374416.2016.1197834

Kingston, B., and Wilensky, R. (2018). Building adult social and emotional capacity: a key ingredient for unleashing the power of prevention. *J. Soc. Soc. Work Res.* 9, 783–797. doi: 10.1086/700655

Kleinke, K. (2017). Multiple imputation under violated distributional assumptions: a systematic evaluation of the assumed robustness of predictive mean matching. *J. Educ. Behav. Stat.* 42, 371–404. doi: 10.3102/1076998616687084

Koo, T. K., and Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* 15, 155–163. doi: 10.1016/j.jcm. 2016. 02. 012

Kraft, M. A., Blazar, D., and Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: a meta-analysis of the causal evidence. *Rev. Educ. Res.* 88, 547–588. doi: 10.3102/0034654318759268

Leite, W. L., Jimenez, F., Kaya, Y., Stapleton, L. M., MacInnes, J. W., and Sandbach, R. (2015). An evaluation of weighting methods based on propensity scores to reduce selection bias in multilevel observational studies. *Multivar. Behav. Res.* 50, 265–284. doi: 10.1080/00273171.2014.991018

Lendrum, A., and Humphrey, N. (2012). The importance of studying the implementation of interventions in school settings. *Oxf. Rev. Educ.* 38, 635–652. doi: 10.1080/03054985.2012.734800

Lendrum, A., Humphrey, N., and Greenberg, M. (2016). "Implementing for success in school-based mental health promotion: the role of quality in resolving the tension between fidelity and adaptation" in *Mental health and wellbeing through schools: The way forward*. eds. R. Shute and P. Slee (Hove: Routledge), 53–63.

Li, J., and Fraser, M. W. (2015). Evaluating dosage effects in a social-emotional skills training program for children: an application of generalized propensity scores. *J. Soc. Serv. Res.* 41, 345–364. doi: 10.1080/01488376.2014.994797

Li, F., Zaslavsky, A. M., and Landrum, M. B. (2013). Propensity score weighting with multilevel data. *Stat. Med.* 32, 3373–3387. doi: 10.1002/sim.5786

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *J. Am. Stat. Assoc.* 83, 1198–1202. doi: 10.1080/01621459.1988.10478722

Louppe, G., Wehenkel, L., Sutera, A., and Geurts, P. (2013). "Understanding variable importances in forests of randomized trees" in *Advances in neural information processing systems* eds. C. J. Burges and L. Bottou and M. Welling and Z. Ghahramani and K.Q. Weinberger (Red Hood, NY: Curran Associates), 431–439.

Lovibond, P. F., and Lovibond, S. H. (1995). The structure of negative emotional states: comparison of the depression anxiety stress scales (DASS) with the Beck depression and anxiety inventories. *Behav. Res. Ther.* 33, 335–343. doi: 10.1016/0005-7967(94)00075-U

MacKinnon, D. P., Lockwood, C. M., and Williams, J. (2004). Confidence limits for the indirect effect: distribution of the product and resampling methods. *Multivar. Behav. Res.* 39, 99–128. doi: 10.1207/s15327906mbr3901_4

Maslach, C., Jackson, S. E., Leiter, M. P., Schaufeli, W. B., and Schwab, R. L. (1986). *Maslach burnout inventory* (Vol. *21*, pp. 3463–3464). Palo Alto, CA: Consulting Psychologists Press.

Maslach, C., Schaufeli, W. B., and Leiter, M. P. (2001). Job burnout. *Annu. Rev. Psychol.* 52, 397–422. doi: 10.1146/annurev.psych.52.1.397

Masten, A. S., and Cicchetti, D. (2010). Developmental cascades. *Dev. Psychopathol.* 22, 491–495. doi: 10.1017/S0954579410000222

McCluskey, C. P., Bynum, T. S., and Patchin, J. W. (2004). Reducing chronic absenteeism: an assessment of an early truancy initiative. *Crime Delinq.* 50, 214–234. doi: 10.1177/0011128703258942

McCoach, D. B., and Adelson, J. L. (2010). Dealing with dependence (part I): understanding the effects of clustered data. *Gift. Child Q.* 54, 152–155. doi: 10.1177/0016986210363076

Meyers, D. C., Durlak, J. A., and Wandersman, A. (2012). The quality implementation framework: a synthesis of critical steps in the implementation process. *Am. J. Community Psychol.* 50, 462–480. doi: 10.1007/s10464-012-9522-x

Musci, R. J., Pas, E. T., Bettencourt, A. F., Masyn, K. E., Ialongo, N. S., and Bradshaw, C. P. (2019). How do collective student behavior and other classroom contextual factors relate to teachers' implementation of an evidence-based intervention? A multilevel structural equation model. *Dev. Psychopathol.* 31, 1827–1835. doi: 10.1017/S095457941900097X

Muthén, B. (2004). "Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data" in *Handbook of quantitative methodology for the social sciences,* ed. D. Kaplan (Newbury Park, CA, US: Sage Publications), 345–368.

Neace, W. P., and Munoz, M. A. (2012). Pushing the boundaries of education: evaluating the impact of second step®: a violence prevention curriculum with psychosocial and non-cognitive measures. *Child Youth Serv.* 33, 46–69. doi: 10.1080/0145935X.2012.665324

Oberle, E., Domitrovich, C. E., Meyers, D. C., and Weissberg, R. P. (2016). Establishing systemic social and emotional learning approaches in schools: a framework for schoolwide implementation. *Camb. J. Educ.* 46, 277–297. doi: 10.1080/0305764X.2015.1125450

Oberski, D. (2016). "Mixture models: latent profile and latent class analysis" in *Modern statistical methods for HCI Human–Computer Interaction Series*, eds. J. Robertson, M. Kaptein, (Cham: Springer), 275–287. doi: 10.1007/978-3-319-26633-6_12

Orpinas, P., and Frankowski, R. (2001). The aggression scale: a self-report measure of aggressive behavior for young adolescents. *J. Early Adolesc.* 21, 50–67. doi: 10.1177/0272431601021001003

Panayiotou, M., Humphrey, N., and Hennessey, A. (2019). Implementation matters: using complier average causal effect estimation to determine the impact of the promoting alternative thinking strategies (PATHS) curriculum on children's quality of life. *J. Educ. Psychol.* 112, 236–253. doi: 10.1037/edu0000360

Pas, E. T., Bradshaw, C. P., Becker, K. D., Domitrovich, C., Berg, J., Musci, R., et al. (2015). Identifying patterns of coaching to support the implementation of the good behavior game: the role of teacher characteristics. *Sch. Ment. Heal.* 7, 61–73. doi: 10.1007/s12310-015-9145-0

Patti, J., Holzer, A. A., Brackett, M. A., and Stern, R. (2015). Twenty-first-century professional development for educators: a coaching approach grounded in emotional intelligence. *Coaching* 8, 96–119. doi: 10.1080/17521882.2015.1061031

Pianta, R. C., and Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: standardized observation can leverage capacity. *Educ. Res.* 38, 109–119. doi: 10.3102/0013189X09332374

Pianta, R. C., Hamre, B. K., and Mintz, S. L. (2012). *Classroom assessment scoring system (CLASS) upper elementary manual*, Charlottesville, VA: Teachstone.

Pianta, R. C., Mashburn, A. J., Downer, J. T., Hamre, B. K., and Justice, L. (2008). Effects of web-mediated professional development resources on teacher–child interactions in pre-kindergarten classrooms. *Early Child Res. Q.* 23, 431–451. doi: 10.1016/j.ecresq.2008.02.001

Portnow, S., Downer, J. T., and Brown, J. (2018). Reductions in aggressive behavior within the context of a universal, social emotional learning program: classroom-and student-level mechanisms. *J. Sch. Psychol.* 68, 38–52. doi: 10.1016/j.jsp.2017.12.004

Preacher, K. J., and Selig, J. P. (2012). Advantages of Monte Carlo confidence intervals for indirect effects. *Commun. Methods Meas.* 6, 77–98. doi: 10.1080/19312458.2012.679848

Reid, K. (2013). *Managing school attendance: Successful intervention strategies for reducing truancy*, London: Routledge.

Reyes, M. R., Brackett, M. A., Rivers, S. E., White, M., and Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. *J. Educ. Psychol.* 104, 700–712. doi: 10.1037/a0027268

Rimm-Kaufman, S. E., Fan, X., Chiu, Y. J., and You, W. (2007). The contribution of the responsive classroom approach on children's academic achievement: results from a three year longitudinal study. *J. Sch. Psychol.* 45, 401–421. doi: 10.1016/j.jsp.2006.10.003

Rockwood, N. J. (2017). Advancing the formulation and testing of multilevel mediation and moderated mediation models. Doctoral dissertation, The Ohio State University.

Roeser, R. W., Eccles, J. S., and Sameroff, A. J. (2000). School as a context of early adolescents' academic and social-emotional development: a summary of research findings. *Elem. Sch. J.* 100, 443–471. doi: 10.1086/499650

Rosenblatt, J. L., and Elias, M. J. (2008). Dosage effects of a preventive social-emotional learning intervention on achievement loss associated with middle school transition. *J. Prim. Prev.* 29, 535–555. doi: 10.1007/s10935-008-0153-9

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York (NY): J. Wiley & Sons.

Rucinski, C. L., Brown, J. L., and Downer, J. T. (2018). Teacher–child relationships, classroom climate, and children's social-emotional and academic development. *J. Educ. Psychol.* 110, 992–1004. doi: 10.1037/edu0000240

Ryff, C. D., and Keyes, C. L. M. (1995). The structure of psychological well-being revisited. *J. Pers. Soc. Psychol.* 69, 719–727. doi: 10.1037/0022-3514.69.4.719

Sagarin, B. J., West, S. G., Ratnikov, A., Homan, W. K., Ritchie, T. D., and Hansen, E. J. (2014). Treatment noncompliance in randomized experiments: statistical approaches and design issues. *Psychol. Methods* 19, 317–333. doi: 10.1037/met0000013

Schaufeli, W. B., Bakker, A. B., Hoogduin, K., Schaap, C., and Kladler, A. (2001). On the clinical validity of the Maslach Burnout Inventory and the Burnout Measure. *Psychology & health* 16, 565–582.

Shah, A. D., Bartlett, J. W., Carpenter, J., Nicholas, O., and Hemingway, H. (2014). Comparison of random forest and parametric imputation models for imputing missing data using MICE: a CALIBER study. *Am. J. Epidemiol.* 179, 764–774. doi: 10.1093/aje/kwt312

Shrout, P. E., and Fleiss, J. L. (1979). Intraclass correlation; uses in assessing rater reliability. *Psychol. Bull.* 86, 420–428. doi: 10.1037/0033-2909.86.2.420

Slee, P. T., Lawson, M. J., Russell, A., Askell-Williams, H., Dix, K. L., Owens, L., et al. (2009). KidsMatter primary evaluation final report. Melbourne: beyond blue. Available at: https://dspace2.flinders.edu.au/xmlui/bitstream/handle/2328/26832/Slee20KidsMatter20Evaluation20final20report.pdf?sequence=1&isAllowed=y

Smith, A., and Hutchinson, K. (2022). "Planning for implementation: why, who, and how" in *Implementation Science*, eds. F. Rapport, R. Clay-Williams, and J. Braithwait (London, England: Routledge), 130–135.

Strobl, C., Boulesteix, A. L., Kneib, T., Augustin, T., and Zeileis, A. (2008). Conditional variable importance for random forests. *BMC Bioinformatics* 9:307. doi: 10.1186/1471-2105-9-307

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

Teasley, M. L. (2004). Absenteeism and truancy: risk, protection, and best practice implications for school social workers. *Child. Sch.* 26, 117–128. doi: 10.1093/cs/26.2.117

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., and Imai, K. (2014). Mediation: R package for causal mediation analysis. Journal of Statistical Software 59, 1–38. doi: 10.1111/cdev.12864

Tolan, P., Elreda, L. M., Bradshaw, C. P., Downer, J. T., and Ialongo, N. (2020). Randomized trial testing the integration of the good behavior game and MyTeachingPartner™: the moderating role of distress among new teachers on student outcomes. *J. Sch. Psychol.* 78, 75–95. doi: 10.1016/j.jsp.2019.12.002

Watson, D., Clark, L. A., and Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS. *J. Pers. Soc. Psychol.* 54, 1063–1070. doi: 10.1037/0022-3514.54.6.1063

Wehby, J. H., Maggin, D. M., Partin, T. C. M., and Robertson, R. (2012). The impact of working alliance, social validity, and teacher burnout on implementation fidelity of the good behavior game. *Sch. Ment. Heal.* 4, 22–33. doi: 10.1007/s12310-011-9067-4

Weiss, M.J., Bloom, H. S., and Brock, T. (2014). A conceptual framework for studying the sources of variation in program effects *Journal of Policy Analysis and Management* 33, 778–808.

Weiss, S., Muckenthaler, M., Heimlich, U., Kuechler, A., and Kiel, E. (2021). Teaching in inclusive schools. Do the demands of inclusive schools cause stress? *Int. J. Incl. Educ.* 25, 588–604. doi: 10.1080/13603116.2018.1563834

Weissberg, R. P. (2019). Promoting the social and emotional learning of millions of school children. *Perspect. Psychol. Sci.* 14, 65–69. doi: 10.1177/1745691618817756

Welch, B. L. (1947). The generalization of Student's problem when several different population variances are involved. *Biometrika* 34, 28–35. doi: 10.2307/2332510

Zhang, Z. (2016). Multiple imputation with multivariate imputation by chained equation (MICE) package. *Annal Transl Med* 4:30. doi: 10.3978/j.issn.2305-5839.2015.12.63

Zhao, P., Su, X., Ge, T., and Fan, J. (2016). Propensity score and proximity matching using random forest. *Contemp. Clin. Trials* 47, 85–92. doi: 10.1016/j.cct.2015.12.012

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 22 May 2023
DOI 10.3389/fpsyg.2023.1134094

OPEN ACCESS

EDITED BY
Stephanie M. Jones,
Harvard University,
United States

REVIEWED BY
Ana Luiza Colagrossi,
Instituto Sedes Sapientiae,
Brazil
Paola Andrea Guerrero-Rosada,
University of Michigan,
United States

*CORRESPONDENCE
Arturo Harker Roa
a.harker@uniandes.edu.co

SPECIALTY SECTION
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

RECEIVED 29 December 2022
ACCEPTED 07 March 2023
PUBLISHED 22 May 2023

CITATION
Harker Roa A, Córdoba Flechas N, Moya A and Pineros-Leano M (2023) Implementing psychosocial support models in contexts of extreme adversity: lessons from a process evaluation in Colombia.
Front. Psychol. 14:1134094.
doi: 10.3389/fpsyg.2023.1134094

Implementing psychosocial support models in contexts of extreme adversity: lessons from a process evaluation in Colombia

Arturo Harker Roa[1]*, Natalia Córdoba Flechas[2], Andrés Moya[3] and María Pineros-Leano[4]

[1]School of Government and IMAGINA Research Center, Universidad de los Andes, Bogota, Colombia, [2]IMAGINA Research Center, Universidad de los Andes, Bogota, Colombia, [3]Department of Economics, Universidad de los Andes, Bogota, Colombia, [4]School of Social Work, Boston College, Chestnut Hill, MA, United States

Introduction: High quality investments during early childhood allow children to achieve their full potential by setting developmental foundations. However, challenges in the scale-up of evidence-based interventions make across-the-board implementation a non-trivial matter. Moreover, extreme contextual conditions -such as community violence, forced displacement, and poverty- impose a double threat. First, by directly affecting early childhood development (ECD), forced displacement and exposure to violence during early childhood, coupled with deficits in nurturing relationships, can trigger toxic stress, affecting children's mental health and social and emotional learning. Second, contexts of extreme adversity exacerbate common implementation pitfalls in the scale-up of interventions. Recognizing and documenting "what it takes" to successfully implement "what works" can contribute to the expansion and effectiveness of evidence-based programs that promote ECD in these settings. *Semillas de Apego* (SA, onward), a community-based psychosocial support model for caregivers, materialized as a strategy to promote ECD in communities affected by violence and forced displacement.

Methods: This article presents the results of the process evaluation of SA during the 2018–2019 implementation in Tumaco, a violence ridden municipality in the south-west border of Colombia, South America. In this phase, the program reached 714 families, 82% direct victims of violence and 57% were internally displaced. The process evaluation combined qualitative and quantitative methodological approaches to produce evidence of factors that promoted implementation quality.

Results: Findings identified salient components of the program that promoted the program's acceptability, adoption, appropriateness, fidelity and sustainability: a rigorous cultural adaptation; well-structured team selection and training methodologies; and a team support and supervision protocol to provide continuous capacity building and prevent burn-out and other occupational hazards common among professionals in mental health and psychosocial support interventions. The statistical analysis using monitoring data identified key predictors of the dosage delivered (a measure of fidelity). Evidence suggests that initial attendance to the program and observable characteristics -such as educational attainment, violence victimization and employment status- predict a successful compliance (in terms of dosage to benefit from the program).

Discussion: This study provides evidence for the development of structural, organizational, and procedural processes for the adoption, appropriate adaptation,

and high-fidelity delivery of psychosocial support models delivered in territories affected by extreme adversity.

KEYWORDS

early childhood, implementation science, process evaluation, community violence, forced displacement and migration, mental health, child development, psychosocial support

1. Introduction

Public policy efforts focused on early childhood are one of the most cost-effective mechanisms to end transmission of poverty across generations and successfully reduce socioeconomic inequalities globally (Noble, 2021). High quality investments during this critical period -from birth to age 5- allow children to achieve their full potential by setting developmental foundations that determine a successful life trajectory in terms of educational, health and labor outcomes (Bowles et al., 2001; Cawley et al., 2001; Heckman, 2006; Borghans et al., 2008; Black et al., 2017). However, challenges in the scale-up of evidence-based interventions make across-the-board implementation a non-trivial matter (Aboud et al., 2018). Moreover, extreme contextual conditions -such as community violence, forceful displacement and poverty- impose a double threat. First, by directly affecting early childhood development (ECD), forced displacement and exposure to violence during early childhood, coupled with deficits in nurturing relationships, can trigger toxic stress, affect children's mental health and social and emotional learning (Shonkoff and Phillips, 2000; Felitti, 2002; Walker et al., 2007; Evans and Schamberg, 2009; Blair, 2010; Molano et al., 2018). Second, contexts of extreme adversity, with low resources, exacerbate common implementation pitfalls in the scale-up of interventions (Al-Ubaydli et al., 2021). Recognizing and documenting "what it takes" to successfully implement "what works" can contribute to the expansion and effectiveness of evidence-based programs that promote ECD in these settings (Gupta et al., 2021).

It is estimated that, by 2021, a total of 36.6 million children in the world have been afflicted by forceful displacement caused by conflict, violence and other crises (UNHCR, 2022), and 449 million children –i.e., one in six children in the world– were living in a conflict zone (Strømme et al., 2022). In Colombia, more than two hundred thousand children between 0 and 5 years of age have been officially registered as victims of the internal armed conflict, and approximately three hundred thousand children have fled with their families from the Venezuelan social and economic crisis and arrived to the country (R4V, 2022; RUV, 2022). This constitutes a latent mechanism for the transmission of poverty and inequality across generations and should set as a priority the design and implementation of evidence-based programs that effectively promote ECD in conflict-affected, low-resourced settings (Ibáñez and Moya, 2010).

In 2014, *Semillas de Apego* (SA, onward), a community-based psychosocial support model for caregivers, materialized as an intervention to promote ECD in communities affected by violence and forced displacement in Colombia. SA promotes maternal mental health through the improvement of healthy and nurturing child–parent relationships, which ultimately foster ECD and overall well-being. This article presents the results of the process evaluation of SA during the 2018–2019 implementation in Tumaco, a violence ridden municipality in the south-west border of Colombia. This work adds to the scarce literature identifying implementation barriers and enablers of ECD interventions in low-resourced settings such as Brazil (Gonçalves et al., 2019; Buccini et al., 2021) and Turkey (Erdemir, 2022a; Erdemir, 2022b), and contributes evidence from a context characterized by community violence, armed conflict and forced displacement.

This study is conceptually framed by the final stage of the "translational pipeline" model, which collects research that focuses on understanding how to guarantee that an intervention with previously proven effectiveness will work at a larger scale, at similar settings (scale-up) or at somehow different settings (scale-out) (WHO and ExpandNet, 2010). This study focuses on identifying key determinants of "implementation effectiveness" (or implementation success) of an innovative psychosocial support intervention that has a proven "treatment effectiveness" (or intervention success) in a community exposed to recurring violence and forced displacement. Specifically, qualitative and quantitative methods were used in the study to answer the following questions: (1) which practices promoted *implementation success* of SA in Tumaco during 2018 and 2019?; (2) what factors should be considered essential during future scale-up phases to guarantee implementation quality?; and (3) are there any improvement opportunities in the implementation protocol to further promote implementation quality? Qualitative approaches were used to better understand specific dimensions of implementation quality – acceptability, adoption, appropriateness, feasibility, fidelity, penetration and sustainability-. Quantitative methods were used to describe mechanisms behind the level of take-up, adherence and "dosage" received by participants, as measures of fidelity in Proctor et al. (2011).

By providing evidence on factors that promoted five of the abovementioned dimensions of implementation quality, this study informs the construction of structural, organizational, and procedural processes for future adaptation and high-fidelity delivery of psychosocial support models in territories affected by community violence and forced displacement. Promoting ECD and overall wellbeing of children living in similar contexts of extreme adversity should be one of the global key priorities, where the number of forcibly displaced families has more than doubled in the last decades and now includes over 30 million.

2. Materials and methods

2.1. Semillas de Apego: Theory of change and implementation teams

Semillas de Apego (SA) promotes mental health and healthy child-caregiver attachment as a pathway for a proper development among children exposed to community violence and forced displacement. By providing psychosocial support to primary caregivers', the program aims to help children reach their full potential amid toxic stress. SA's theory of change is characterized by three

premises and seven objectives that lead to three short-term outcomes and one long-term outcome (see Figure 1). The three premises are: (i) Adverse Childhood Experiences (ACEs or potentially traumatic events experienced during infancy), such as violence victimization and forced displacement have devastating effects on ECD (Shonkoff and Phillips, 2000; Felitti, 2002; Walker et al., 2007); (ii) a healthy child–parent emotional bond can promote resilience and a proper ECD, even in contexts of extreme adversity (Lieberman et al., 2006; Ippen et al., 2011); and (iii) Exposure to traumatic experiences affects caregiver's mental health and hinders their capacity to provide a secure and healthy attachment (Lieberman and Van Horn, 2011).

Given these premises, the program's curriculum is structured to achieve the following objectives: (1) generate capacity for self-reflection and a non-judgmental curiosity about own and children's internal emotional universe; (2) promote sensory integration and adopt self-regulation and stress management tools; (3) raise awareness on the capability of resilience that caregivers and their children have; (4) give new meaning to past traumatic experiences and restore trust; (5) give a new meaning to childrearing support networks and build nurturing teams; (6) strengthen the relationship between caregivers and children; and (7) Increase the repertoire of context-relevant and culturally appropriate parenting strategies. As proposed by the theory of change, if these objectives are achieved, in the short-term SA should: (I) improve caregiver's mental health, (II) build trust and strengthen childrearing support network and nurturing team, and (III) promote a strong and healthy emotional bond between the child and their caregivers. In turn, these transformations are expected to support the socioemotional and cognitive development of the child in the long run.

Semillas de Apego builds upon the framework and structure of the Child–Parent Psychotherapy (CPP) (Lieberman and Van Horn, 2011) and Building Bridges programs (Reyes and Lieberman, 2010). The CPP is a clinical intervention with a multi-theory framework -including attachment, cognitive behavioral, developmental, psychodynamic, and trauma informed theories, among others- that has proven to be effective in improving the mental health and behavior of children, and strengthening child–parent relationships in households exposed to traumatizing events such as: domestic violence (Lieberman et al., 2006; Ippen et al., 2011), child maltreatment (Cicchetti et al., 1999; Toth et al., 2002, 2015), and caregivers struggling with depression (Cicchetti et al., 1999; Guild et al., 2017). Building Bridges is a non-clinical group-based intervention, also inspired by CPP, that has suggestive evidence on its relevance and scalability, but its impact has not been yet evaluated.

Semillas de Apego consists of 15 multi-caregiver group sessions, delivered once per week. Each session lasts approximately two and a half hours. Each group includes between 12 and 16 participants, all primary caregivers of at least one child 0–5 years old. Maintaining the same structure in all sessions and the same group composition makes the weekly meetings a predictable space for the participants, which generates a sense of safety and promotes trust and closeness among the group members. All the sessions have the following sequence of activities: a welcoming moment (to socialize, share important life events and talk about exercises left to practice at home during the week), a warm-up activity (to connect to emotional and physical state through mindfulness and body practices), a core activity (usually involving self-reflection and arts and crafts) and a closing moment (to share experiences and collect learnings from the session). The main objective of the core activities for each session is broadly described in Appendix A. SA uses a *task shifting* approach, which is a human resource management strategy originally developed to address public health crises in contexts where highly qualified professionals are scarce (Orkin et al., 2021). Specifically, task shifting aims at re-distributing tasks from highly qualified workers to workers with fewer qualifications to make more efficient use of the available human recourses (WHO, 2008). In the case of SA, all group sessions are led by two "facilitators," who are community members, are caregivers themselves, and have participated in a six-week experiential training. Although the facilitators are not required to have prior experience nor

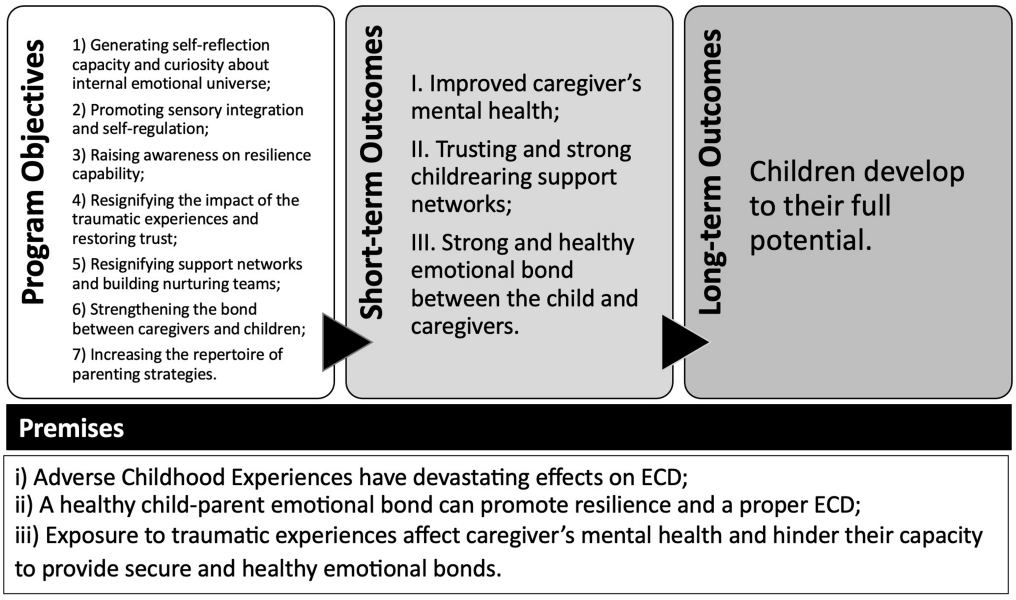


FIGURE 1
Theory of change for *Semillas de Apego*.

formal training in psychosocial interventions, the selection process of SA does favor local community members who have some experience leading group activities. Facilitators have the responsibility of recruiting participants, delivering the program sessions, reporting monitoring data, engaging in reflective supervision, and handling the necessary logistics to successfully conduct the group sessions.

Semillas de Apego adapted the *reflective supervision* model from the one implemented in the CPP. The reflective supervision protocol in SA includes weekly face-to-face group meetings led by a "technical supervisor," and individual meetings or calls between the same supervisor and each facilitator. These meetings are mandatory for all the facilitators. The reflective supervision protocol is an essential component of the program for two reasons. First, the meetings seek to deepen core elements of the intervention by discussing the progress of participants, following-up on at-risk cases, and revising technical and logistical issues. In addition to the monitoring system delivered through a digital platform (described in section 3.3.2), the reflective supervision protocol is a fundamental mechanism to monitor the quality of the sessions. Second, this protocol creates safe spaces to address personal matters that can potentially affect the facilitators and could thus have a negative impact on implementation quality. In addition to providing continuing technical support, the supervision protocol aims to mitigate occupational hazards that are common among professionals working in the mental health field (e.g., vicarious trauma, compassion fatigue, psychosocial distress, and burnout syndrome) (Susman-Stillman et al., 2020; Wilson et al., 2021; Barron et al., 2022).

The final role in the implementation team is the "general technical director," who ensures there is a successful adaptation and implementation of SA to the context where it will be delivered, oversees all technical issues (throughout implementation processes), leads the reflective supervision for the team of supervisors, and intermittently accompanies reflective supervision sessions for the facilitators. Additionally, the director oversees strengthening and managing the network of local institutional allies that support the implementation of SA, and reports to the program's management team about the progress and challenges in the program implementation.

During the 2018–2019 implementation phase, the program had the following support materials: the curriculum, a written document that presented the detailed the objectives each session, scripted all the activities in the session, described the needed arts, crafts and other materials; a training manual, which included the detailed workplan for the training of the facilitators and a summary of the theory supporting the design of the intervention; a technical guide, a written document summarizing theoretical and technical concepts related to the science supporting the program; and a simple guide for basic breath technics and mindfulness practices.

2.2. Context: Tumaco

Tumaco is a 200,000-inhabitant municipality that lies in the Pacific coast of Colombia, in the border with Ecuador. Historically, it has been a setting extremely affected by violence and poverty. In 2017, 243 homicides were reported in Tumaco, representing a rate per 100,000 inhabitants of 116.6. These violence figures exceed those recorded in the most violent cities in the world -such as Los Cabos (111.3), Caracas (111.2), and Acapulco (106.6) (Ortega, 2018)-, and is more than 10 times higher than those observed in many countries amid active civil conflict, —such as Afghanistan (6.5), Iraq (8.0), Somalia (8.0), and South Sudan (13.9) (UNODC, 2019). The fact that this municipality has the largest number of illicit crops in the country (23,148 hectares), half of its population lives with unsatisfied basic needs, 23% of the working-age population are unemployed and 92% of the employed population work in the informal sector, reveals several of the structural determinants of violence in Tumaco (UNODC, 2019; DANE, 2011; UNICEF, 2017).

2.3. Study design

In 2015, the team led a pilot study of SA in Bogotá with 64 participants (divided into three groups), all of them victims of forced displacement. The pilot, funded by the Ministry of Health and Social Protection (MoH) and the Interamerican Development Bank (IADB), was the first implementation of the SA curriculum and included a process evaluation and a results evaluation. This implementation phase was the result of a collaboration with the Mayor's Office of Bogotá that aimed at integrating SA to the service portfolio offered to internally displaced families in the city. The pilot provided evidence that suggested the validity and appropriateness of the SA curriculum (see Harker Roa et al., 2017).

To further strengthen the program in the face of a future scale-up phase by identifying the program's impact and implementation enablers and barriers, between 2018 and 2019 a process evaluation and an impact evaluation (Moya et al., 2022) were carried out simultaneously. The process evaluation, which is the focus of this article, concentrated on seven measures to assess implementation quality: (i) fidelity, (ii) acceptability, (iii) adoption, (iv) appropriateness, (v) feasibility, (vi) penetration, and (vii) sustainability (Proctor et al., 2011; Smith et al., 2018). These constructs are widely used to define implementation outcomes for mental health and behavioral interventions (Proctor et al., 2009; Lewis et al., 2015a,b). To gather information on the process evaluation, qualitative and quantitative data were gathered from different sources (Table 1). All the procedures were approved by the Institutional Review Board of Universidad de los Andes (record #1303, February 2021).

2.3.1. Qualitative approach

Qualitative data was collected through Key Informant Interviews (KIIs) to explore the proposed research questions (see section 1). KIIs were guided by the *Phenomenological Interviewing* framework, aiming at understanding the factors associated with implementation success (i.e., the phenomenon of interest), through the testimonies of actors who had first-hand experiences in the program (Englander, 2012). Semi-structured interview guides were designed for each of the three groups of key informants included in the field work. The first group was the technical team of SA, which included three professionals: the Technical Consultant (affiliated to the University of California San Francisco), the General Technical Director (affiliated to SA) and Technical Supervisors (affiliated to SA, $N = 2$). The second group of key informants were the team of program facilitators ($N = 6$), which was made up by community agents who were affiliated to Genesis Foundation, the NGO implementing the program. Interviews with

TABLE 1 Measures of "implementation quality" and information source.

Construct	Definition	Key Informant			Monitoring System
		Technical Team	Program facilitators	Program Participants	
Acceptability	"The perception among implementation stakeholders that a given treatment, service, practice, or innovation is agreeable, palatable, or satisfactory" (Proctor et al., 2011, p. 67)	No	No	Yes	No
Adoption	"The intention, initial decision, or action to try or employ an innovation or evidence-based practice" (Idem, p. 69)	No	No	Yes	No
Appropriateness	"The perceived fit, relevance, or compatibility of the innovation or evidence-based practice for a given practice setting, provider, or consumer; and/or perceived fit of the innovation to address a particular issue or problem" (Idem, p. 69)	Yes	Yes	Yes	No
Feasibility	"The extent to which a new treatment, or an innovation, can be successfully used or carried out within a given agency or setting" (Idem, p. 69)	Yes	Yes	No	No
Fidelity	"The degree to which an intervention was implemented as it was prescribed in the original protocol or as it was intended by the program developers." (Idem, p. 69)	Yes	Yes	No	Yes
Penetration	"The integration of a practice within a service setting and its subsystems" (Idem, p. 70)	Yes	Yes	No	No
Sustainability	"The extent to which a newly implemented treatment is maintained or institutionalized within a service setting's ongoing, stable operations" (Idem, p. 70)	Yes	No	No	No

Source: Excerpts taken from Proctor et al. (2011).

members of these two groups aimed at collecting information on general implementation challenges, and focused on the appropriateness, feasibility, and potential penetration of SA. Finally, the third group included a sample of primary caregivers ($N = 9$; 8 females, 1 male) that participated in the program. The interviews were designed to collect information about the implementation quality of the program. The interviews also explored the appropriateness, acceptability and adoption of the psychosocial and caregiving strategies discussed throughout the sessions. Each one of the interview guides explored the constructs described in Table 1. Examples of the questions used are presented in Table 2.

All interviews were conducted between December 2020 and February 2021. The interviews were recorded, transcribed by research assistants, organized and systematized using Nvivo Software®. The process of data analysis was done in two phases. Initially, the interviews were grouped together based on the key actor represented in each interview and were analyzed accordingly. Each question asked during the interviews was classified based on the constructs from the interview guide (Table 1). In the second phase, the information that was previously sorted into the different constructs was coded. Three coders were trained to code the data. Before they started coding, a training session on open coding was led by one of the co-authors (NC). Open coding activities were then used to break down the data into emergent categories, which were then used to classify the information within each construct (see Table 3) (Strauss, 1990; Strauss and Corbin, 2002; Nathaniel, 2021).

2.3.2. Quantitative approach

The quantitative research design focuses on leveraging evidence on the fidelity of the implementation, defined as the extent to which SA was executed in Tumaco as it was prescribed in the original protocol (Dusenbury et al., 2003). As explained by Proctor et al. (2011), fidelity is measured "typically by comparing the original evidence based intervention and the disseminated/implemented intervention in terms of (1) adherence to the program protocol, (2) dose or amount of program delivered, and (3) quality of program delivery" (p. 70). Specifically, in this study the quantitative approach provides measures of dosage, but not of adherence -defined as the degree to which the sessions occurred as intended (Hogue et al., 1996)- or the quality of the delivery.

The quantitative information used in this study was collected through a monitoring system designed for SA, which is structured and delivered through the digital platform KoboToolbox®. The system is comprised of custom-made digital forms that collect information provided by the program facilitators, after the execution of each of the 15 sessions in the program's curriculum. Each form includes 4 to 5 open- and close-ended items that inquire about: achievement of goals or milestones in each weekly session and logistics (i.e., availability and quality of spaces, materials, among others). All program facilitators were trained and initially received weekly support to use the digital platform.

At this stage, the monitoring system was able only to consistently collect information on participant attendance. Recurrent structural changes made to the monitoring system during the 2018–2019 implementation phase made it impossible to create indicators of adherence and quality of program delivery, across time and cohorts. Therefore, the quantitative analysis in this process evaluation study focuses on the degree to which SA was implemented as it was prescribed in the original protocol (i.e., fidelity) in terms of the number of sessions delivered to each participant.

TABLE 2 Example questions to explore each construct of implementation quality.

Construct	Questions
Acceptability	How satisfied are you with the information and content of Semillas de Apego? What did you like about the program? What would have you liked to be different [regarding the program]?
Adoption	Can you give us an example of something you learned in the program and how you use it with your family/children? What has motivated you to make it a habit or daily practice?
Appropriateness	What makes the Semillas de Apego program suitable for the reality of Tumaco?
Feasibility	In your opinion, what allows Semillas de Apego to be successfully implemented in Tumaco? What does the program have that makes it possible to be implemented in Tumaco?
Fidelity	In your own words, what does Semillas de Apego aim to achieve? What did the trainings consist of? How often were they [the trainings] carried out? How do you think these trainings help you in your role as facilitator? What trainings help you the most?
Penetration	How does the economic incentive help your participation in the program? How can Semillas the Apego integrate to local or government institutional service settings?
Sustainability	What would you say are the biggest challenges that Semillas de Apego faces today? What would help you do your job better? What are your recommendations for the future expansion of the program?

Questions translated to English from the KII guides.

TABLE 3 Emergent categories.

Construct	Emergent categories
Fidelity	- Strong process of recruitment, selection, and training of facilitators
	- Continuous support of facilitators
Appropriateness	- Flexibility for adaptation to context and culture
Acceptability	- Relative importance of subsidies
Adoption	- Positive parenting
	- New support networks
	- Mindful breathing
Sustainability	- Security
	- Local support and commitment
	- Funding

In addition to descriptive analyses of the proposed dosage measures, a survival analysis (see Allison, 1984) was conducted using participant attendance data to identify the factors associated with the probability of participating in a "sufficient" number sessions - or, alternatively, the probability of not dropping out. To do this, a *time of survival* measure ("t") was created based on the information on the last session attended by each participant before either dropping out or successfully graduating from the program. A participant who continued in SA until session 15 was assigned a value of 15 ($t = 15$), while a participant who did not return to the program after session 6, was assigned a value of 6 ($t = 6$). Afterwards, two different dropout indicators were created using this *time of survival* measure. The first indicator, denominated "observed dropout," is solely based on being present until the end of the program: all individuals who did not attend the 15th session are considered to have dropped out of the program. The second dropout indicator, the "normative dropout" measure, is based on a "minimum dose" that a person should receive in order to expect SA to have a significant impact. Thus, all individuals who did not attend at least 12 sessions are considered to have dropped out of the program. In the design phase, this threshold was defined by the technical team and is formally considered an "expulsion rule" in the set of norms of SA. Moreover, before program initiation, all potential participants are forewarned that if they miss three or more sessions, they will be asked to leave the program. Also, to capitalize the trust-building activities that are at the core of the initial weeks of SA, participants are cautioned that they would also be asked to leave if they miss more than one of the first 3 sessions of the program.

The objective of the survival analysis is to estimate the hazard ratios (HR) across different "participant profiles" using a Cox Proportional Hazards Model (Allison, 2014). To create participant profiles, the data from the monitoring system was merged with data collected in the baseline survey of the impact evaluation study of SA (see Moya et al., 2022), which includes characteristics of the caregiver (i.e., the program participant), her household (family and dwelling) and her child (details in Section 3). The initial raw sample for this analysis included a total of 712 participants, all who were enrolled to the program and participated in the baseline survey. After excluding observations with missing values and outliers, our analytical sample includes 647 observations (90.9% of the total participants enrolled).

3. Results

The results of the process evaluation study aimed to identify information around seven spheres, namely: acceptability, adoption, appropriateness, feasibility, fidelity, penetration, and sustainability (Proctor et al., 2011; Smith et al., 2018). However, the evidence produced by the qualitative and quantitative analyses was not robust enough to support conclusions related to the "feasibility" and "penetration"

dimensions. The themes that arose from the qualitative data analysis were particularly related to the spheres of acceptability, adoption, appropriateness, fidelity and sustainability. The results from the quantitative analysis focused only on measures related to the fidelity dimension. This section presents qualitative and quantitative evidence of factors that enabled the implementation of SA, which could eventually inform the development and expansion of psychosocial support interventions with primary caregivers to promote ECD in contexts impacted by violence and forced displacement.

3.1. Fidelity

Evidence on the fidelity in this phase of implementation (2018–2019) of SA comes from both the quantitative and qualitative analyses. On the one hand, the quantitative approach provides evidence on the dosage (i.e., the number of sessions delivered) and the factors associated with a higher probability of compliance (i.e., participating in enough sessions to perceive the benefits of the intervention). On the other hand, results from the qualitative approach relate to the level of adherence to the protocol and the quality of the delivery. There were two mechanisms to monitor adherence and quality of delivery in the implementation phase this study evaluates. The first one was the monitoring system. The second mechanism was the *reflective supervision* protocol. The qualitative information collected provides evidence on the successful implementation of processes that play a key role to guarantee adherence and quality in the delivery (as explained in section 2): (i) recruitment, selection, and training of program facilitators; and (ii) the continuous support of facilitators.

3.1.1. Dosage measures and trends

Table 4 presents, separately for each of the four cohorts implemented in Tumaco during 2018–2019, the count of persons invited to participate in the program, the count of persons that enrolled, and the count of persons that attended to at least one session – which we define as an "initial take-up." Using information collected in the monitoring system, we constructed three measures of dosage. The first measure is the enrollment rate, which shows the proportion of caregivers that, after receiving a general description of the program and a formal invitation to participate, confirmed their interest to join the 4-month program and participated in a baseline survey. On average, 63% of the persons invited made an initial commitment to participate. This proportion was relatively homogenous across cohorts and oscillated between 60% (cohort 4) and 65% (cohort 3).

The second measure, the take-up rate, is the proportion of the initially committed caregivers that attended to at least one session. For the total 2018–2019 implementation phase in Tumaco, 77% of the persons that enrolled, had an initial take-up of the program. An important result is that this proportion increased significantly, suggesting that important implementation barriers were surpassed from the second cohort onwards. This result is confirmed by the trajectory of the third dosage measure, the average number of sessions attended (after take-up): for the first cohort the average number of sessions was 7.6, while for the second, third and fourth it was 11.9, 11.0, 11.6, and 10.8, respectively. At an ideal level of fidelity (in terms of dosage) -that is, when SA is implemented as it was prescribed in the original protocol-, take-up rates would be 100% and the average number of sessions attended would be 15.

TABLE 4 Take-up and adherence measures, across cohorts.

		Cohort				Total
		1	2	3	4	
[1]	Invited	215	238	362	324	1,139
[2]	Enrolled	132	151	237	193	713
[3]	Initial take-up*	91	117	187	154	549
[4]	Enrollment rate ([2]/[1])	61%	63%	65%	60%	63%
[5]	Take-up rate ([3]/[2])	69%	77%	79%	80%	77%
[6]	Avg. sessions attended (after take-up)⁺	7.6	11.9	11.0	11.6	10.8

Constructed using administrative ledgers of SA. *Attended to at least one session. ⁺Average number of sessions attended by all caregivers that had an initial take-up (i.e., persons who were invited, then enrolled and finally attended to at least one session).

Figure 2 presents the average attendance rate (the third dosage measure), by session and cohort, for the caregivers that had an initial take-up. At least three messages come up from the graph. First, there is a negative time gradient: attendance rates drop as the sessions advance, in all cohorts. Second, while the attendance rate for the first cohort is above 60% for only two sessions, for the subsequent cohorts most of the sessions had attendance rates are above 70%. Third, there is a high volatility in attendance rates, within and across cohorts: maximum attendance rate levels are observed in sessions 2, 3, and 4 (above 90% for cohort 4), and minimum levels are observed in sessions 14 and 15 (around 45%, for cohort 1). To better understand the determinants behind this variation on the level of fidelity in terms of dosage, a survival analysis is presented in the next section.

3.1.2. Determinants of non-compliance

As mentioned before, attendance registries were used to construct two proxy measures of non-compliance: "*observed* dropouts" and "*normative* dropouts." The distribution of these two indicators is used to capture the average level and variation of fidelity -i.e., how close was the implementation to the original protocol's prescription in terms of the amount of program delivered or dosage-, overall and across participant profiles (described by the set of observable characteristics described in Table 5). Participants in the study (98% woman) are extremely vulnerable: only 11% have formal employment, 50% report no monthly labor income, 75% have at most secondary education, 59% report having been forcefully displaced and 84% being direct victims of violence (see Table 5). In addition, 40% of the households do not have access to public water supply or sewage service, 43% are beneficiaries of a conditional cash transfer program, and 28% are mono-parental. It is important to highlight that there is still an important heterogeneity in the severity of all sources of vulnerability, at the participant and the household level. We observe that most of the abovementioned variables have a relatively large coefficient of variation (defined as the relative magnitude of the standard deviation relative to the mean). This is an important result given that it provides an opportunity to explore potential determinants of the amount of program delivered to a participant, which is a proxy of fidelity.

According to the time of survival indicator, of the 647 participants in the analytical sample 25% attended to only the 1st session, 50% reached the 12th session, and 75% reached the 14th session. This pattern is replicated by the two non-compliance indicators: 49% did

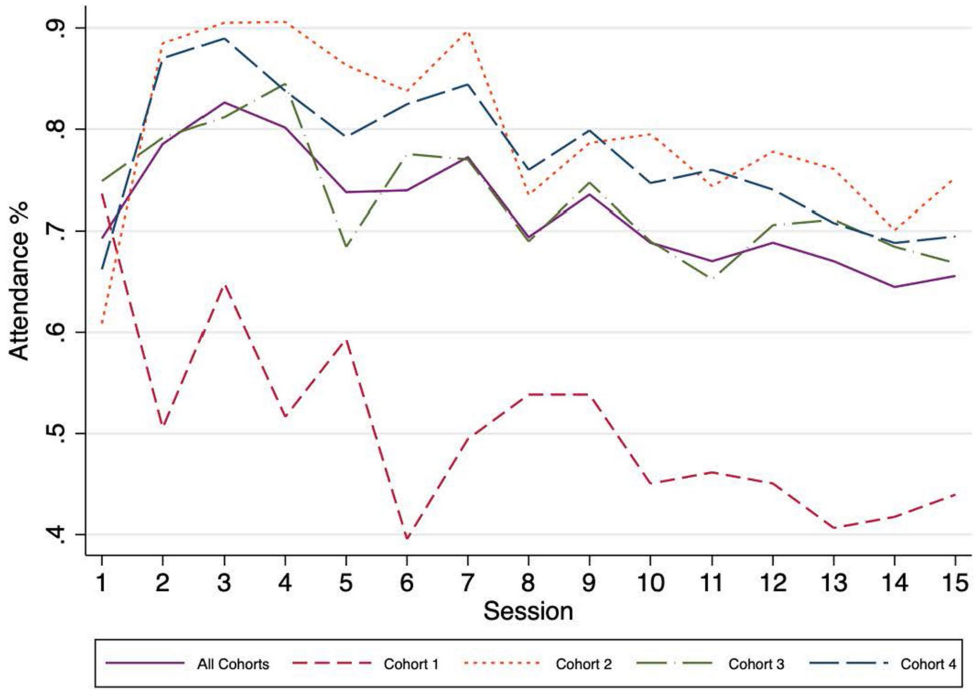


FIGURE 2
Attendance rates after initial take-up, by session and cohort.

not reach the program's 15th session (thus classified as "observed dropouts"), and 54% did not attend at least 12 of the 15 sessions ("normative dropouts").

Table 6 presents the estimated hazard ratios (HRs) for the two proposed Cox Proportional Hazards models. Column 1 presents estimates for the model that predicts "observed dropout" events and Column 2 for the model that predicts "normative dropout" events. In its rows, the table presents the estimated HR across two groups (or profiles) defined by the dichotomous explanatory variable in the row and the standard error for this HR below (in parenthesis). The results show that, after controlling for the full set of covariates included in the models, the following characteristics reduce the probability of not reaching the 15th session (and thus improve the level of fidelity): attending to at least 2 of the first 3 sessions (93.2% lower for this group), female participants (72.4% lower), and having secondary or tertiary education (45.5 and 44.5% lower, respectively). On the contrary, "observed dropout" probability increases if the caregiver is employed in the formal sector (195% higher) and if she is a direct victim of violence (66.1% higher). Similar results -in terms of the sign and at the same level of statistical significance of the HRs- are obtained when the predicted event is the "normative dropout" (see column 2, Table 6).

3.1.3. Strong process of recruitment, selection, and training of facilitators

In the interviews with the technical team, they indicated that the recruitment and selection of facilitators was one of the main steps to ensure an effective training and the successful implementation of SA. The technical team agreed on three key factors to achieve a solid team: (i) doing face-to-face interviews with applicants, (ii) prioritizing the willingness of the candidates to deal with emotional and therapeutic processes and (iii) selecting persons that demonstrate honesty and transparency. One of the members of the technical team mentioned:

> *"We decided to have a selection process. We held a group session where we provided them with a first experience of what the selection process would be like. This taught us a great deal! We received resumes that, given their academic training, made them seem as though they were the right fit. But we ended up choosing more intuitively and not according to the resumes, in favor of people whom we truly felt would be able to deal with the process". BN Technical Team*

According to the facilitators interviewed, the training for the program went beyond the initial training sessions and it implied constant learning. In fact, the training sessions and program reviews were recurrent; they took place once a week and included on-site visits by the team of supervisors, every 15 days. The team's perception of the training is one of constant support and availability from the supervision team. Regarding training, the facilitators mentioned:

> *"I feel that throughout all these years we have been constantly training because we never stop learning, there's always something new. The people who have been in charge have always been very much focused on us as facilitators making use of all the tools so that we leave nothing behind." Facilitator O*

> *"It's wonderful, that we are not alone. We are always accompanied by them, by our supervisor, our coordinator. If something happens*

10.3389/fpsyg.2023.1134094

TABLE 5 Descriptive statistics for the variables included in the survival analysis.

	N	Mean	SD	Min	Max	P25	P50	P75
Dosage measures								
Time of survival (t)	647	8.43	6.05	0	15	1	12	14
Normative dropout (=1)	647	0.54	0.50	0	1	0	1	1
Observed dropout (=1)	647	0.49	0.50	0	1	0	0	1
Attended at least 2 of first 3 (=1)	647	0.64	0.48	0	1	0	1	1
Caregiver characteristics								
SA Cohort (1, 2, 3, or 4)	647	2.74	1.04	1	4	2	3	4
Caregiver's age	647	28.94	9.51	16	78	22	26	33
Caregiver is female (=1)	647	0.98	0.14	0	1	1	1	1
Monthly labor income (USD)	647	199	280	0	1,177	0	0	313
Caregiver is formal worker (=1)	647	0.11	0.31	0	1	0	0	0
Caregiver's education level (1, 2, or 3)	647	1.07	0.64	0	2	1	1	1
Caregiver is child's parent (=1)	647	0.91	0.28	0	1	1	1	1
Severity index > risk threshold (=1)	647	0.21	0.41	0	1	0	0	0
Caregiver is IDP (=1)	647	0.59	0.49	0	1	0	1	1
Caregiver is violence victim (=1)	647	0.84	0.37	0	1	1	1	1
Household characteristics								
Number of children under 5	647	1.32	0.61	1	5	1	1	2
Two parent household (=1)	647	0.72	0.45	0	1	0	1	1
Household asset index	647	−1.64	1.18	−4.90	2.87	−2.42	−1.63	−0.82
Access to water and sanitation (=1)	647	0.60	0.49	0	1	0	1	1
CCT beneficiary (=1)	647	0.43	0.50	0	1	0	0	1
Child characteristics								
Child's age	647	2.55	0.72	1	5	2	2	3

Impact evaluation study baseline survey (Moya et al., 2022) and SA monitoring system. "Cohort" indicates the cohort of the program for each participant. Caregiver's education level: = 0 if the participant has primary education or less, = 1 if the participant has completed or some secondary education, and =2 if the participant has completed or some tertiary education. Severity index above risk threshold: 1 if the severity index is above the risk threshold. Caregiver is an internally displaced person: = 1 if the participant has suffered from internal displacement (= 0, otherwise). Caregiver is victim of direct violence: = 1 if the participant has been a victim of direct violence (= 0, otherwise). Number of children under 5 years: is the count of children under 5 years old who live in the household. Two parent household: = 1 if the household is composed of two parents (= 0, otherwise). Household asset index: Standardized measure of structural wealth based on self-reported ownership of assets.

> *all of a sudden, they are around 24-7 and will always give us the support we need. I think we haven't been left there, abandoned, they always give us their support". Facilitator V*

The facilitators also mentioned that one of the most effective learning experiences was participating in the program itself, as women and as mothers, before facilitating the sessions. In other words, the facilitators agreed that it was important to understand the program based on their own processes of childrearing and life experiences (frequently marked by violence, abuse, and trauma). This led the facilitators and participants in the program to establish relationships of empathy and trust, and to use the experience of the facilitators as real-life examples that can be achieved thanks to the processes proposed in SA. The facilitators referred only to the experiential processes as the most appropriate for achieving good implementation.

> *"For example, in the activities that we carry out during the training, we include abuse, which is what I personally worked on with my son. I can now let go, speak, and recognize what I did with my son, which helps me to help mothers who will be entering the process, so that I can say: 'Look, you can overcome this, you can transform this maltreatment into something positive for your child, in strengthening these relationships.' Why? Because I went through this myself, I did it". Facilitator*

> *"So, during the training, this could allow me to acknowledge my mistakes and where I am failing as a mother; this is a very specific tool; that helps mothers be open to this relationship, and they begin to believe in the facilitator. That is one of the marvelous touches of Semillas, the fact that they can see the facilitator as an average person who this also happens to and who has also gone through what they are experiencing. We do not judge them, nor do we look at them as if to say: 'you are the worst mother'." Facilitator*

> "*Our script is rigorous, but there are parts where you must provide an example from your personal life. We tell them: 'I also beat them,*

TABLE 6 Descriptive statistics for the variables included in the survival analysis.

Predicted variable	Observed dropout (1)	Normative dropout (2)
Caregiver characteristics		
Cohort 2 [=1]	0.743	0.823
	(0.353)	(0.389)
Cohort 3 [=1]	1.654	2.287**
	(0.665)	(0.878)
Cohort 4 [=1]	0.714	0.969
	(0.351)	(0.430)
Attended at least 2 of the first 3 sessions [=1]	0.068***	0.100***
	(0.024)	(0.030)
Caregiver's age	0.948	0.968
	(0.046)	(0.040)
Caregiver's age squared	1.000	1.000
	(0.001)	(0.001)
Caregiver is female [=1]	0.276***	0.258***
	(0.110)	(0.099)
Caregiver's income (log)	0.830	0.831
	(0.129)	(0.121)
Caregiver's income squared (log)	1.026	1.024
	(0.023)	(0.022)
Caregiver is formal worker [=1]	2.951***	1.425
	(1.138)	(0.351)
Secondary Education [=1]	0.545**	0.354***
	(0.130)	(0.136)
Tertiary Education [=1]	0.555**	0.593
	(0.153)	(0.245)
Caregiver is the child's parent [=1]	1.370	1.031
	(0.596)	(0.321)
Severity index above mental health risk threshold [=1]	1.172	1.251
	(0.223)	(0.211)
Internally displaced person [=1]	1.139	1.169
	(0.214)	(0.195)
Victim of direct violence [=1]	1.661**	1.453*
	(0.423)	(0.314)
Household characteristics		
Number of children under 5 years	0.841	0.791*
	(0.109)	(0.095)
Two parent household [=1]	0.999	0.991
	(0.186)	(0.174)
Household asset index	1.122	1.150*
	(0.093)	(0.087)
Access to public water supply or sewage [=1]	0.805	0.903
	(0.163)	(0.169)
CCT beneficiary [=1]	1.230	1.330*
	(0.209)	(0.211)
Child characteristics		
Child's age	1.046	1.060
	(0.109)	(0.102)
Observations	501	501

(Continued)

TABLE 6 (Continued)

Exponentiated coefficients. Robust standard errors in parenthesis. Specification estimated through Cox Proportional-Hazards Model. A formal statistical test for the "proportional hazards" assumption was conducted using Schoenfeld residuals. Given that the evidence that some covariates were time-varying and thus the original model violated the assumption, interaction terms between these time-varying covariates and the time of survival variable were included in the final model. Also, as a robustness check the models were estimated using a rescaled version of the time of survival variable (by adding 1 to each t). *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

> *I also did such and such, I was also violent, I did this…'. So, the mother feels that the person who is speaking is human, that we are connected." Facilitator V*

3.1.4. Continuous support of facilitators

As mentioned before, SA adopted the reflective supervision model which provides a structure for the continuous training process of the facilitating team. Through individual and group supervision activities, facilitators received feedback from the technical team regarding the fidelity of the implementation of the curriculum, in terms of the adherence to the protocol and the quality of the delivery. According to the technical team, this reflective supervision takes place regularly and it considers the relationships between supervisors and professionals, between professionals and caregivers, and finally between caregivers and their children. A member of the technical team summarized the continuous support in the following way:

> "*In their first trip, the primary team began with the pilot project in Bogotá. It then continued its support for the team in this context, which we call reflexive supervision, in which they deal with clinical dilemmas regarding how to apply the model for different circumstances or groups and they are worked on collaboratively." VR Technical Team*

> *"We think that this type of reflection and supervision, which is reflective, creates emotional support spaces that become part of a parallel process. And I think that they are critical in this type of work, because it evokes a lot in a person. One sees a lot of traumas, poverty, injustice. […] We think that you need someone a bit more distanced from the system who can help you think about what you can come and say, 'Oh, I feel this way!'. And you can receive this type of support". VR Technical Team*

As stated by the technical team, *reflective supervision* is conceived to provide *technical support*, but also as a strategy of *psycho-emotional well-being* for professionals who work with families and children in

highly adverse contexts. According to the interviews, this type of supervision enables the team to manage strong emotions and carry out reflexive actions to tailor the implementation practices. The facilitators discussed at great length the benefits of reflective supervision by stating:

> *"Reflective group supervision is about how to understand or ask how the group is doing. And when a particular circumstance, situation or weakness is generated, it must be strengthened through exercises, breathing, workshops, or homework that they assign to us. So, it's wonderful to see how we are growing little by little and strengthen things that we were perhaps not so strong in, even theoretical things." Facilitator M*

> *"Well, at the individual level, we talk about more personal and deeper matters that sometimes affect both the work and the person. So, it is more a matter of getting deeper into personal matters, asking: 'how do we feel? Did it help the session run smoothly? Did it prevent us from carrying out the session […]? What is happening in our lives at that moment?' That kind of things." Facilitator O*

The interviews with the facilitators further revealed that reflective supervision is a model that favors professional well-being and allows the team to feel recognized from the emotional perspective, which helps them not only to deliver the sessions with the expected quality but also to improve their responses to conflict in their family and daily lives.

> *"I have never had this kind of attention towards myself. Do you understand? It was also a matter of picking up the form, going out there, applying and taking it and getting things done. They are concerned about us here, how we are doing, our relationships with our colleagues, bosses, children, family; it's comprehensive. To me this is a beautiful aspect of the program and hopefully it will never change, the end goal, of not just dealing with the community or the mothers, or the people that we'll be working with, the children, early childhood, but also with the facilitator who can also recharge and be nourished by the experiences projected by the community." Facilitator O*

> *"One of the training sessions was very intense. It focused on [past] trauma. The things I carried with me from childhood, and they made me see that the world can be different. And they made me start to change patterns that I had formed in my childhood, negative things, things that are not good for me. And the therapy that they performed on us during the training sessions helped me a great deal. I had fallen into a very similar pattern to my mother's. It was not a positive thing, it was negative. And I really understood that we can change those child-rearing patterns and be different. I do not want to be like my mother, I want to be a different person. And that was one of the things for which I am most grateful to the supervisor and to the program, because I really cut out that pattern and now, I am another person, with different qualities. This left a mark on me and helps me when I am in the field, and I am working with my group. To speak to them based on my own example, the experiences I had and went through to transmit it to them." Facilitator G*

3.2. Appropriateness

Appropriateness refers to the relevance the intervention has among providers and/or participants (Proctor et al., 2011). In this particular study, appropriateness was brought up several times by the facilitators and the technical team and it focused specifically on the flexibility of the curriculum and the subtle, yet meaningful changes that needed to be made to adapt it to the specific conditions of Tumaco.

3.2.1. Flexibility for adaptation to context and culture

Through the interviews with facilitators, it became evident that the SA curriculum allows for flexibility to recognize and integrate the contextual and cultural factors of the community within which the program is being implemented. Although the program's curriculum has core components that should not be modified, the implementation protocol of SA explicitly recognizes that certain components must be adapted to the local context. This adaptation processes involves collaborating with community agents and recognizing their knowledge and understanding the social, economic, cultural, and political realities of the context. It was also necessary to consider different pedagogical strategies in order to teach the content of the manual to the facilitators. In the case of Tumaco, the technical team mentioned that an "experiential learning" methodology was necessary to effectively train the local team of program facilitators. The team arrived at this approach after an initial -and mostly unsuccessful- attempt to do the training based on autonomous reading of the program's manual and additional supporting lectures. Members of the technical team mentioned:

> *"The team (in Tumaco), for example, had not read the curriculum. This was a challenge that we faced in the beginning because there was a cultural difference in the context of Tumaco. […] We decided in our consultation that instead of forcing our way of working, which mostly comes from Bogotá, why don't we learn to work as they do? Maybe it's more organic and maybe it's not a matter of reading. Since they are not going to read what is assigned to them, what if change our methodology to match theirs? What if we go at their pace instead of imposing our own? We should not claim that they are not paying attention or that they don't care, when in reality, something else is happening there." VR Technical Team*

> *"The arrival of Semillas de Apego in Tumaco involved various trial and error exercises through which the technical team reached the conclusion that the team should be trained in an experiential manner. The importance of living the program, of incorporating the tools in the day-to-day lives of the facilitators, was consolidated as one of the criteria for training the local teams on the theoretical contents and group management and including it as a factor to consider in training for adults" BN Technical Team*

Facilitators also agreed that acknowledging cultural differences was crucial for the delivery of SA in Tumaco. Most cultural changes had to do with the integration of colloquial language that people could easily understand. But also, the cultural adaptation implied a consideration on how the traditions of the Colombian Pacific could

be integrated to the curriculum to relate its content and objectives to the culture of the participants. For instance, oral traditions of Afro-Colombians were integrated through children's songs and lullabies that are very present in local child-rearing practices (Meneses Copete, 2022). One of the facilitators summarized this by saying:

> *"The first part was the most difficult. It involved dealing with this new type of work. So, it was about the approach and how to reach this community. Because sometimes it is not just a matter of looking at how the script [in the curriculum] is structured. No, you must speak as they do there, using colloquialisms, words that are spoken in my community and that we are familiar with. Because if you use very technical words in the mother's group, they will not understand you. So, it is better to follow the script but to change that 'particular word.' The way we communicate among ourselves here in our territory and bring it down to that level so that they can understand. Because there are questions in some parts of the script where, if you ask the mothers the way it is written in the script, they will not understand. But we know how to change it, so that it will end up being the same question, but using words spoken here in this environment." Facilitator G*

Also, adaptations to the context included changes in the music and materials used during arts and crafts workshops. The implementation team incorporated traditional instruments and music from Tumaco, which generated engagement and familiarity in the participants. Moreover, in an effort to reduce costs and environmental impacts, the implementation team incorporated materials endemic to the region to substitute the ones that had to be imported from other places.

> *"We have done many things to adapt the program to the context, above all regarding the music that is played during the sessions, which we have been improving. We have been including things that are originally from the Pacific in the topics of the sessions. Things such as materials, or things like that, that can be obtained here, that we have been including, and this has also been helpful". Facilitator O*

3.3. Acceptability

Interviews with participants provided important evidence on the acceptability of SA, in terms of how satisfied the caregivers were in the program, and if they perceived that the curriculum addressed their needs. For instance, caregivers highlighted how the program helped them to increase their capacity to regulate their emotions, understand the developmental needs of their children and increase the ability to interact positively with them. Participants mentioned:

> *"It was a very good program, one that helps moms -not only new moms- to try to control their anger. Because all of us can suddenly become desperate. One has to be honest; it happens to everybody. But to control your breathing and say 'well, just breathe'". Participant A.*

> *"It's not that I'm the most explosive person. But [the program] taught you to control yourself, to calm down when the children are getting frantic, and to correct them with patience and to be a little more tolerant". Participant I.*

3.3.1. Relative importance of cash subsidies

Through our data collected from mothers, fathers, and caregivers, it became clear that although the incentives to attend the sessions were motivating - as participants could buy food and necessary items for their children - their main motivation were the benefits obtained from the program. We believe this is evidence that also supports the acceptability of the program. Program participants frequently mentioned this:

> *"Yes, there was an incentive. They gave me 20,000 pesos [$4 USD] for each session. It [the money] was for the girl or whatever I wanted. […] But for me, the most important thing was not the money, but the learning received from this type of training, right? The most important thing for me was the training." Participant B*

> *"I went mostly to receive the talks, which taught us useful things. But the money was also useful because I went to the supermarket and bought something for the kid or whatever I needed. As a single mother, everything helps me." Participant F*

3.4. Adoption

Through the interviews, participants also shared how they integrated into their lives tools and strategies provided by SA. The adoption of self-care practices and parenting practices strongly suggests not only that the program is relevant for the context, but also that it can be successfully implemented.

3.4.1. Parenting strategies and practices

In particular, frequently caregivers explained how their intention to change their way of interacting with their children is a consequence of them understanding the potential positive impact this has on the behavior, wellbeing, and development of their children.

> *"It is also important to place ourselves in the child's perspective and see how we can educate them in a way that does not affect their development. They are in that stage of development, and they are discovering everything. But we also must instill in them what is good and what is bad. Because why are we going to lie and say 'this is good', when it is bad […]. My girl, she likes the phone a lot, so I told my husband not to lend her the phone so much because every day she asks 'mommy the phone'. She wants to be stuck watching videos of dolls all the time. So, I had to control those things more and try to teach them other types of games". Participant E*

> *"Sitting down to talk and then use the techniques they gave us there [in the program]. We are adults, obviously we are tall... but sitting*

> *or squatting down to'he child's height so that the child looks at us in the same position […] at that moment, we can talk in front of'he child's eyes and tell the child what is wrong and why should correct it. Then she will understand […]". Participant O*

Participating in SA also helped to expand the caregivers' comprehension regarding child development and to prioritize their needs as individuals. A participant mother mentioned:

> *"There were many things that I 'skipped' or did my own way, because I thought that was the way to do it. In the program, they taught us how things should be in raising children. Because they are a reflection of bad parenting. [Before SA] my son, when he ate, would induce vomiting. And they explained to me that this happened if I forced him to eat and suddenly yelled at him, attacked him: 'eat, eat, eat' - it was something that the child did because I attacked him. All of that helped me with my kid. I was suffocating him too much at mealtime." Participant C*

3.4.2. Support networks

Additionally, program participants highlighted, thanks to the program, they were willing to share and create new spaces to support peers, thus creating new networks. Essentially, the participants expressed that they were open to build new relationships in their communities as a consequence of their participation in SA.

> *Well, uh, sharing! Sharing with others, talking to others. Being more of a neighbor to my neighbor, being more attentive to my neighbor, who may be living in a worse situation than I am, and putting into practice what Semillas de Apego taught me" Participant A*

> *"We shared a lot among all those who went to the meetings. But more than anything, we got to know each other." Participant G.*

3.4.3. Mindful breathing

According to the interviews with participants, the activities related to mindfulness were the most memorable aspect of the program. Participants mentioned that mindful breathing exercises were particularly effective for managing stress and promoting relaxation. The participants highly valued practicing mindful breathing and recognized that this skill helped them improve their parenting skills and the communication with their children.

> *"Well, the truth is, one of the things is learning to breathe, right? When you feel, you are going to explode! So, you go - think about it before doing it. You think and breathe, and it calms you down!" Participant G*

> *"The breathing part. Learning to breathe. Not only to breathe when you have problems with the children but also when you have many problems in your head. So, learning to breathe well, to breathe deeply! Participant E*

> *"The way of breathing at the moment when the children drive you crazy. This therapy of how to breathe to calm down." Participant O*

3.5. Sustainability

The process evaluation found that the sustainability (and potential scale-up) of SA faces at least 3 important challenges: (1) guaranteeing the safety of the team, (2) procuring local support and commitment, and (3) ensuring funding to expand and maintain operations.

3.5.1. Security

Tumaco is a highly complex territory where there is an amalgamation of social issues, including the presence of illegal armed groups, illegal drug trafficking, and forced recruitment and displacement. Confrontations between the existing legal and illegal armed groups generate invisible barriers at the territorial level. Therefore, entry into the neighborhoods posed the question if it was really possible to guarantee the safety of the implementation team. A facilitator mentioned:

> *"Another challenge, that is also important, was the fear that I suppose many of us feel when entering certain places where violence is present. Where you didn't know whether it was better to go in or not to go in. But love for the program, as well as our professionalism, helped us to keep moving forward". Facilitator O*

Additionally, the interviewees mentioned that the necessity of developing strong partnerships and communication mechanisms with the communities where the program was going to be implemented to have their buy-in and support, prior to start any work. The facilitators highlighted the importance to engage key community stakeholders such as teachers, educators, government employees or other agents recognized and respected by the community. Facilitators underlined how constant interaction and communication with the community leaders helps to mitigate security risks.

> *"We were going to the children's kindergarten. 'They' [illegal armed-group members] knew that. We went in with our [program-labelled] vests; we brought our ID. They knew that we were there to help, contributing to the community, to the children. In other words, 'They' don't interfere with aspects involving children. 'They' might get involved in other people's affairs, but not the children's: they also take care of the children. 'They' knew that we were not doing anything wrong, we weren't spies of any kind or infiltrators; we were there to help the families, the children." Facilitator H*

3.5.2. Local support and commitment

The evidence gathered in the conversations with the team of facilitators revealed a few key challenges when scaling-out SA in other municipalities or territories. First, the facilitators mentioned establishing partnerships at the institutional level to ensure visibility and a better coordination of the program. A facilitator mentioned:

> *"I think that the most important thing would be to foster connections with the main entity or institution in the municipality. Because we can't arrive at a municipality with a program that no one knows about and simply go in and say: 'okay, I'm going to work here and that's that'. The idea is to foster those connections or partnerships once we arrive, similarly to when we arrive in the neighborhoods through the community action board. So, there we would connect with the Mayor and the [government] entities that exist in the territory. How can we discuss [with these partners about] the program? Obviously by […] asking for their perspectives about the municipality […]; what they think, whether they approve of the project taking place in the municipality." Facilitator O*

3.5.3. Funding

According to the technical team, the future donors of SA must take on a long-term commitment, of at least 3 years. That is the approximate time to reach the territories, adapt the curriculum to the context, and achieve the minimum quality level of training and implementation.

> *"Well, the funding has always been… It's a challenge for any program, right? And that determines many things, because once you are allied with the person that provides funding, well, after that we become responsible for what they want and how they want to measure progress and all of that, so having the flexibility to have funding that includes the level of support that we know works and reflective supervision is essential. These resources are necessary but worthwhile! We know that it has been a challenge to find partners that believe in this work and who want to develop high quality. And that it would be, a kind of commitment. We know that it takes time and that it would be a three-year commitment. At least three years, because normally the first year involves great efforts to connect with key people to start the training. It took pretty much that much time for them to truly internalize this, experience it, make it their own." VR Technical Team*

4. Discussion

SA aims to fill a salient void in the portfolio of ECD services in Colombia, Latin America and other regions with ongoing armed conflicts or persistently high levels of community violence. Given the devastating effects of early exposure to violence, there is an urgency to expand the reach of evidence-based programs that can effectively promote resilience among families rearing children in contexts extreme adversity (Shonkoff, 2010; UNICEF, 2017). Building upon the successful experience of the CPP (Lieberman and Van Horn, 2011) and a thorough tailoring to the local culture, resources, and characteristics of victims of violence in Colombia (Molano et al., 2019), SA constitutes a scalable and sustainable effort to foster early-childhood development and protection in the context of community violence and forced displacement.

The process evaluation of the 2018–2019 phase of SA in Tumaco contributes to future efforts to expand ECD programs by advancing in the understanding of "threats to scalability" and enablers of "implementation success." The evidence provided by this study suggests that the program's curriculum is perceived as relevant and fit to the reality of participants, that caregivers have the intention to adopt the provided self-care and parenting tools, and that the overall level of fidelity is acceptable. Moreover, this study identifies two key implementation enablers that will be extrapolated to future scale-up and scale-out phases of SA, and that can be valuable to other ECD programs that aim at thriving in similar contexts.

The first implementation enabler was the intentional flexibility of the program for its adaptation to the context and culture. It is important to highlight that this adaptability is intentional, and follows a structured process that defines the contents in the curriculum that: (a) are core elements and "should never be adjusted" to the context; (b) "must be contextualized" every time the program lands into a new community, to promote appropriateness, adherence and effectiveness; and (c) "could be adapted," if the implementation team perceives that the adjustment will improve the level of implementation success. This evaluation shows that a thorough and structured cultural adaptation is a strategic practice that promotes implementation success, and probably is also at the core of the program's intervention success.

The second implementation enabler is the integration of a *task shifting* approach. Previous studies have demonstrated that task-shifting is particularly important and valuable in settings with few professionals available (Galvin and Byansi, 2020). Also, recent studies suggest that constraints on the locally available professionals is an important threat to scalability and partly explains differences in the observed impact of similar ECD programs across similar contexts, such as Jamaica, Colombia and Perú (Araujo et al., 2021). In the case of SA, having community agents as program facilitators was initially viewed as compulsory, given the unavailability of qualified professionals. However, the task-shifting approach has proven to be a fundamental enabler and a pilar for the program's implementation success. The findings suggest that training community agents to deliver an evidence-based intervention is effective only when using innovative and appropriate teaching approaches such as experiential learning. Also, that the task-shifting strategy is only viable if there is a well-structured protocol that guarantees the continuous support by trained professionals to facilitators, all throughout the implementation of the program.

The results similarly suggest that task-shifting is not only valuable, but it is also a possibility in a low-resourced setting in Colombia, where there is a huge shortage of mental health professionals. By successfully training and supporting community health workers, the SA aims at building local capacity to develop a sustainable path to scale in places where: psychosocial support services are scarce (or inexistent) and, given the widespread of diverse expressions of violence, there is relatively large population of victimized families and children. Given that they already are part of the community, training lay health workers is particularly important because the intervention fosters an already existing trust and rapport with the participants, which is a highly important factor in conflict-affected areas where it is difficult to trust outsiders. Yet, it is important to highlight that the collected evidence suggests that the task-shifting model requires a robust process of recruitment, selection, and training of facilitators, and a structured mechanism that provides continuous support of program facilitators, such as the reflective supervision protocol.

Despite the encouraging findings, the study found that SA still has at least four areas of improvement. First, the safety of the work teams remains a concern in the implementation and scaling-up of the intervention. Previous studies have indicated that in contexts of violence, safety has consistently come up as a central topic in the ethics committees and has been subject to much supervision by universities and organizations (Sluka, 2020). While Tumaco could be cataloged as a difficult zone for implementation in terms of safety, the results of the implementation assessment suggest that a safety protocol that includes possible risks and solutions, must be developed before teams engage in field work.

Second, in terms of fidelity there is an important "learning curve" between the first time the program is implemented in a territory and the subsequent iterations. Evidence from the 2018–2019 implementation shows the dosage delivered to participants (measured with attendance rates and average total sessions attended) was much lower for the first cohort. This probably is the result of a combination of implementation threats, such as: lack of engagement and trust from the community and key allies (e.g., ECD Centers), inadequate training of the team, insufficiently deep adaptations to the context and culture, among other factors.

Third, the program needs to integrate differentiated strategies to prevent the low-dosage and higher drop-out rates of particularly vulnerable participants. For instance, the results from the statistical analysis show that participants that are less educated and have been direct victims of violence systematically have higher odds of dropping out of the program and not participating in enough sessions to benefit from the program. Improving the compliance of participants and attaining the minimum planned dosage for most (if not all) the participants that have an initial take-up is a direct way of improving the cost-effectiveness of the program.

The fourth area of improvement is related to the strengthening of the program's monitoring system. In the evaluated phase of implementation (2018–2019), SA did not have any direct feedback channels with caregivers (i.e., the participants) to monitor program engagement and perception of the outcome. This is a key issue that should be addressed as a future improvement. An additional, and probably more challenging improvement opportunity for the monitoring system would be to develop an analytical tool to systematize the information exchanged in the reflective supervision sessions.

A recommended strategy to integrate the lessons learned in the first implementation efforts in Tumaco and prepare to expand the program to other communities, is to explicitly develop protocols around the four implementation stages proposed in the "EPIS conceptual model" (Aarons et al., 2011): (E)xploration stage, which includes all strategies for key stakeholder identification and engagement; (P)reparation stage, which includes a collaborative curriculum adaptation process and the procurement of strategic alliances and key input suppliers; (I)mplementation stage, that includes team selection and training processes, curriculum implementation, continuous team support and supervision and program monitoring; and (S)ustainment stage, which focuses on evidence production efforts (e.g., process evaluation), program adjustment and implementing a communication strategy to or extending the support of all key stakeholders.

Future implementation science studies could contribute by advancing designing and testing strategies to minimize the "learning curve" when deploying an evidence-based intervention in a new context. Future studies could also provide evidence on how to increase the cost-effectiveness of the program, which is a non-trivial determinant of scale-up efforts. Finally, subsequent research could integrate more the voices of participants to better understand their experience with the intervention and even have a *participatory action* approach by including caregivers in the study design.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

All study procedures were reviewed and approved by the Institutional Review Board of Universidad de los Andes (record #1303, February 2021). The participants provided their written informed consent to participate in this study.

Author contributions

AHR led the development of the field work and led the design and development of the manuscript. AHR conducted the analysis of the quantitative data. NCF and MP-L conducted the analysis of the qualitative data. AHR, NCF, and MP-L drafted the manuscript. All authors collectively conceived and conceptualized this original research article, contributed to the discussion, reviewed the manuscript, and provided critical feedback.

Funding

We acknowledge generous financial support the Saving Brains partnership (Grant Agreement ID SB-POC-1809-19091), United Way Colombia, Fundación Exito, Fundacion FEMSA, Genesis Foundation, Fundación Coca-Cola, Primero lo Primero, and the Facultad de Typo: Economía at Universidad de los Andes.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Aarons, G. A., Hurlburt, M., and Horwitz, S. M. (2011). Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Admin. Pol. Ment. Health* 38, 4–23. doi: 10.1007/s10488-010-0327-7

Aboud, F. E., Yousafzai, A. K., and Nores, M. (2018). State of the science on implementation research in early child development and future directions. *Ann. N. Y. Acad. Sci.* 1419, 264–271. doi: 10.1111/nyas.13722

Allison, P. D. (1984). "Proportional hazards and partial likelihood" in *Event history analysis* (Thousand Oaks, CA: SAGE Publications, Inc.), 34–42. doi: 10.4135/9781412984195.n4

Allison, PD. (2014). Event history and survival analysis: Regression for longitudinal event data. Available at: https://books.google.com.co/books?id=_DJ1AwAAQBAJ

Al-Ubaydli, O., Lee, M. S., List, J. A., and Suskind, D. (2021). "The science of using science: a framework for understanding the threats to scaling evidence-based policies" in *The scale-up effect in early childhood & public policy.: Why interventions lose impact at scale and what we can do about it*. eds. J. List, D. Suskind and L. Suplee (New York, NY: Routledge)

Araujo, M. C., Rubio-Codina, M., and Schady, N. (2021). "70 to 700 to 70,000: lessons from the Jamaica experiment" in *The scale-up effect in early childhood & public policy: Why interventions lose impact at scale and what we can do about it*. eds. J. List, D. Suskind and L. Suplee (New York, NY: Routledge) doi: 10.4324/9780367822972-15

Barron, C. C., Dayton, C. J., and Goletz, J. L. (2022). From the voices of supervisees: what is reflective supervision and how does it support their work? (part I). *Infant Ment. Health J.* 43, 207–225. doi: 10.1002/imhj.21972

Black, M. M., Walker, S. P., Fernald, L. C. H., Andersen, C. T., DiGirolamo, A. M., Lu, C., et al. (2017). Early childhood development coming of age: science through the life course. *Lancet* 389, 77–90. doi: 10.1016/S0140-6736(16)31389-7

Blair, C. (2010). Stress and the development of self-regulation in context. *Child Dev. Perspect.* 4, 181–188. doi: 10.1111/j.1750-8606.2010.00145.x

Borghans, L., Duckworth, A., Heckman, J., and Ter Weel, B. (2008). The economics and psychology of personality traits. *J. Hum. Resour.* 43, 972–1059. doi: 10.1353/jhr.2008.0017

Bowles, S., Gintis, H., and Osborne, M. (2001). The determinants of earnings: a behavioral approach. *J. Econ. Lit.* 39, 1137–1176. doi: 10.1257/jel.39.4.1137

Buccini, G., Venancio, S. I., and Pérez-Escamilla, R. (2021). Scaling up of Brazil's Criança Feliz early childhood development program: an implementation science analysis. *Ann. N. Y. Acad. Sci.* 1497, 57–73. doi: 10.1111/nyas.14589

Cawley, J., Heckman, J., and Vytlacil, E. (2001). Three observations on wages and measured cognitive ability. *Labour Econ.* 8, 419–442. doi: 10.1016/S0927-5371(01)00039-2

Cicchetti, D., Toth, S. L., and Rogosch, F. A. (1999). The efficacy of toddler-parent psychotherapy to increase attachment security in offspring of depressed mothers. *Attach. Hum. Dev.* 1, 34–66. doi: 10.1080/14616739900134021

Dusenbury, L., Brannigan, R., Falco, M., and Hansen, W. B. (2003). A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Educ. Res.* 18, 237–256. doi: 10.1093/her/18.2.237

Englander, M. (2012). The interview: data collection in descriptive phenomenological human scientific research. *J. Phenomenol. Psychol.* 43, 13–35. doi: 10.1163/156916212X632943

Erdemir, E. (2022a). Home-based early education for refugee and local children via mothers: a model of contextually sensitive early intervention. *J. Child Fam. Stud.* 31, 1121–1144. doi: 10.1007/s10826-021-02197-7

Erdemir, E. (2022b). Uncovering community cultural wealth through an early intervention program: Syrian refugee children speaking. *Early Childhood Educ. J.* 50, 259–278. doi: 10.1007/s10643-020-01140-7

Evans, G., and Schamberg, M. (2009). Childhood poverty, chronic stress, and adult working memory. *Proc. Natl. Acad. Sci.* 106, 6545–6549. doi: 10.1073/pnas.0811910106

Felitti, V. (2002). The relation between childhood experiences and adult health: turning gold into lead. *Perm. J.* 6, 44–47.

Galvin, M., and Byansi, W. (2020). A systematic review of task shifting for mental health in sub-Saharan Africa. *Int. J. Ment. Health* 49, 336–360. doi: 10.1080/00207411.2020.1798720

Gonçalves, T. R., Duku, E., and Janus, M. (2019). Developmental health in the context of an early childhood program in Brazil: the "Primeira Infância Melhor" experience. *Cad. Saude Publica.* 35:e00224317. doi: 10.1590/0102-

Guild, D. J., Toth, S. L., Handley, E. D., Rogosch, F. A., and Cicchetti, D. (2017). Attachment security mediates the longitudinal association between child-parent psychotherapy and peer relations for toddlers of depressed mothers. *Dev. Psychopathol.* 29, 587–600. doi: 10.1017/S0954579417000207

Gupta, S., Supplee, L. H., Suskind, D., and List, J. A. (2021). "Failed to scale: embracing the challenge of scaling in early childhood" in *The scale-up effect in early childhood & public policy. Why interventions lose impact at scale and what we can do about it*. eds. J. List, D. Suskind and L. Suplee (New York, NY: Routledge) doi: 10.4324/9780367822972-1

Harker Roa, A, Moya, A, and Camperos, LM. (2017). Semillas de Apego: Un Programa para la Protección de la Primera Infancia Víctima del Desplazamiento Forzado y la Violencia en Colombia. Informe de Final Preparado para el Banco Interamericano de Desarrollo - BID. Dirección de Políticas Públicas para la Prevención y Reducción de la Criminalidad en Áreas Urbanas de América Latina y el Caribe. Unpublished Document.

Heckman, J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science* 312, 1900–1902. doi: 10.1126/science.1128898

Hogue, A., Liddle, H. A., and Rowe, C. (1996). Treatment adherence process research in family therapy: a rationale and some practical guidelines. *Psychotherapy* 33, 332–345. doi: 10.1037/0033-3204.33.2.332

Ibáñez, A. M., and Moya, A. (2010). Vulnerability of victims of civil conflicts: empirical evidence for the displaced population in Colombia. *World Dev.* 38, 647–663. doi: 10.1016/j.worlddev.2009.11.015

Ippen, C. G., Harris, W. H., Van Horn, P., and Lieberman, A. F. (2011). Traumatic and stressful events in early childhood: can treatment help those at highest risk? *Child Abuse Negl.* 35, 504–513. doi: 10.1016/j.chiabu.2011.03.009

Lewis, C. C., Fischer, S., Weiner, B. J., Stanick, C., Kim, M., and Martinez, R. G. (2015a). Outcomes for implementation science: an enhanced systematic review of instruments using evidence-based rating criteria. *Implement. Sci.* 10:155. doi: 10.1186/s13012-015-0342-x

Lewis, C. C., Stanick, C. F., Martinez, R. G., Weiner, B. J., Kim, M., Barwick, M., et al. (2015b). The society for implementation research collaboration instrument review project: a methodology to promote rigorous evaluation. *Implement. Sci.* 2015 10:2. doi: 10.1186/s13012-014-0193-x

Lieberman, A. F., Ippen, C. G., and Van Horn, P. (2006). Child-parent psychotherapy: 6-month follow-up of a randomized controlled trial. *J. Am. Acad. Child Adolesc. Psychiatry* 45, 913–918. doi: 10.1097/01.chi.0000222784.03735.92

Lieberman, AF, and Van Horn, P. (2011). *Psychotherapy with infants and young children: Repairing the effects of stress and trauma on early attachment*, New York: Guilford Press.

Meneses Copete, Y. A. (2022). Alabaos Y currulaos: resignificación de prácticas identitarias Y culturales afropacíficas en El Marco Del Conflicto Armado. *Investig. Desarro.* 30, 374–407. doi: 10.14482/INDES.30.1.305.8

Molano, A., Harker Roa, A., and Cristancho, J. C. (2018). Effects of indirect exposure to homicide events on Children's mental health: evidence from urban settings in Colombia. *J. Youth Adolesc.* 47, 2060–2072. doi: 10.1007/s10964-018-0876-8

Molano, A, and Harker Roa, A., Moya, A, Camperos, LM, Gamboa, LA, Niño, B Y, and Reyes, V. (2019). Evaluación de la calidad en la implementación del programa Semillas de Apego: Análisis documental de la fidelidad al diseño curricular. Universidad de los Andes. Unpublished document.

Moya, A, Lieberman, AF, Harker Roa, A, Niño, B, Sanchez, J, Torres, MJ, et al. (2022). Maternal health and early-childhood development in conflict-affected settings: Experimental evidence from Colombia. Working Paper. Unpublished document.

Nathaniel, A. (2021). From the editor's desk: Classic Grouded Theory: What it is and what it is not. *Grounded Theory Review* 20, 1–7.

Noble, K. G. (2021). "Early childhood: the opportunity to untap human potential" in *The scale-up effect in early childhood & public policy.: Why interventions lose impact at scale and what we can do about it*. eds. J. List, D. Suskind and L. Suplee (Routledge)

Orkin, A. M., Rao, S., Venugopal, J., Kithulegoda, N., Wegier, P., Ritchie, S. D., et al. (2021). Conceptual framework for task shifting and task sharing: an international Delphi study. *Hum. Resour. Health* 19:61. doi: 10.1186/s12960-021-00605-z

Ortega, JA. (2018). Metodología del ranking (2017) de las 50 ciudades más violentas del mundo. Seguridad Justicia y Paz – Consejo Ciudadano para la Seguridad Pública y Justicia Penal A.C. Available at: https://www.oas.org/ext/es/seguridad/red-prevencion-crimen/Recursos/Biblioteca-Digital/ranking-2017-de-las-50-ciudades-mas-violentas-del-mundo

Proctor, E. K., Landsverk, J., Aarons, G., Chambers, D., Glisson, C., and Mittman, B. (2009). Implementation research in mental health services: an emerging science with conceptual, methodological, and training challenges. *Adm. Policy Ment. Health Ment. Health Serv. Res.* 36, 24–34. doi: 10.1007/s10488-008-0197-4

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., et al. (2011). Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm. Policy Ment. Health Ment. Health Serv. Res.* 38, 65–76. doi: 10.1007/s10488-010-0319-7

R4V/Inter-Agency Coordination Platform for Migrants and Refugees from Venezuela. (2022). Regional refugee and migrant response plan (RRMRP). January 2023–December 2024. Available at: https://www.r4v.info/en/rmrp2023-2024

Reyes, Vilma, and Lieberman, Alicia, *Building bridges: Child-parent psychotherapy group model* (2010).

RUV – Registro Único de Víctimas. (2022). Red Nacional de Información, RUV. Unidad Para la Atención y Reparación Integral a las Víctimas. Available at: https://www.unidadvictimas.gov.co/es/registro-unico-de-victimas-ruv/37394

Shonkoff, J. P. (2010). Building a new biodevelopmental framework to guide the future of early childhood policy. *Child Dev.* 81, 357–367. doi: 10.1111/j.1467-8624.2009.01399.x

Shonkoff, J, and Phillips, D. (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: National Academy Press.

Sluka, J. A. (2020). Too dangerous for fieldwork? The challenge of institutional risk-management in primary research on conflict, violence and 'Terrorism'. *Contemp. Soc. Sci.* 15, 241–257. doi: 10.1080/21582041.2018.1498534

Smith, J. A., Baker-Henningham, H., Brentani, A., Mugweni, R., and Walker, S. P. (2018). Implementation of reach up early childhood parenting program: acceptability, appropriateness, and feasibility in Brazil and Zimbabwe. *Ann. N. Y. Acad. Sci.* 1419, 120–140. doi: 10.1111/nyas.13678

Strauss, AL. (1990). *Qualitative analysis for social scientists*. New York: Cambridge University Press.

Strauss, AL, and Corbin, J. (2002). *Bases de la investigación cualitativa: técnicas y procedimientos para desarrollar la teoría fundada* (*1st*). Medellín: Editorial Universidad de Antioquia.

Strømme, A., Fylkesnes, G. K., Denselow, J., Mangan, R., Podieh, P., and Kamøy, K. (2022). "Stop the War on Children: The forgotten ones". *Save the Children International*.

Susman-Stillman, A., Lim, S., Meuwissen, A., and Watson, C. (2020). Reflective supervision/consultation and early childhood professionals' well-being: a qualitative analysis of supervisors' perspectives. *Early Educ. Dev.* 31, 1151–1168. doi: 10.1080/10409289.2020.1793654

Toth, S. L., Maughan, A., Manly, J. T., Spagnola, M., and Cicchetti, D. (2002). The relative efficacy of two interventions in altering maltreated preschool children's representational models: implications for attachment theory. *Dev. Psychopathol.* 14, 877–908. doi: 10.1017/s095457940200411x

Toth, S. L, Sturge-Apple, M. L., Rogosch, F. A., and Cicchetti, D. (2015). Mechanisms of change: Testing how preventative interventions impact psychological and physiological stress functioning in mothers in neglectful families. *Dev. Psychopathol.* 27, 1661–1674. doi: 10.1017/S0954579415001017

UNHCR. (2022). *Global trends forced displacement in 2021*.

UNICEF. (2017). *Violence in early childhood. Regional framework for UNICEF in Latin America and the Caribbean. ECD and child protection sections – UNICEF LACRO*.

UNODC (2019). Global study on homicide. Available at: https://www.unodc.org/unodc/en/data-and-analysis/global-study-on-homicide.html

Walker, S., Wachs, T., Gardner, J. M., Lozoff, B., Wasserman, G., Pollitt, E., et al. (2007). Child development: risk factors for adverse outcomes in developing countries. *Lancet* 369, 145–157. doi: 10.1016/S0140-6736(07)60076-2

WHO. (2008). Task shifting: Rational redistribution of tasks among health workforce teams – global recommendations and guidelines. World Health Organization – PEPFAR – UNAIDS. NLM classification: WC 503.6

WHO and ExpandNet. (2010). Nine steps for developing a scaling-up strategy. World Health Organization. Available at: https://apps.who.int/iris/handle/10665/44432

Wilson, K., Weatherston, D. J., and Hill, S. (2021). "Introduction to reflective supervision: through the lens of culture, diversity, equity, and inclusion" in *Therapeutic cultural routines to build family relationships*. eds. M. L. Lewis and D. J. Weatherston (Cham: Springer).

10.3389/fpsyg.2023.1134094

Appendix

APPENDIX A Main objectives of the modules and sessions in the Semillas de Apego curriculum.

Session	Main objective of core activity
Module 1	Promoting maternal health
Session 1	Recognizing the importance of non-judgmental and accepting relationships, as the base of constructive processes.
Session 2	Recognizing personal strengths and skills to promote self-esteem and group integration.
Session 3	Increasing reflexive capacity to identify the roles assumed in life and better understand how they have impacted family dynamics and parenting.
Session 4	Increasing awareness of how parenting choices and practices impact who their children will be in the future.
Session 5	Reflecting on the possibility of adopting an alternative childrearing pattern, different from the way they were raised.
Session 6	Recognizing the impact of experiences of violence, and reflecting on the present needs and future actions.
Session 7	Recognizing the each person's journey so far and the impact on the upbringing of their sons and daughters.
Module 2	Promoting early childhood development
Session 8	Recognizing the responsibility and capacity to ensure a nurturing environment, protection and safety to their children.
Session 9	Identifying assertive ways of relating with their children, given the stages of child development.
Session 10	Understanding basic principles of child development and how children communicate their needs through behavior.
Session 11	Learning strategies to respond appropriately to children's needs.
Session 12	Increasing skills and confidence when talking to children about adverse life experiences (potentially traumatic).
Module 3	Strengthening social support networks
Session 13	Increasing confidence to engage other adults to create a childrearing team and a support network.
Session 14	Increasing reflective capacity on the impact of changes and transitions on their own life and on their childrens' lives.
Session 15	Increasing reflective capacity on closure and farewells, for themselves and their children.

frontiers | Frontiers in Psychology

TYPE Conceptual Analysis
PUBLISHED 21 July 2023
DOI 10.3389/fpsyg.2023.1059317

OPEN ACCESS

EDITED BY
Celene Domitrovich,
Georgetown University, United States

REVIEWED BY
Sheila L. Macrine,
University of Massachusetts Dartmouth, United States
Valerie Shapiro,
University of California, Berkeley, United States

*CORRESPONDENCE
Emily Anne Meland
✉ emily_meland@gse.harvard.edu

RECEIVED 01 October 2022
ACCEPTED 28 June 2023
PUBLISHED 21 July 2023

CITATION
Meland EA and Brion-Meisels G (2023) Integrity over fidelity: transformational lessons from youth participatory action research to nurture SEL with adolescents.
Front. Psychol. 14:1059317.
doi: 10.3389/fpsyg.2023.1059317

Integrity over fidelity: transformational lessons from youth participatory action research to nurture SEL with adolescents

Emily Anne Meland* and Gretchen Brion-Meisels

Harvard University, Cambridge, MA, United States

Much has been written about social and emotional learning (SEL) and its positive impact on young people's academic and life outcomes, yet most of this research is based in early childhood and elementary settings. SEL programming for adolescents has shown mixed results, with many programs proving to be largely ineffective or even showing slightly negative impacts for some youth. Adherence to scripted SEL curricula, or "fidelity" to the program components, is often seen by young people to be "lame", inauthentic, and condescending, failing to connect to the topics and issues that feel most critical to them in this stage of their development. For all students, and especially for those whose identities have been systematically marginalized or oppressed by the dominant culture, SEL programming that fails to explicitly address these experiences of injustice often feels inauthentic and out of touch for youth. Therefore, effective implementation of SEL for adolescents is likely to require skillful adaptation and responsiveness to the identities, interests, and motivations of students by educators. In this case, effective SEL may look less like fidelity to a specific set of scripts, sessions, or activities, but rather a commitment to the wholeness of a set of core principles, relationships, and opportunities for adolescent exploration and leadership/empowerment, or what we will call *integrity* of implementation. In this paper, we present one promising approach to adolescent social and emotional development – youth participatory action research (YPAR) – and the ways in which studying the YPAR process (in addition to the research topics selected by youth) can provide key insights into the social and emotional learning and development of youth.

KEYWORDS

social emotional learning, youth participatory action research, critical participatory action research, fidelity of implementation, adolescent development

Introduction

In recent decades, there has been a growing consensus about the contributions that high quality social and emotional learning (SEL) can make to young people's positive life outcomes (Durlak et al., 2011; Jones and Kahn, 2017). While SEL programming likely supports both children and adolescents, most of the research on effective SEL interventions is based on work done in early childhood and elementary settings (Domitrovich et al., 2017; Yeager, 2017). In these settings, social and emotional learning often looks like a

decontextualized, or predetermined, set of lessons and activities that have previously demonstrated a positive impact on students' social and emotional skills (Jones S. M. et al., 2017). This preset or "boxed" approach to social and emotional learning has shown mixed results with adolescents; in large-scale studies, many programs seem to be largely ineffective or even show slightly negative impacts for some youth (Ciocanel et al., 2017; Domitrovich et al., 2017; Yeager, 2017). Qualitative data suggests that adherence to scripted SEL curricula, or "fidelity" to the program components, is often seen by young people to be "lame" (Sawchuk, 2021), inauthentic, and condescending, failing to connect to the topics and issues that feel most critical to them in this stage of their development (Yeager, 2017). This is likely related to the unique developmental needs of adolescence – a growing need for autonomy, identity exploration and resolution, and relationships that provide a sense of belonging – which strict fidelity to an SEL program may undermine or fail to address (Roeser et al., 2000; Ciocanel et al., 2017; Domitrovich et al., 2017; Yeager, 2017; Jagers et al., 2019; Owens et al., 2022).

Faced with this set of challenges, understanding SEL implementation beyond *fidelity to a manual* or *set of scripts* – at the level of its essential mechanisms of change – becomes essential (O'Donnell, 2008; Abry et al., 2015; Jones S. M. et al., 2017). A 2011 meta-analysis of over 200 school-based SEL programs showed that implementation is a key moderator of evidence-based SEL program outcomes. With academic and social emotional impacts almost twice as large for programs that were implemented effectively, as compared to those that encountered problems with implementation (Durlak et al., 2011; Durlak, 2016). Yet, even in this meta-analysis, only 57% of the studies monitored implementation at all, and implementation problems encompassed any implementation issues reported by the authors. This points to the challenge of studying implementation, which itself is a multi-dimensional construct around which the field continues to theorize. We must more precisely understand what aspects of implementation are critical to impacting SEL program outcomes (e.g., Dusenbury et al., 2003; Durlak and DuPre, 2008; Berkel et al., 2011; Proctor et al., 2011; Durlak, 2016).

In this paper, we share some insights from youth participatory action research (YPAR), which we believe is an approach to research, youth development, and systems change that promotes social and emotional skill development by providing adolescents with tools, relationships, and collective opportunities to advocate for more just and equitable environments. As a result, YPAR not only impacts the environments in which youth live, but can also provide youth with feelings of autonomy, connection, and agency. While widely studied as an onto-epistemological approach to research, the lessons of youth participatory action research are often overlooked in the field of social and emotional development. There is limited research connecting these two fields of work or helping them to learn from each other. To some extent, this gap may be the result of epistemological differences in how YPAR and SEL researchers tend to design their studies and conceptualize their outcomes. YPAR studies tend to be critically oriented and focus on setting-level outcomes that indicate improvements in educational equity or justice, whereas many SEL studies seek to measure the efficacy of a program through the aggregated individual-level outcomes of students. Despite this, we believe that there are critical lessons that the field of SEL can learn from youth participatory action research.

In this article, we share one important lesson about how social and emotional learning might be better understood if we were to measure *integrity of implementation* over fidelity, drawing from LeMahieu (2011). LeMahieu (2011) cites the need for "less fidelity of implementation (do exactly what they say to do) [and] more integrity of implementation (do what matters most and works best while accommodating local needs and circumstances)" in implementation science. Identifying what matters most to nurturing social emotional development and wellbeing – the true active ingredients or "kernels of SEL" for youth (Jones and Bouffard, 2012; Li and Julian, 2012; Jones S. et al., 2017) – allows us to shift our conceptualization away from fidelity to standardized activities and toward culturally and contextually responsive *integrity*. This is something we can learn by studying the YPAR process, which grounds itself in a set of core onto-epistemological principles, or commitments.

Moving from fidelity to integrity

One challenge in the implementation science literature is that we lack a clear set of operational definitions for the different aspects of implementation that we might hope to measure (Dusenbury et al., 2003; Durlak and DuPre, 2008; Proctor et al., 2011). For example, fidelity has been used interchangeably with terms such as "adherence," "compliance," "integrity," and "faithful representation" (Durlak and DuPre, 2008, p. 329). These terms and how they are operationalized may be interpreted differently by both researchers and practitioners, ranging from perfect adherence to a scripted and sequenced set of activities, to implementation of a program to an acceptable level or target compliance rate, while allowing for some changes or adaptation (Durlak and DuPre, 2008; O'Donnell, 2008; Berkel et al., 2011). We argue that further clarity and distillation of the fidelity concept is necessary. Future conceptualizations must distinguish between fidelity as defined as "the degree to which an intervention was implemented as it was prescribed in the original protocol" (Proctor et al., 2011, p. 69), and integrity, defined as the degree to which an intervention was implemented maintaining its core active ingredients, while authentically and fully integrating the assets and needs of the local community. In our work researching SEL implementation with K-12 teachers, we find that practitioners grapple with this tension in tangible ways. Teachers have expressed worries to us that if they do not follow a set of program scripts word-for-word, they may undermine the efficacy of the research-based program being studied. Yet, in trying to stick so closely to the script and carry out the intervention "with fidelity," these same educators may undercut the authenticity through which they execute the program and overlook opportunities to be responsive to the backgrounds, needs, and interests of their students. In doing so, they actually miss the critical active ingredients of the intervention (e.g., the development of authentic and reciprocal relationships). It is also common for teachers to share that SEL programs do not resonate with and are not responsive to their students. Without being privy to and having a clear understanding of the theoretically important program components, practitioners are not confident as to when they can and cannot deviate from the script, even to make changes that may make the program more effective for their own students. We hypothesize that increased integrity of implementation would positively impact implementation quality, defined as "the processes used to convey program material to

participants "(Berkel et al., 2011, p. 26). Providing teachers with the tools to know when and how to adapt the curriculum to meet the needs and interests of their students and the strengths of their teaching practice, while maintaining integrity of implementation, has the potential to positively impact all components of quality of process, including: "(1) teacher–student interactivity, (2) teacher enthusiasm, (3) teachers' communication of goals and objectives, (4) student engagement, (5) student attentiveness, and (6) students expressing their opinions" (Dusenbury et al., 2005, p. 310). Thus, specifying the theoretically important program components – often called active ingredients, or "kernels of SEL" – becomes critical (Durlak and DuPre, 2008; Berkel et al., 2011; Abry et al., 2015; Durlak, 2016; Jones S. et al., 2017).

This tension has often been framed in the implementation science literature as a tension between fidelity and adaptation, a debate which first challenged the relevance of strict fidelity of implementation to program success (Durlak and DuPre, 2008; Berkel et al., 2011; Durlak, 2016). Indeed, studies have shown that adaptations during implementation can improve the effectiveness of interventions, by potentially "increasing: ownership on the part of community implementers, perceived relevance on the part of participants, and the match between the program and the ecological niche" (Berkel et al., 2011, p. 26). Yet, not all adaptations are associated with program outcomes. Studies thus far have shown that additions to programs, in the context of high fidelity, are associated with improved outcomes, but changes or modifications in the absence of fidelity, are typically not (Berkel et al., 2011). More research is needed to understand "whether and under what conditions adaptation or reinvention might enhance program outcomes, and under what conditions adaptation or reinvention results in a loss of program effectiveness" (Dusenbury et al., 2003, p. 252). We believe that *integrity* of implementation in SEL, or the degree to which an intervention was implemented maintaining its core active ingredients, while authentically and fully integrating the assets and needs of the local community, would allow us to see more clearly inside the "black box" of implementation. It would allow us to understand why and how adaptations can contribute to program effectiveness, and preempt when they would fall short (Durlak, 2016). Indeed, adaptations to increase program efficacy become the expectation, rather than a deviation. With integrity of implementation, it is made clear to those delivering the intervention exactly what matters most for young people's social and emotional development, those core components without which we would not expect to see change. This is also what should be described and measured in evaluations, increasing our understanding of the true mechanisms of change in our interventions. From there, while the program may provide suggested activities, adaptation to the local context is encouraged, expected, and supported through the program design and implementation, rather than seen as a deviation from the program's intent.

The distinction between fidelity and integrity may help us to better understand the research on program adaptation that shows a positive correlation between adaptation and program efficacy (Durlak and DuPre, 2008). For example, decades of research points to the importance of relationships as the active ingredient, or key mechanism, in human development (Li and Julian, 2012; Osher et al., 2018). In studying program implementation, Durlak and DuPre (2008) describe findings from a Mitchell (1983) study of youth mentorship that "found that the types of activities performed during a mentoring program were unrelated to outcomes, perhaps because the quality of the relationship formed between mentor and youth was more important. In mentoring, it may not be what you do but how you do it that counts" (p. 341). In this case, integrity of implementation, specifically with regard to the development of authentic relationships between youth and their mentors, was the most critical lever of change for positive youth development in the mentorship program. When it comes to SEL with adolescents, integrity of implementation may trump fidelity of implementation in promoting positive youth development. While currently, "efforts to empirically validate hypothesized core components are quite rare" (Berkel et al., 2011, p. 25), we argue that this is an essential piece of effective SEL research and implementation.

Social emotional learning

Social emotional learning (SEL) commonly refers to the process through which people acquire skills, attitudes, behaviors, and values essential for success in school, work, and life (Jones and Bouffard, 2012). These skills and competencies can primarily be grouped into three large buckets, or domains: cognitive, social, and emotional (Jones and Kahn, 2017). In addition, SEL is considered to include the development of mindsets, character, values, and identity (Jagers et al., 2019; National Commission on Social, Emotional, and Academic Development, 2019). Definitions and measures of social and emotional skills often vary across programs, sometimes making it challenging to isolate the effects of SEL programs on specific social and emotional skills and competencies (Jones and Doolittle, 2017; Jones et al., 2019). Despite this, there is an extensive body of literature linking SEL programming to individual-level outcomes, such as academic performance, behavior, mental health, and positive youth development (Durlak et al., 2011), and fewer studies linking SEL to teacher, classroom, school, or community-level outcomes (Jones S. M. et al., 2017).

Decades of research in developmental science tells us that social and emotional skills build and become increasingly complex over time, and that more basic social and emotional skills learned in early and middle childhood become the building blocks for more complex social emotional skills and competencies in adolescence and adulthood. For example, children must first learn to identify and name their own and others' emotions before they are able to acquire more complex problem-solving and perspective-taking skills (Jones S. M. et al., 2017). This means that both the targeted skills and the ways in which they are taught must be aligned with a young person's age and stage of development, and that SEL should support young people in meeting the unique demands of their contexts (Jones S. M. et al., 2017). What is often overlooked is the importance of culturally responsive and sustaining SEL (Jagers et al., 2019; Meland et al., 2019), which requires that educators align their curriculum with the cultural strengths and contextual experiences of youth in order to honor and sustain students' diverse cultures and ways of being, while simultaneously disrupting systems of oppression that privilege certain ways of being (e.g., White, middle-class, heteronormative, and neurotypical, etc.) over others. For example, many SEL programs fail to explicitly address structural inequity, rendering themselves inauthentic for youth (Kaler-Jones, 2020). While calls have increased in recent years for SEL that is transformative, fearless, abolitionist, liberatory, and humanizing (Jagers et al., 2019; Simmons, 2021; Camangian and Cariaga, 2022; DeMartino et al., 2022), the systematic translation of these ideas into SEL classroom practice has not

been actualized. This lack of cultural and contextual responsiveness often leaves adolescents with SEL programming that feels out of touch. Unfortunately, as Jones S. M. et al. (2017) point out, "SEL programs and interventions frequently target the same skills in the same ways across multiple years" (p. 52), exacerbating this issue. For adolescents, whose social and emotional skills and their applications are becoming increasingly integrated and complex, the approaches to SEL that were used in their early years no longer meet the developmental moment.

Effective implementation of SEL for adolescents requires skillful adaptation and responsiveness to the experiences, interests, and motivations of students, and must explicitly attend to the stage-salient tasks of fostering identity development, agency, and belonging (Jagers et al., 2019). This type of adaptation can be planned for with the creation of flexible curricula, but requires that educators and youth have control over the specific content of local activities. When conceptualizing high-quality implementation of SEL programming for adolescents in this way, it is especially important to understand conditions of the environment that support or hinder social and emotional development. In other words, each local context and moment brings its own challenges to health and wellness for young people, requiring a nuanced set of social emotional skills to navigate. Often, there are structural factors that create systemic inequity, which contribute to these challenges, as well as interpersonal and internal dynamics. Given this reality, effective SEL may look less like fidelity to a specific set of scripts, sessions, or activities, but rather a commitment to a set of core principles, relationships, and opportunities for adolescent exploration and empowerment. This is what we refer to as *integrity of implementation.*

Youth participatory action research as a means of nurturing adolescent social and emotional development

One promising approach for fostering adolescent social and emotional development, which capitalizes on the opportunities and strengths of this developmental period *and* responds to cultural and contextual factors, is Youth Participatory Action Research (YPAR) (Cammarota and Fine, 2008; Mirra et al., 2016; Ozer, 2017; Fine and Torre, 2019; Jagers et al., 2019; Ozer et al., 2020). We do not believe that YPAR should be instrumentalized as another version of an SEL program, since it is intended as an approach to research and social change. However, understanding the YPAR process can provide key insights into the active ingredients or "kernels" of adolescent social and emotional learning that take place through this type of engagement. We might better understand integrity of implementation in SEL through looking at the YPAR process, because its implementation centers on a set of core commitments rather than a set of predefined activities to accomplish its aims.

YPAR is a form of participatory action research (PAR)[1] in which youth are full participants in the research process and seen as the experts of their own lives and contexts (Caraballo et al., 2017). Youth identify topics of inquiry relevant to their life experiences, in which they may interrogate the structural, interpersonal, and psychological factors influencing their lives, collect data on these topics, and engage in systemic analysis. This work is supported by the presence of trusted adults who are knowledgeable of the research process, and who often facilitate or teach some specific research tools, and partner with youth through democratic participation in this process. Ultimately, YPAR projects seek to create some form of collective action that aims to disrupt systems of inequity and promote positive change in communities (Rodríguez and Brown, 2009; Mirra et al., 2016; Brion-Meisels and Alter, 2018). These action projects may take many forms including public art, presentations, recommendations to school and community leaders, and other forms of advocacy. YPAR harnesses the energy, passion, and potential of this critical developmental period by providing youth the chance to build strong relationships with adults and peers, better understand themselves and their communities, study sociopolitical questions of interest, and take action on issues that affect their lives (Rodríguez and Brown, 2009; Ozer, 2017). Through this process, youth have the opportunity to build and display a myriad of social emotional skills, including communication, collaborative problem-solving, critical thinking, identifying and managing emotions, empathy and perspective-taking, and civic values/participation (Ozer et al., 2020).

Despite sharing goals and outcomes around positive youth development, the fields of SEL and YPAR have not always been in dialogue, in part because of differences in their epistemological and ontological origins. Research on the impact of YPAR on SEL specifically has been limited, but studies of YPAR in school settings have shown that YPAR can promote the development of critical thinking skills (e.g., Kirshner et al., 2011), support sociopolitical development (e.g., Cammarota and Romero, 2011; Zeal and Terry, 2013), increase the diversity and depth of social connections (e.g., Flores, 2007), and increase youth voice in school-based decision making (Mitra, 2008; Kirshner et al., 2011; Chou et al., 2015). Studies of YPAR often utilize mixed-methods approaches to understand the impacts of YPAR on structural and cultural aspects of a setting, and sometimes also study its effects on youth over time. For example, in the largest quasi-experimental study of school-based YPAR, researchers found that adolescents who were randomly assigned to the YPAR class "showed increases in sociopolitical skills, motivation to influence their schools and communities, and participatory behavior" as compared to the control group who took part in a direct service peer mentor project (Ozer and Douglas, 2013, p. 66). Future studies of the YPAR *process* (distinct from the research that the youth and their co-researchers conduct in their communities) may consider measuring changes in youth social and emotional competence to further our understanding of these links.

YPAR projects each engage a unique set of research questions and processes that are responsive to the context and priorities that youth participants identify. Unlike in traditional SEL research, where outcomes are often predetermined based on research-driven priorities, in these projects, the outcomes under investigation are generated by the youth/community, and the set of research activities undertaken

1 YPAR, PAR, and CPAR can all be used to describe an approach to participatory action research (PAR). In this paper, we use YPAR to refer to intergenerational critical participatory action research with youth, highlighting the key role adolescents play in the process, and how this collaboration with youth might impact their social and emotional development. We also reference the CPAR commitments, as we believe that a core principle of YPAR is its critical engagement with issues of power and social justice.

depends on the research questions. Therefore, fidelity to a specific set of activities cannot be fixed, or pre-determined; rather, the YPAR process must be fluid and adaptive to meet the questions and needs of the communities in which the projects are carried out. In this way, YPAR projects vary significantly in their content and chosen outcomes. At the same time, most of these projects share a set of core commitments that guide the YPAR process. These commitments, outlined below, might be seen as "active ingredients" or "kernels" of the work (Li and Julian, 2012; Jones S. et al., 2017) – they help us to identify *why* YPAR projects often nurture adolescents' SEL skills and to understand *how* these projects nurture SEL in culturally and contextually responsive ways. In the remainder of this paper, we outline these core commitments and propose that studying *integrity* of implementation may provide us with important information about the mechanisms through which youth build social and emotional skills in the YPAR process.

A focus on core commitments

Scholars in the field of participatory action research (PAR) have described the core principles and commitments of the work in different ways (see Rodríguez and Brown, 2009; Torre et al., 2012; Brion-Meisels and Alter, 2018; Cammarota et al., 2018). In this paper, we choose to summarize the commitments identified by scholars in three broad buckets, or groups. By describing the commitments in this way, we intend to help SEL scholars draw explicit connections between the commitments of Y/C/PAR and the mechanisms of transformation for youth (see Table 1). In addition to these three buckets/groups of commitments, we add a fourth bucket focused on the commitment to authentic and reciprocal relationships.[2] From this point forward, we will refer to these as the YPAR commitments to highlight our focus on youth development, while recognizing that these commitments are grounded in critically-oriented approaches to participatory action research.

The first three commitments outlined below are grounded in the work of María Torre and her colleagues at the Public Science Project, who describe three epistemological commitments of critical participatory action research, each of which move the knowledge produced "toward a stronger validity" (Torre et al., 2012, p. 179) for those closest to the issue at hand. These commitments are: (1) "reframing the problem through critical theory" (ecological and construct validity), (2) "deep and broad participation" (expert validity), and (3) "action and accountability to social change movements" (impact validity) (p. 180). Grounding these commitments is a set of ontological beliefs, or assumptions, which underpin the intergenerational work. These include: "all people have valuable knowledge about their lives and experience; all people have the ability to develop strong critical analyses; all people have multiple identities and carry important histories, connections, and responsibilities to various communities; all people and institutions are embedded in complex social, cultural, and political systems historically defined by power and privilege; the production of knowledge is not objective, or value-free; social research is most valid using multiple/triangulated methods to help capture interconnected individual, social, institutional and cultural layers; participation is not automatic; and change is an ongoing process" (Torre, 2009). Rather than include each of these commitments separately in our table, we bucket them into groups that help illuminate how they nurture social and emotional development. In everyday practice, SEL is fostered throughout the YPAR process in complex and overlapping ways at both the setting and individual level. Our table over-simplifies this, for the purpose of helping scholars in the SEL field better understand the ways in which YPAR supports social emotional development for adolescents *and* how we might begin to assess integrity of implementation for each commitment.

Understanding the connection between each of the core commitments of YPAR and the central goals of SEL can help us to imagine a framework through which we might understand implementation in more iterative and flexible ways. In other words, if these commitments themselves are active ingredients of YPAR that nurture social and emotional development, then we can measure integrity with respect to these commitments, rather than fidelity to a specific set of activities.

Before sharing our thoughts about the ways in which SEL scholars might learn from the practice of YPAR, we believe it is important to share a bit about our own identities and backgrounds. The current perspectives emerged from our personal journeys as scholars committed to social and emotional development, adolescent agency/voice, and critically-oriented research. Each of us has spent time in K-12 settings as a classroom teacher, and each of us entered the world of academia because we believed that the tools of this world would help us better advocate for educational justice. Over time, we each became increasingly concerned with the ways in which traditional SEL research and practice placed dominant ways of being at the center of "good" social and emotional development – a critique that has been echoed by many others in the field (e.g., Jagers et al., 2019; Simmons, 2021; Camangian and Cariaga, 2022; DeMartino et al., 2022). This concern about SEL research is likely heightened by our positionalities, which provide us with significant social privilege in many contexts. [Author 1] identifies as a White, cisgender, heterosexual female with Italian immigrant ancestry who embodies many dominant social identities. She has benefited from contemporary educational systems that operate through the perpetuation of White supremacy, whether that be through what knowledge is valued, how success is measured, or what is deemed as an appropriate way to be and express in educational settings. As a classroom teacher, she was confronted most directly with the ways in which our U.S. school systems are often not set up to value and support students' diverse backgrounds, experiences, and ways of being, setting her on a journey to learn and unlearn how to create educational spaces in which all children, youth, and adults can thrive. [Author 2] identifies as a queer White, cisgender female whose ancestors were a part of the Jewish diaspora, and who has benefited economically from contemporary educational systems. She, also, is working to unlearn colonial ways of being, which is a challenging process that pushes her to slow down, decenter her own thinking, and recenter embodied ways of knowing.

In the following sections we explore each commitment within the context of a YPAR project, and then define how it might help us think about measuring integrity in SEL with adolescents.

2 A commitment to trusting, equitable, and reciprocal relationships underlies all high quality Y/C/PAR work and is threaded throughout the process. We pull out this commitment as a separate bucket, because of its centrality in SEL.

TABLE 1 Mechanisms of transformation through YPAR commitments.

YPAR commitment	Mechanism of transformation*	Assessing integrity
Content. The YPAR team agrees to the interrogation of real-life, relevant issues identified by youth, at least in part through the lens of the structural factors that promote or inhibit community thriving.	Engagement in content that is culturally and contextually relevant (Ladson-Billings, 1995; Paris, 2012) Critical consciousness development (Freire, 1973; Seider and Graves, 2020) Motivational processes– competence, autonomy, relatedness (Roeser et al., 2000; Eccles and Wigfield, 2002)	Review of research questions for relevance to students' lives and opportunities to interrogate power and/or structural inequity Observing the process through which students arrive at their topic of inquiry Engaging youth in conversation (e.g., focus groups, interviews, photovoice, video reflection)
Process. The YPAR team agrees to deep and democratic participation in which youth expertise is essential and those most impacted by the research are centered in its design.	Reciprocal engagement (Li and Julian, 2012; Osher et al., 2018) Shifting the balance of power toward young people (Sameroff, 2010; Li and Julian, 2012; Osher et al., 2018) Motivational processes – competence, autonomy, relatedness (Roeser et al., 2000; Eccles and Wigfield, 2002) Experiences of Competence and Confidence (Lerner and Lerner, 2013)	Reviewing documentation of participatory processes Observing the YPAR team Engaging youth in conversation (e.g., focus groups, interviews, photovoice, video reflection)
Purpose. The YPAR team agrees to engaging in collective action toward a more socially just community/world.	Deeper learning (Mehta and Fine, 2019) Critical consciousness development (Freire, 1973; Seider and Graves, 2020) Opportunities to display Character, Caring, and Contribution (Lerner and Lerner, 2013) Motivational processes – competence, autonomy, relatedness (Roeser et al., 2000; Eccles and Wigfield, 2002)	Reviewing documentation/ materials created for action Observing the YPAR team carry out their action projects Engaging youth in conversation (e.g., focus groups, interviews, photovoice, video reflection)
Core. The YPAR team agrees to authentic and trusted relationships amongst co-researchers, especially between youth and adult partners.	Caring, authentic, and reciprocal relationships (Valenzuela, 1999; Li and Julian, 2012; Osher et al., 2018) Sense of Connection (Lerner and Lerner, 2013)	Relationship surveys and self-reports Observing the YPAR team Engaging youth in conversation (e.g., focus groups, interviews, photovoice, video reflection)

*Transformation takes place at both the setting and individual-level.

Commitment one: youth driven and contextually relevant content

The first commitment describes the *content* of the project. The focus of inquiry in YPAR projects is on issues identified by youth as impacting their lives, viewed through the lens of critical theory (Torre et al., 2012). Rodríguez and Brown (2009) describe this as "a commitment to research and learning in which the topics of inquiry, the content of learning, and the knowledge produced reflect and address the real life problems, needs, desires, and experiences of youth researchers" (p. 24–25). These issues should be ones that can be interrogated through the lens of the structural factors that promote or inhibit community thriving. As Torre et al. (2012) explain, "Critical inquiry deliberately shifts the gaze from 'what's wrong with that person?' to 'what are the policies, institutions, and social arrangements that help to form and deform, enrich and limit, human development?' and 'how do people resist the weight of injustice in their lives?'" (p. 179).

In our work with youth, upholding this commitment begins with inviting young people to identify issues or opportunities that impact themselves and their communities, which they want to investigate. As adult partners, we serve as facilitators for this conversation, trying to ensure that all voices are heard, and we ask probing questions that help students interrogate the root causes of some of issues they identify. Because YPAR projects aim to investigate the structural factors that contribute to a given issue (in addition to the psychological and interpersonal factors), we often begin our brainstorming process with the visual metaphor of a tree. Youth are asked to brainstorm the ways they see, feel, and hear inequality in schools (the "leaves") and things they think might be underlying or causing the inequality that they see ("the roots"). We then ask them to do the same exercise, but on a tree of liberation. What are examples of assets, strengths, moments of joy, or resistance to inequality that they have experienced in their educational journey (the "leaves")? And what are the deeper structural, cultural, or institutional policies or practices that have supported these moments? Looking at all that they have brainstormed, students then

consider what issues or opportunities they are interested in studying more deeply. Youth researchers might choose a topic for their research that aims to better understand one of the "leaves" and how it is connected to the roots in service of making their school a more equitable and liberatory space; or, they might choose to focus on a root, such that multiple leaves might be impacted.

Here is an example of this work in practice. One group of students at a working-class suburban high school noticed that their classes tended to be segregated, with wealthier and White students concentrated in the advanced placement classes, and less wealthy students and students of color in the regular tracks. In discussion, they called out similar trends in access across a range of school-based opportunities, as well as knowledge of and access to school-based supports (e.g., tutoring, mental health, guidance counseling). These students wanted to better understand which students had knowledge of how to access these opportunities and supports and/or found the supports useful, and why, so that they might propose ways to increase equitable access. This became the focus of inquiry for their project, upholding commitment one.

The tree activity is just one way to get students thinking about the issues they might want to address through YPAR; it is not the only way to honor this commitment. Reading across the literature on YPAR, one can find examples of projects where students have begun by studying social theory, and then extrapolated from the theory to consider their own context. One can find examples of projects where students have begun by talking about what frustrates or upsets them about their local context, and then dug into social theory about those issues. And, one can find examples of projects where a critical incident has propelled youth to action. What is important is that ultimately, the group selects a topic of inquiry that feels meaningful and relevant to the youth researchers and allows them to identify possible avenues to create change. By honoring this commitment, adult partners honor adolescents' desires for autonomy and agency, as well as their naturally salient critical thinking skills (Roeser et al., 2000; Eccles and Wigfield, 2002). Regardless of *how* the students come to choose their topic, the process itself can scaffold mutual understanding and connection among participants by highlighting shared experiences and fostering empathy. It fosters perspective taking by giving young people (and adults!) an opportunity to listen to and learn from each other's experiences. And, it can provide students with critical analytic tools that foster feelings of agency, as the team comes to consensus on what they want to study (or influence) in that context (Jagers et al., 2019).

Measuring integrity with regard to this commitment, rather than fidelity to the specific activities that it might entail, allows researchers to ensure that specific YPAR projects are including the mechanism necessary for social emotional learning to occur, while *also* providing local educators and organizers with the flexibility to design activities that best meet the cultural and contextual needs of their students. We believe that upholding this commitment increases the possibility that students will engage with topics that are culturally and contextually relevant (Ladson-Billings, 1995; Paris, 2012), harness their motivational processes (Eccles and Wigfield, 2002), and support the development of critical consciousness (Freire, 1973; Seider and Graves, 2020). At the setting level, this may represent a shift toward more youth-driven and culturally sustaining pedagogical practice (Paris, 2012; Ladson-Billings, 2014). Measuring this type of integrity of implementation could look like an analysis of the conversations that led to the project topic/question, an audit of the activities that students completed in order to pick their topic/question, or an analysis of how the research questions reflect the lived experiences of youth and their communities. It might also look like interviewing or surveying students after this step of the process, to understand how they viewed the adult partnerships to be upholding this first commitment, and how that commitment to youth- and community-driven content may have impacted their own thinking/behavior/sense of belonging. Ultimately, measuring the integrity of implementation for this commitment requires asking youth, as their perception of and connection to the research question may impact their motivation to engage in the YPAR activities and their sense of agency in creating positive change in their communities.

Commitment two: participatory processes

The second commitment describes the *process* of YPAR projects, which requires deep and democratic participation from multiple stakeholders, with explicit efforts to amplify and center the voices of those most impacted by the issue under investigation (Torre et al., 2012; Brion-Meisels et al., 2020). This is described by Rodríguez and Brown (2009) as, "a commitment to genuinely collaborative methodological and pedagogical processes that validate, incorporate, and build on the knowledge and skills of youth researchers and support critical and creative engagement in research and learning" (p. 27). In this sense, knowledge is co-constructed in such a way that youth expertise is valued and viewed as essential to the validity of the process, and power is shared between youth and adults through democratic decision-making processes. Torre et al. (2012) further articulate that this process requires "co-constructing what questions most need asking; collaborating to develop both theory and method; [and] co-analyzing data" (p. 175).

In our work with youth, honoring this commitment has begun with collaborative decision-making about the issue/opportunity that they would like to investigate. In this early moment, we see our role as adult facilitators, in part, to ensure that all voices are given space and time. Often, we use this moment to talk about the ways in which research has harmed and helped communities in the past. As we teach students about what makes a good research question, and how the language in our research questions will guide our methodology, we unearth additional opportunities for participants' voices to shape our process.

After collecting everyone's ideas about the topic of inquiry, we work with students to refine the language of their research question. Using a protocol to gain consensus, we have students indicate "fist to five," (National Center for Family Philanthropy, n.d.) whether they feel comfortable moving forward with the final language of the research question, or if we need to pause and continue to revise it. We have found that collaborative tools like Google documents, slides, and forms/surveys allow students to contribute their ideas and suggestions both in real time and asynchronously, after some time for reflection. Through this process, students are thinking deeply and critically about what they wish to understand, for whom, and how they will go about collecting data. As a team, we work to ask ourselves difficult questions about participation and voice – and wonder (aloud) about whose voices might be missing. After coming to consensus on a research question, we work with students to align their data-collection methodology to the question. Here, again, is an opportunity for students' perspectives and prior knowledge to inform the shape of our project. For example, the students who wished to better

understand how their peers gained access to opportunities and supports at their high school chose to create a school-wide survey to gather this information because they believed that this method would allow them to represent the most diverse range of voices. In the process of co-designing a study, adult partners and youth researchers have a chance to discuss many of the social and emotional dynamics in their local context, what participation means, and how different people can best access participation. These conversations often raise awareness about structural and interpersonal factors that shape wellbeing, as well as providing students with analytic tools to better understand their local context.

Once a study has been designed, the protocols themselves must be created. In the case of the project described above, this meant co-constructing a survey with youth researchers. Protocol creation and piloting is a time-consuming and arduous process, and different contexts require different levels of scaffolding and support. In our case, time constraints meant that we sometimes put sample questions in front of youth researchers to react to and to revise using language that would be most clear and accessible to their peers; while at other times, we invite students to develop their own survey questions. Every piece of the survey was reviewed and approved by the students through multiple rounds of review and discussion. Students then designed the recruitment strategy and set out to collect their data by encouraging their peers to take the survey and spreading the word through multiple channels (e.g., lunchroom tabling, email, Google classroom, school assemblies, etc.). Once their data had been collected, the students self-organized into groups to analyze various portions of the data. These groups discussed and came to a collective understanding of the key themes and interpretations of their findings. Finally, the students worked together to decide who would present what piece of information in their final presentation to the community. In each step of the analytic process, adult partners worked to provide students with the tools that they might need and to scaffold their ability to learn these tools; but the commitment to participatory processes required that the youth researchers collaboratively controlled the study design, analytic process, and findings. In this sense, the role of the adult partners was largely to continue to raise up questions about democratic participation and decision-making, provide students with models for how they might honor these commitments, and allow students to experiment with building a process that worked for their context.

Since SEL skills related to communication, collaborative problem-solving, decision-making, and planning (Jagers et al., 2019) are critical to this second YPAR commitment, measuring integrity of implementation for this commitment provides a flexible way of measuring whether the critical components that lead to setting-level and individual change/transformation are present. This commitment is supported in the implementation science and community psychology literature as well, which finds that "shared decision-making (community participation, collaboration) enhances implementation" and increases the chances that the program will be sustained over time (Durlak and DuPre, 2008, p. 340). We can imagine that participatory processes could be measured in multiple ways. It is possible to document instances of shared decision-making throughout the process, through observation or participant self-reporting. One might also interview youth or hold focus groups about their experiences with collaboration and collaborative decision-making. In addition to having these decision-making processes in place, we believe that it is equally important that youth feel that the process was truly democratic, and that they feel a sense of agency throughout the process. This can be documented through youth surveys or focus groups and triangulated with data on documented decision-making processes.

Commitment three: purpose through collective action

The third commitment describes one of the central purposes of any YPAR project – to engage in collective action toward more socially just communities and societies. YPAR projects are designed to inform action to improve the lives of marginalized youth and their communities (Rodríguez and Brown, 2009). These actions might look like developing theory, engaging in social policy, "performing data" through arts-based methods, or making evidence available to organizing allies and activists (Torre et al., 2012). Research is conducted to understand and thereby take action against unjust systems that constrain the ability of all youth to thrive and to build up structures and supports that are protective and promotive.

In our work with youth, collective action has taken many forms. Often, youth research teams present their findings alongside actionable recommendations to those in positions of decision-making power in their communities – school leadership, policy makers, parents, and others. Sometimes, this first action leads to other actions, as the adults work with youth researchers to implement some of the recommendations proposed. Other times, collective action has taken the form of public art, photovoice projects, and community awareness campaigns – students have held events for their peers, designed infographics for school leaders, or provided professional development workshops to their teachers. Regardless of what collective action is taken, through the process of collective action, youth must think through the implications of their findings for different sub-groups. They must begin to develop theories about what lies underneath the findings – how routines, policies, structures, and interactions may be shaping different peoples' experiences. And, they must cooperate to imagine how taking action might contribute to positive change.

Collective action requires that adults and youth practice a number of SEL skills, including demonstrating civic awareness and values, perspective taking, communication and consensus-building, planning and organizing, and adapting to shifts or changes in the plan. In the positive youth development literature, this might be framed as youth opportunities to demonstrate *character* by taking action to promote equity and social justice and *caring* for their communities, thereby leading to opportunities for *contribution* (Lerner and Lerner, 2013). Youth are supported to do so through the intergenerational research process in which adults can co-construct and scaffold skill-building.

While collective action is rarely considered explicitly in the implementation science literature, it builds from findings that empowering communities is an effective way to address local challenges, and that participation enhances implementation (Israel et al., 1998; Durlak and DuPre, 2008; Berkel et al., 2011; Proctor et al., 2011). Measurement of integrity for this commitment may look like documenting the action youth choose to pursue, and the process through which they agree upon this action. It might look like observing the conversations that youth researchers and adults have, as they come to consensus about the action steps they choose to take; or, asking youth to reflect on how the process of collective action shifted their ability to act in other settings (if at all). It is also important to

assess, from the youth perspective, whether they believe that their actions can make a difference in their community, even in small ways, and that they are not simply going through the motions. This is important for the development of their civic consciousness (Sherrod et al., 2010), as well as their own wellbeing. Ultimately, much of the students' experience in and perception of the YPAR project rests on the strength of the relationships that are formed throughout the process.

Commitment four: relationships at the core

The fourth and final commitment describes the interpersonal conditions that drive learning throughout the YPAR process. We believe that, if the commitments above are embodied throughout the research process, then strong, trusting, and authentic relationships will form between co-researchers, especially between youth and adults. At the same time, we recognize that relationships are central to each of the commitments listed above, and forming these relationships builds over time and space. Relationships in which there is an ethic of care, reciprocity, scaffolded joint activity, and intentionality toward shifting the balance of power towards youth, provide the foundation for successful YPAR projects. Importantly, these relationships also support the positive social and emotional development of youth (Li and Julian, 2012; Osher et al., 2018). Given the essential role relationships play in positive youth development, we believe it is important to call out a focus on these relationships as the crucial fourth commitment, and a force that drives the other three.

In our work with youth, we prioritize commitment four by creating space for relationship building in each interaction with youth researchers. This might look like a fun icebreaker at the start of a meeting, checking in on what went well and what's been challenging in our weeks, or demonstrating care around aspects of students' lives outside of our project. This is common practice in SEL programming as well, grounded in a vast knowledge base on relationships as a core mechanism of youth social and emotional development (Li and Julian, 2012; Osher et al., 2018). Often, early in a project, we use games to practice collective decision-making, problem-solving, and action; debriefing these games can help scaffold our relationships and communication for future events. As a project moves forward, we are more likely to use check-ins to give students a chance to describe what they need from the community, how they are doing, and what is "up" for them on a particular day. We express interest and provide support for the other priorities in the young people's lives as much as possible, sometimes forgoing the YPAR meeting agenda altogether so that the students might study for upcoming exams or prioritize other pressing commitments. As adults, we participate fully and model vulnerability in these activities; we consider ourselves co-researchers and team members. We work to ask for support when we need it, while carefully balancing our desire to be vulnerable with our desire to center young peoples' voices and needs in the space.

Relationships are not only a critical mechanism of social and emotional development, but interpersonal relationships are also a crucial contextual factor that can enable or inhibit implementation (Lacouture et al., 2015). There is a vast body of literature on how researchers can measure quality relationships (Sabol and Pianta, 2012), but here, measuring integrity to this commitment means ensuring that this active ingredient/driver of transformation is present in YPAR projects, regardless of their specific content, methods, or collective action. Assessing integrity of implementation might look like administering a survey to students and adults regarding the developmental relationships they experienced through the project (e.g., Search Institute Developmental Relationships survey). It might also look like focus groups with students in which they are asked to reflect on the relationships built over the course of the project and how they believe they have impacted their trajectory. Additionally, it might involve asking students how their relationships on the research team have impacted their relationships to others outside the team (if at all).

Contextual factors influencing integrity of implementation

As a final note, in documenting integrity of implementation in a YPAR project, it may be equally important to document the conditions of the environment (structural, interpersonal, political) that promote or constrain the ability for youth to carry out their projects and uphold the commitments described above. For example, youth researchers may encounter political resistance to their proposed collective action at the school or community level, or a lack of time or physical space may make it difficult to engage in fully participatory processes throughout the research project. The paper's second author and others have written about the risks inherent in the "schoolification" of YPAR, which would need to be taken into consideration, should this approach be considered in a school-based setting (Brion-Meisels and Alter, 2018). These tensions include,

> …authenticity around power sharing (Kohfeldt et al., 2011; Rubin et al., 2017); limited time, student, and staff turnover; imbalances of power (Rubin et al., 2017); centralized control over policies affecting the school (Kirshner, 2007; Ozer et al., 2008; Kohfeldt et al., 2011; Ozer and Douglas, 2013; Rubin et al., 2017); and student agency versus the structural constraints of schooling (Ozer and Douglas, 2013; Herr, 2017; Rubin et al., 2017). (Brion-Meisels and Alter, 2018, p. 432).

In this sense, partnering with youth in authentic research processes within the context of dominant institutional structures can be particularly challenging, and requires careful attention to upholding the core commitments of the process in the face of these tensions and constraints. Learning and unlearning will need to take place to support school-based adults to partner with, rather than act on or for, young people to understand, to understand the YPAR epistemology and what it means to systematically co-construct knowledge with youth, to address underlying adultism, to understand their positionality, power, and core purpose in carrying out this work with youth, and to embody the core commitments in order to conduct YPAR with integrity.

Understanding structural barriers to integrity of implementation provides critical information about the processes underlying the SEL outcomes we observe (Durlak, 2016). This is echoed by Lacouture et al. (2015), who describe:

> …four layers of contextual factors that shape the implementation of the social programs: (1) the individual capabilities of the key actors to take the intervention forward (e.g., values, roles,

> knowledge, purpose), (2) the interpersonal relationships supporting the intervention (e.g., communication, collaboration, network, influences), (3) the institutional settings (e.g., informal rules, organizational culture, leadership, resource allocation, local priorities), and (4) the infra-structural system (e.g., political support) (p. 6).

In understanding integrity of implementation, we must also understand and document these critical contextual factors. All four of the factors described by Lacouture et al. (2015) above are relevant to the success of YPAR. With regard to the individual capabilities of key actors, YPAR is an approach to research built on the assumption that all human beings have the capacity and wisdom to engage in investigations of their lives. For this reason, academic "experts" or university researchers are not a necessary part of the YPAR process. Indeed, there are many projects that could be considered YPAR but tend to fall under the umbrella of community organizing because their primary purpose is action rather systematic study for generalizable knowledge as research is traditionally defined in academia. With that said, YPAR is a complex approach to the co-creation of knowledge that requires specific understandings about power, participation, and purpose, and therefore requires training, experience, and apprenticeship like any other approach to research or skill-development. This echoes the implementation science literature, which indicates at effective professional development is necessary for quality implementation, including an understanding of the theory behind an intervention and its core components or active ingredients (Durlak, 2016). As we have discussed in detail throughout this paper, the interpersonal relationships supporting YPAR are critical and central to the process. This is true not only for the adult and youth co-researchers, but of the relationships surrounding them, which may serve to support or to hinder or undermine the process. YPAR aims to explicitly impact the institutional and infra-structural contexts in service of creating more equitable communities; it is therefore influenced by and acts upon these features of the context, likely even more directly than the majority of traditional school-based interventions. Many of the same contextual factors that have been found to influence implementation quality are relevant considerations for the integrity of YPAR implementation; for further discussion of these factors from an ecological perspective see Domitrovich et al. (2008) and Durlak (2016). Future research may help us to better understand these conditions by explicitly studying questions of YPAR in a school-based context, including: What structural conditions need to change to enable adult-youth relationships in schools to flourish? What adult expertise and support is required to enact core commitments with integrity? And how might educators build relationships with students as co-conspirators in their search for justice when the educators themselves might also be the subject of student change efforts?

Integrity over fidelity for authenticity and impact

In this paper, we share a set of commitments from the field of youth participatory action research (YPAR), which we believe contribute to the social and emotional development of youth in culturally, contextually, and developmentally aligned ways through both setting-level and individual transformation. It is important to note that YPAR was not designed as a social emotional learning intervention targeting individuals; quite the opposite, YPAR intends to create change at the organizational, institutional, cultural, or community-level, by supporting youth to investigate and act upon the forces that shape their lives. Still, through integrity of implementation in the YPAR process – upholding a set of core commitments – we see that both setting-level transformation and individual-level social emotional learning often take place. This is likely because measuring integrity of implementation gets us much closer to understanding the key active ingredients and mechanisms of change at both of these levels. We believe that this is an important lesson for the field of social emotional learning. Perhaps it is not by understanding fidelity to a set of predetermined activities, but rather integrity to a set of core principles or commitments, that we can glean more powerful insights into the drivers of social emotional learning and development for adolescents.

Author contributions

EM and GB-M contributed conception and design of the paper. EM wrote the first draft of the manuscript. GB-M reviewed and wrote sections of the manuscript. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Abry, T., Hulleman, C. S., and Rimm-Kaufman, S. E. (2015). Using indices of fidelity to intervention core components to identify program active ingredients. *Am. J. Eval.* 36, 320–338. doi: 10.1177/1098214014557009

Berkel, C., Mauricio, A. M., Schoenfelder, E., and Sandler, I. N. (2011). Putting the pieces together: an integrated model of program implementation. *Prev. Sci.* 12, 23–33. doi: 10.1007/s11121-010-0186-1

Brion-Meisels, G., and Alter, Z. (2018). The quandary of youth participatory action research in school settings: a framework for reflecting on the factors that influence purpose and process. *Harv. Educ. Rev.* 88, 429–454. doi: 10.17763/1943-5045-88.4.429

Brion-Meisels, G., Fei, J. T., and Vasudevan, D. S. (2020). ""At our best:" youth-adult partnerships and the struggle for collective well-being" in *At our best: Building youth-adult partnerships in out-of-school time settings*. eds. D. S. Vasudevan, G. Brion-Meisels and J. T. Fei (Charlotte, NC: IAP), 289–304.

Camangian, P., and Cariaga, S. (2022). Social and emotional learning is hegemonic miseducation: students deserve humanization instead. *Race Ethn. Educ.* 25, 901–921. doi: 10.1080/13613324.2020.1798374

Cammarota, J., Ayala, J., Rivera, M., Rodríguez, L. F., Berta-Ávila, M. I., and Torre, M. E. (2018). "Theoretical lineages & guiding principles" in *Par entremundos: a*

pedagogy of the Américas. eds. D. S. Vasudevan, G. Brion-Meisels and J. T. Fei (New York, NY: Peter Lang)

Cammarota, J., and Fine, M. (2008). *Revolutionizing education: youth participatory action research in motion*. New York, NY: Routledge.

Cammarota, J., and Romero, A. (2011). Participatory action research for high school students: transforming policy, practice and the personal with social justice education. *Educ. Policy* 25, 488–506. doi: 10.1177/0895904810361722

Caraballo, L., Lozenski, B. D., Lyiscott, J. J., and Morrell, E. (2017). YPAR and critical epistemologies: rethinking education research. *Rev. Res. Educ.* 41, 311–336. doi: 10.3102/0091732X16686948

Chou, F., Kwee, J., Lees, R., Firth, K., Florence, J., Harms, J., et al. (2015). Nothing about us without us! Youth-led solutions to improve high school completion rates. *Educ. Action Res.* 23, 436–459. doi: 10.1080/09650792.2015.1013047

Ciocanel, O., Power, K., Eriksen, A., and Gillings, K. (2017). Effectiveness of positive youth development interventions: a meta-analysis of randomized controlled trials. *J. Youth Adolesc.* 46, 483–504. doi: 10.1007/s10964-016-0555-6

DeMartino, L., Fetman, L., Tucker-White, D., and Brown, A. (2022). From freedom dreams to realities: adopting transformative abolitionist social emotional learning (TASEL) in schools. *Theory Pract.* 61, 156–167. doi: 10.1080/00405841.2022.2036062

Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., et al. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: a conceptual framework. *Adv. School Ment. Health Promot.* 1, 6–28. doi: 10.1080/1754730X.2008.9715730

Domitrovich, C., Syvertsen, A., and Calin, S. (2017). *Promoting social and emotional learning in the middle and high school years*. University Park, PA: Edna Bennett Pierce Prevention Research Center, The Pennsylvania State University.

Durlak, J. A. (2016). Programme implementation in social and emotional learning: basic issues and research findings. *Camb. J. Educ.* 46, 333–345. doi: 10.1080/0305764X.2016.1142504

Durlak, J. A., and DuPre, E. P. (2008). Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am. J. Community Psychol.* 41, 327–350. doi: 10.1007/s10464-008-9165-0

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

Dusenbury, L., Brannigan, R., Falco, M., and Hansen, W. B. (2003). A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Educ. Res.* 18, 237–256. doi: 10.1093/her/18.2.237

Dusenbury, L., Brannigan, R., Hansen, W. B., Walsh, J., and Falco, M. (2005). Quality of implementation: developing measures crucial to understanding the diffusion of preventive interventions. *Health Educ. Res.* 20, 308–313. doi: 10.1093/her/cyg134

Eccles, J. S., and Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annu. Rev. Psychol.* 53, 109–132. doi: 10.1146/annurev.psych.53.100901.135153

Fine, M., and Torre, M. E. (2019). Critical participatory action research: a feminist project for validity and solidarity. *Psychol. Women Q.* 43, 433–444. doi: 10.1177/0361684319865255

Flores, K. S. (2007). *Youth participatory evaluation: strategies for engaging young people* (Vol. *14*). New York: Wiley.

Freire, P. (1973). *Education for critical consciousness*. New York, NY: Continuum.

Herr, K. (2017). Insiders doing PAR with youth in their schools: negotiating professional boundaries and healing justice. *Int. J. Qual. Stud. Educ.* 30, 450–463. doi: 10.1080/09518398.2017.1303213

Israel, B. A., Schulz, A. J., Parker, E. A., and Becker, A. B. (1998). Review of community-based research: assessing partnership approaches to improve public health. *Annu. Rev. Public Health* 19, 173–202. doi: 10.1146/annurev.publhealth.19.1.173

Jagers, R. J., Rivas-Drake, D., and Williams, B. (2019). Transformative social and emotional learning (SEL): toward SEL in service of educational equity and excellence. *Educ. Psychol.* 54, 162–184. doi: 10.1080/00461520.2019.1623032

Jones, S., Bailey, R., Brush, K., and Kahn, J. (2017). Kernels of practice for SEL: low-cost, low-burden strategies. Harvard Graduate School of Education. Available at: https://www.wallacefoundation.org/knowledge-center/pages/kernels-of-practice-for-sel-low-cost-low-burden-strategies.aspx (Accessed September 30, 2022).

Jones, S., Bailey, R., Brush, K., and Nelson, B. (2019). Introduction to the taxonomy project: tools for selecting & aligning SEL frameworks. Establishing Practical Social-Emotional Competence Assessments Work Group. Available at: https://measuringsel.casel.org/wp-content/uploads/2019/02/Frameworks-C.1.pdf (Accessed September 30, 2022).

Jones, S. M., Barnes, S. P., Bailey, R., and Doolittle, E. J. (2017). Promoting social and emotional competencies in elementary school. *Futur. Child.* 27, 49–72. doi: 10.1353/foc.2017.0003

Jones, S. M., and Bouffard, S. M. (2012). Social and emotional learning in schools: from programs to strategies. *Soc. Policy Rep.* 26, 1–33. doi: 10.1002/j.2379-3988.2012.tb00073.x

Jones, S. M., and Doolittle, E. J. (2017). Social and emotional learning: introducing the issue. *Futur. Child.* 27, 3–11. doi: 10.1353/foc.2017.0000

Jones, S. M., and Kahn, J. (2017). The evidence base for how we learn: supporting students' social, emotional, and academic development. Consensus statements of evidence from the council of distinguished scientists. Aspen Institute. Available at: https://eric.ed.gov/?id=ED577039.

Kaler-Jones, C. (2020). When SEL is used as another form of policing. Communities for Just Schools Fund. Available at: https://medium.com/@justschools/when-sel-is-used-as-another-form-of-policing-fa53cf85dce4.

Kirshner, B. (2007). Supporting youth participation in school reform: preliminary notes from a university- community partnership. *Child. Youth Environ.* 17, 354–363. doi: 10.1353/cye.2007.0066

Kirshner, B., Pozzoboni, K., and Jones, H. (2011). Learning how to manage bias: a case study of youth participatory action research. *Appl. Dev. Sci.* 15, 140–155. doi: 10.1080/10888691.2011.587720

Kohfeldt, D., Chhun, L., Grace, S., and Langhout, R. D. (2011). Youth empowerment in context: exploring tensions in school-based YPAR. *Am. J. Community Psychol.* 47, 28–45. doi: 10.1007/s10464-010-9376-z

Lacouture, A., Breton, E., Guichard, A., and Ridde, V. (2015). The concept of mechanism from a realist approach: a scoping review to facilitate its operationalization in public health program evaluation. *Implement. Sci.* 10:153. doi: 10.1186/s13012-015-0345-7

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *Am. Educ. Res. J.* 32, 465–491. doi: 10.3102/00028312032003465

Ladson-Billings, G. (2014). Culturally relevant pedagogy 2.0: a.k.a. the remix. *Harv. Educ. Rev.* 84, 74–84. doi: 10.17763/haer.84.1.p2rj131485484751

LeMahieu, M. (2011). What we need in education is more integrity (and less Fidelity) of implementation. Carnegie Foundation for the Advancement of Teaching. Available at: https://www.carnegiefoundation.org/blog/what-we-need-in-education-is-more-integrity-and-less-fidelity-of-implementation/.

Lerner, R. M., and Lerner, J. V. (2013). The positive development of youth: comprehensive findings from the 4- H study of positive youth development. Available at: https://4-h.org/wp-content/uploads/2016/02/4-H-Study-of-Positive-Youth-Development-Full-Report.pdf.

Li, J., and Julian, M. M. (2012). Developmental relationships as the active ingredient: a unifying working hypothesis of "what works" across intervention settings. *Am. J. Orthopsychiatry* 82, 157–166. doi: 10.1111/j.1939-0025.2012.01151.x

Mehta, J., and Fine, S. (2019). *In search of deeper learning: the quest to remake the American high school*. Cambridge, MA: Harvard University Press.

Meland, E., Brion-Meisels, G., and Jones, S. M. (2019). Toward a model of culturally-sustaining social emotional learning. Working paper shared at the *Social Emotional Learning and Equity Institute*, hosted by EASEL Lab and the Radcliffe Institute for Advanced Studies, Harvard University, November 2019.

Mirra, N., Garcia, A., and Morrell, E. (2016). *Doing youth participatory action research: transforming inquiry with researchers, educators and students*. New York, NY: Routledge.

Mitchell, C. M. (1983). The dissemination of a social intervention: Process and effectiveness of two types of paraprofessional change agents. *Am. J. Community Psychol.* 11, 723–739.

Mitra, D. (2008). *Student voice in school reform: building youth-adult partnerships that strengthen schools and empower youth*. Albany: State University of New York Press.

National Center for Family Philanthropy. (n.d.). Fist to five voting and consensus. Available at: https://www.ncfp.org/knowledge/fist-to-five-voting-and-consensus/ (Accessed February 21, 2023).

National Commission on Social, Emotional, and Academic Development. (2019). From a nation at risk to a nation at hope: recommendations from the national commission on social, emotional, & academic development. Aspen Institute. Available at: https://www.aspeninstitute.org/programs/national-commission-on-social-emotional-and-academic-development/.

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Rev. Educ. Res.* 78, 33–84. doi: 10.3102/0034654307313793

Osher, D., Cantor, P., Berg, J., Steyer, L., and Rose, T. (2018). Drivers of human development: how relationships and context shape learning and development. *Appl. Dev. Sci.* 24, 6–36. doi: 10.1080/10888691.2017.1398650

Owens, C. L., Johnson, A. H., and Thornton, A. (2022). Addressing equity in schools: youth participatory action research and transformative social and emotional learning during COVID-19. *Child. Sch.* 44, 48–54. doi: 10.1093/cs/cdab029

Ozer, E. J. (2017). Youth-led participatory action research: overview and potential for enhancing adolescent development. *Child Dev. Perspect.* 11, 173–177. doi: 10.1111/cdep.12228

Ozer, E. J., Cantor, J. P., Cruz, G. W., Fox, B., Hubbard, E., and Moret, L. (2008). The diffusion of youth-led participatory research in urban schools: the role of the prevention

support system in implementation and sustainability. *Am. J. Community Psychol.* 41, 278–289. doi: 10.1007/s10464-008-9173-0

Ozer, E. J., and Douglas, L. (2013). The impact of participatory research on urban teens: an experimental evaluation. *Am. J. Community Psychol.* 51, 66–75. doi: 10.1007/s10464-012-9546-2

Ozer, E. J., Shapiro, V. B., and Duarte, C. (2020). *Opportunities to strengthen SEL impact through youth-led participatory action research (YPAR)*. University Park, PA: Edna Bennett Pierce Prevention Research Center, The Pennsylvania State University.

Paris, D. (2012). Culturally sustaining pedagogy: a needed change in stance, terminology, and practice. *Educ. Res.* 41, 93–97. doi: 10.3102/0013189X12441244

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., et al. (2011). Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm. Policy Ment. Health Ment. Health Serv. Res.* 38, 65–76. doi: 10.1007/s10488-010-0319-7

Rodríguez, L. F., and Brown, T. M. (2009). From voice to agency: guiding principles for participatory action research with youth. *New Dir. Youth* 2009, 19–34. doi: 10.1002/yd.312

Roeser, R. W., Eccles, J. S., and Sameroff, A. J. (2000). School as a context of early adolescents' academic and social-emotional development: a summary of research findings. *Elem. Sch. J.* 100, 443–471. doi: 10.1086/499650

Rubin, B. C., Ayala, J., and Zaal, M. (2017). Authenticity, aims and authority: Navigating youth participatory action research in the classroom. *Curric. Inq.* 47:175194. doi: 10.1080/03626784.2017.129896

Sabol, T. J., and Pianta, R. C. (2012). Recent trends in research on teacher–child relationships. *Attach Hum. Dev.* 14, 213–231. doi: 10.1080/14616734.2012.672262

Sameroff, A. (2010). A unified theory of development: a dialectic integration of nature and nurture. *Child Dev.* 81, 6–22. doi: 10.1111/j.1467-8624.2009.01378.x

Sawchuk, S. (2021). Why high school SEL programs feel "lame"—and how to fix them. *Educ. Week* 41, 13–16. Available at: http://www.edweek.org/leadership/why-high-school-sel-programs-feel-lame-and-how-to-fix-them/2021/10

Seider, S., and Graves, D. (2020). *Schooling for critical consciousness: Engaging black and Latinx youth in analyzing, navigating, and challenging racial injustice*. Cambridge, MA: Harvard Education Press.

Sherrod, L., Torney-Purta, J., and Flanagan, C. (Eds.) (2010). *Handbook of research on civic engagement in youth*. Hoboken, N.J.: Wiley.

Simmons, D. N. (2021). Why SEL alone isn't enough. *Educ. Leadersh.* 78, 30–34.

Torre, M. E. (2009). Participatory action research map. The public science project. Available at: https://publicscienceproject.org/par-map/ (Accessed September 29, 2022).

Torre, M. E., Fine, M., Stoudt, B. G., and Fox, M. (2012). "Critical participatory action research as public science" in *APA handbook of research methods in psychology*. ed. H. Cooper, vol. *2* (Washington, DC: American Psychological Association), 171–184.

Valenzuela, A. (1999). *Subtractive schooling: U.S.-Mexican youth and the politics of caring*. Albany, NY: State University of New York Press.

Yeager, D. S. (2017). Social and emotional learning programs for adolescents. *Futur. Child.* 27, 73–94. doi: 10.1353/foc.2017.0004

Zeal, M., and Terry, J. (2013). Knowing what I can do and who I can be: youth identify transformational benefits of participatory action research. *J. Ethnogr. Qual. Res.* 8, 42–55. Available at: http://www.jeqr.org/previous-volumes/volume-8-issue-1

frontiers | Frontiers in Psychology

TYPE Policy and Practice Reviews
PUBLISHED 26 July 2023
DOI 10.3389/fpsyg.2023.1052877

OPEN ACCESS

EDITED BY
Stephanie M. Jones,
Harvard University, United States

REVIEWED BY
Ben Dyson,
University of North Carolina at Greensboro, United States
Laura Davidson,
WCSD, United States

*CORRESPONDENCE
Michael J. Strambler
✉ michael.strambler@yale.edu

RECEIVED 24 September 2022
ACCEPTED 30 March 2023
PUBLISHED 26 July 2023

CITATION
Grant N, Meyer JL and Strambler MJ (2023) Measuring social and emotional learning implementation in a research-practice partnership.
Front. Psychol. 14:1052877.
doi: 10.3389/fpsyg.2023.1052877

Measuring social and emotional learning implementation in a research-practice partnership

Nickholas Grant, Joanna L. Meyer and Michael J. Strambler*

Division of Prevention and Community Research, Yale School of Medicine, New Haven, CT, United States

The measurement of social and emotional learning (SEL) implementation is a critical part of enhancing and understanding the effects of SEL programming. Research has shown that high-quality SEL implementation is associated with social, emotional, and academic outcomes. Schools achieve these outcomes in part through organizational practices that emphasize ongoing communication, collaboration, coordination, shared decision making, and strategic planning, processes that are ideally informed by evidence. The application of implementation science to SEL has advanced our understanding of the role of implementation in achieving student outcomes. However, the development of practical approaches for measuring and supporting SEL implementation have lagged behind work on measuring student SEL outcomes. Research-practitioner partnerships (RPP), long-term, mutually-beneficial collaborations geared toward identifying problems of practice and testing solutions for improvement, are a promising means for addressing this important gap. Though implementation science and RPPs have complementary aims, there has been limited attention to the integration of these approaches in the context of SEL programming. The goal of this paper is to offer practical strategies for measuring and using SEL implementation data in schools, using the example of an RPP that used implementation science practices to guide SEL implementation. We give special attention to structures that can support the collection and use of implementation data to improve practice, as well as considerations around developing measures, considering trade-offs of data collection decisions, and conducting data analysis.

KEYWORDS

implementation, social and emotional learning (SEL), researcher practitioner partnerships, implementation science, measurement

Introduction

When education practitioners implement a social and emotional learning (SEL) approach, they usually are hoping to enhance students' social and emotional skills. Although there is a great deal of evidence on the impact of SEL programs on a range of student outcomes, SEL practices and the contexts in which they are implemented vary widely (Durlak et al., 2011; Cipriano et al., 2022). Therefore, in most cases, we cannot assume that the effects of a given SEL approach will be the same as the evidence from prior studies. In short, to know whether SEL practices "work" in a specific case, we first need to know about what was implemented, how much of it was provided, and how well it was delivered. However, this essential step in understanding effectiveness if often overlooked and the development of SEL implementation

measurement tools has been far outpaced by measures of SEL skills and school climate. As a result, much less is known about questions such as: How much of a program needs to be implemented to see meaningful effects? Which aspects of programs are most associated with effects? Perhaps more importantly, the lack of use of SEL measures in school settings among school staff makes it challenging for schools to monitor the progress of their implementation and to act on ways of improving it. It is this last point on which we place the greatest emphasis in this paper—how SEL implementation measures can be developed and used in efficient ways to support SEL practices.

The slower growth of SEL implementation measures for school use is not for a lack of emphasis from researchers, as it is well-known that the role of implementation is central to understanding program effectiveness. In fact, for decades, there has existed a sub-field of implementation science dedicated to understanding and ensuring strong program and intervention implementation (Bauer and Kirchner, 2020). Implementation science is "the scientific study of methods to promote the systematic uptake of research findings and other evidence-based practices into routine practice, and, hence, to improve the quality and effectiveness of health service" (Eccles and Mittman, 2006). Rather than solely focusing on the impact of an evidence-based intervention on outcomes, implementation science tends to focus on measuring the impact of implementation practices on intervention effectiveness in "real-world" settings (Bauer et al., 2015). These intervention practices evaluated may include program fidelity, quality of delivery, dosage, participant responsiveness or engagement, program differentiation, monitoring of comparison/control conditions, adaptation, and program reach, all of which are important when evaluating the strengths of and barriers to implementation (Durlak and DuPre, 2008).

Implementation science and social and emotional learning

In the SEL field, researchers have also stressed the importance of implementation when evaluating social and emotional learning (SEL) programs (Meyers et al., 2012; Durlak, 2016; Oberle et al., 2016). Many of these arguments emphasized the importance of studying how implementation strategies are executed to provide information about processes (e.g., school resources and values, decision making processes, team and school staff responsibilities for evaluation, and teachers and staff attitudes) that helped promote implementation success. Further, studies of SEL point to the importance of certain implementation characteristics promoting outcomes in students. For example, a 2011 meta-analysis found four qualities of effective SEL programs: (1) sequenced training approach, (2) active forms of learning, (3) focused and adequate time spent on skill development, and (4) explicit learning goals (Durlak et al., 2011). In general, evaluating these aspects of implementation quality can encompass three forms, including (a) a process evaluation in which there is simply an observation and collection of data related to characteristics of a program either before, during, and/or after it is been implemented; (b) a formative evaluation in which data are collected and shared with the implementation team in order to improve and modify processes of implementation; or (c) a summative evaluation in which data are collected to study the impact of the implementation strategies on program outcomes (e.g., rates or quality improvement of an program; Bauer et al., 2015). In this paper, we primarily focus on ways in which implementation measures can be used in formative ways, but we also address their use in summative evaluation.

At the core of implementation, science is an over-arching goal: to bridge the gap between prevention research and practice by way of developing and evaluating evidence-based interventions and enhancing their use (Chambers, 2012). One framework that illustrates these processes is the Interactive Systems Framework (ISF), which includes three core systems that co-function to improve dissemination and implementation practices: (a) the Prevention Synthesis and Translation System, (b) the Prevention Support System, and (c) the Prevention Delivery System (Wandersman et al., 2008). The Prevention Synthesis and Translation System involves gathering, synthesizing, and translating research literature for practitioner use; the Prevention Support System involves providing innovation-specific support (i.e., intervention related training and providing information and technical assistance with intervention goals) and general support with building the organizational infrastructure and support; and the Prevention Delivery System involves implementing the service activities planned after building capacity. In the example presented in this paper, we discuss what might be considered yet another more overarching framework for supporting implementation—research-practice partnerships (RPPs).

In the context of SEL programs, there are other key implementation-related questions that need to be addressed such as: How much of a program needs to be implemented to see meaningful effects? Which aspects of SEL programs are most associated with effects? Perhaps, more importantly, the lack of practical SEL measures for use in school settings by school staff makes it challenging for schools to monitor the progress of their implementation and to make improvements as needed. It is this last point that we place the greatest emphasis on in this paper—how SEL implementation measures can be developed and used in efficient ways to support SEL practices. We demonstrate these measures using an example of a research-practice partnership (RPP) that used implementation science practices to guide the implementation of SEL in the Bridgeport Public Schools. In the following sections of the paper, we first briefly define and explain the purpose and practices of implementation science. Lastly, we define and outline a framework for RPPs and discuss the implications for implementation science methods within SEL program development.

Research-practice partnerships

Research-practice partnerships (RPPs) are long-term collaborations between researchers and practitioners that aim to improve education by conducting mutually beneficial research (Coburn et al., 2013; Farrell et al., 2021). RPPs bring together stakeholders from the fields of education research, policy, and practice—fields that are sometimes siloed—to engage the diverse expertise of these stakeholders. RPPs use a variety of strategies to manage the challenges of working in collaboration, including power dynamics that arise from differences in professional backgrounds, individual perspectives, organizational cultures, inter-organizational politics, and much more (Denner et al., 2019; Farrell et al., 2021; Yamashiro et al., 2023). According to a review by Phelps (2019) of 56 studies on challenges in research-practice partnerships in education,

building *organizational infrastructure* (e.g., defining roles, decision-making processes, and communication strategies), *shared meaning* (i.e., identifying shared values and understanding of goals), and *trusting relationships* (e.g., favoring equality over hierarchy, respecting the value of diverse contributions) are essential in RPPs.

The guiding principles inherent in the ISF framework align well with the research-practice partnership model. First, ISF posits that research and practice should mutually build upon one another through using scientific literature and evidence-based research methods. Secondly, the ISF invites shared decision-making and collaboration, communication, and strategic planning and coordination among all parties involved in the dissemination and implementation of the intervention (Wandersman et al., 2008; Chambers, 2012). Regarding this latter point, the ISF proposes that multiple parties (i.e., researchers, prevention practitioners, funding agencies, and support agencies) be involved and utilize their scientific knowledge and expertise to (1) understand the capacity required to deliver a specific service and (2) engage in data driven practices to build organizational capacity to promote an intervention's success (Wandersman et al., 2008; Chambers, 2012).

RPPs and implementation science in social–emotional programming

The principles of RPPs and implementation science are especially useful within education partnerships that aim to promote SEL competencies among students, such as self-awareness, self-management, social awareness, relationship skills, and decision-making skills that are especially useful for supporting developmental transitions into adulthood (Oberle et al., 2016). Historically, schools have primarily focused on academic outcomes and performance, however, schools have been increasingly integrating SEL programming given its connection with improvements in academic performance, student conduct, school climate, peer relationships, and teacher well-being (Durlak et al., 2011; Oberle et al., 2016; Herrenkohl et al., 2020).

Though impactful, the process of adopting and implementing SEL programming school-wide can be challenging; without buy-in from teachers, school staff, and district leaders, SEL practices and policies will be unsustainable and difficult to implement (Herrenkohl et al., 2020). An RPP can help to address these challenges if it attends to the strategies identified above: building *organizational infrastructure* (e.g., the availability of school resources to assist in coordinating and communicating about SEL), *shared meaning* (i.e., establishing values and goals related to SEL that are shared by stakeholders throughout the research and practice organizations), and *trusting relationships* (i.e., teachers' perception that SEL programming is in their best interest, district and school leaders' belief that they will benefit from partnering with researchers, and researchers valuing the expertise of practice-side partners).

Research-practitioner partnership approaches have also been used to enhance the fidelity and sustainability of SEL practices (Ackerman and Skoog-Hoffman, 2020). One notable example of a large scale SEL-related RPP is the Collaborative Districts Initiative of the Collaborative for Academic, Social, and Emotional Learning (CASEL). In 2011, CASEL began partnering with eight large school districts to support and study high-quality SEL implementation. For example, the CASEL-Lowell partnership aimed to understand how to integrate and leverage SEL programming in elementary math classes in order to support teachers (Ackerman and Skoog-Hoffman, 2020). Their collaboration yielded insights indicating that the practices were vital for fostering equitable learning and development for children from diverse backgrounds.

Purpose of the current paper

The main goal of this paper is to offer practical strategies for measuring and using SEL implementation data in schools that draw upon practices drawn from implementation science and the RPPs. Throughout the paper, we highlight examples from an RPP focused on the implementation of a social and emotional learning initiative in an urban district. After providing context about the partners and the SEL initiative, we discuss ways in which researchers and practitioners can work together to develop implementation measures and structures that facilitate the sustainable collection and utilization of data. We also discuss methodological trade-offs concerning data privacy and data linking important considerations when analyzing collected data and reporting findings.

The Bridgeport public schools SEL initiative

We begin the remainder of this paper by describing the Bridgeport Public Schools (BPS) SEL Initiative, which grew out of the Yale-BPS SEL Partnership, a research-practice partnership that began in 2013 as a collaboration between BPS, The Consultation Center at Yale School of Medicine (YSM), and Yale Center for Emotional Intelligence. This partnership aimed to build the social and emotional skills of BPS administrators, teachers, staff, students, and their families. At the start of the Yale-BPS SEL Partnership in the 2013–14 school year, the school district was serving 19,231 students enrolled in grades PK-12 at 28 elementary/middle schools, seven high schools, one early childhood center, and two alternative schools. Approximately 49% of students identified as Hispanic or Latino of any race, 37% of students identified as Black or African American (and not Hispanic or Latino), and 100% of students qualified for free or reduced-price lunch [Connecticut State Department of Education (CSDE), 2023]. In 2014, Bridgeport was the Connecticut's most populous city with about 148,000 residents; it had an estimated median household income of approximately $43,000 for 2012–2016 (compared to approximately $72,000 for the state), making it one of the poorest cities in the state (Connecticut Data Collaborative, 2023). In addition, the school district faced challenges related to discipline concerns and low academic performance. In 2013–2014, the BPS rates of chronic absenteeism and suspensions were more than double the statewide rates [Connecticut State Department of Education (CSDE), 2023].

The Yale-BPS SEL partnership included a collaboration with a diversity of stakeholders/practitioners with expertise in school policy and practice. For instance, members of the university team worked directly with representatives from the district leadership team (e.g., superintendents, assistant superintendents, and SEL coordinators) and school leaders (principals and assistant principals). As the practice partner, the district led SEL decision-making and managed the implementation of SEL programming. As the research partner, The

Consultation Center supported SEL implementation and continuous improvement through data collection, analysis, and reporting, serving as a formative and summative evaluation partner over the first 5 years of the partnership.

The Yale-BPS SEL Partnership began in the summer of 2013 with a 5-year grant from the Tauck Family Foundation awarded to The Consultation Center. The funding was initially intended to support SEL implementation in one school with the intention of gradually scaling up the work to include 3–4 schools. However, when the superintendent left his role mid-way through the 2013–14 school year, the new superintendent was so enthusiastic about the work that she charged the partnership to expand the work districtwide; additional funding was sought and acquired to do so. The overall goal of the partnership was to: (a) promote learning, healthy interpersonal relationships, and sound decision-making; (b) foster safe, supportive, and respectful classrooms and schools; (c) utilize measures relevant to these goals that can be used to measure progress, gage impact and guide improvements; and (d) create a model for school improvement that actively engages all stakeholders (Strambler and Meyer, 2018). Table 1 describes the focus of the 5-year Yale-BPS SEL Partnership; related materials are available on Open Science Framework at https://osf.io/nwzrs/. In 2018, The Consultation Center transitioned to an as-needed consultative role and the Bridgeport Child Advocacy Coalition (BCAC) at RYASAP became the partnership's co-leader alongside the district. As of the last quarter of 2022, the SEL initiative is still in place and has persisted across four superintendent transitions.

RULER

The BPS SEL initiative began with the introduction of RULER, which was developed at the Yale Center for Emotional Intelligence (Brackett et al., 2019). Unlike other SEL frameworks that focus on various inter-and intra-personal competencies (e.g., CASEL; see Blyth et al., 2018), RULER is an evidence-based approach to social and emotional learning designed to enhance emotional intelligence in educators and students. RULER stands for the five key emotion intelligence skills, this approach intends to promote: recognizing, understanding, labeling, expressing, and regulating emotions. The RULER approach relies on first teaching educators (principals, teachers, and school staff) to appreciate the significance of their own and their students' emotions. The RULER approach asks educators to value the skills of recognizing, understanding, and managing emotions; to learn and model these skills; and to support, teach, and encourage students to develop these skills. Instead of being taught as a separate lesson or set of activities, RULER is designed to be integrated into the everyday routine of teaching and learning, by infusing it into classroom practices and the curriculum. For example, the RULER *feeling words* approach lays out a process for building students' emotion vocabulary that can be applied to fiction or non-fiction texts in the curriculum. As described by Brackett et al. (2019), RULER also provides four *anchor tools* that can be used across the day and the school year to support the development of social–emotional skills: the *Classroom Charter*, *Mood Meter*, *Meta-Moment*, and *Blueprint*. For example, the RULER *Mood Meter* is a tool that teachers and students can use together or independently to develop awareness of their emotions and how to shift among emotions to enhance learning.

Yale Center for Emotional Intelligence typically uses a train-the-trainer model, in which a school or district identifies a small group of school or district personnel (known as a RULER Implementation Team, RIT) to attend RULER trainings conducted by YCEI. A school RIT typically includes at least three people: a school leader (principal, assistant principal, dean of students, etc.), a school counselor or social worker, and at least one teacher. When they return to their school, the members of an RIT are expected lead professional development for their colleagues and support RULER implementation. RITs are encouraged to first implement RULER among their faculty before classroom implementation begins. In Bridgeport, the districtwide SEL initiative began with a readiness/leadership development year, when all district and school leaders participated in a series of workshops, meetings, and individual coaching focused on the development of emotional-intelligence leadership mindsets and skills *before* school teams began RULER training. The BPS SEL initiative also had the benefit of a full-time SEL coordinator, an experienced educator with RULER training, who provided focused support to school teams starting when the teams began RULER training.

In practice, schools vary in their readiness to implement RULER, which may relate to school administrators' willingness or ability to dedicate professional development time to RULER, the preparedness of RIT members to train their colleagues, or teacher buy-in. Some schools launch RULER quickly and with fidelity, while other schools are slower to introduce the approach to their teachers and ultimately, their students. As noted above, the implementation quality for any intervention is likely to influence the intervention's effects. The central goal of the Yale-BPS SEL Partnership was to monitor SEL implementation across schools to identify areas of strength and areas of need, so that resources and support could be allocated appropriately.

TABLE 1 Timeline for the Yale-BPS SEL Partnership.

School year	Focus
2013–2014	Strategic planning, capacity building, and RULER pilot
2014–2015	Leadership development and capacity building at district and school level
2015–2016	Integration of SEL at all schools
2016–2017	Continued SEL implementation and evaluation
2017–2018	Sustainability of SEL implementation

A description of each year's activities is available on Open Science Framework at https://osf.io/kdh5t.

Developing implementation measures

When we set out to develop implementation measures for the partnership, a high priority was placed on measures that were useful, practical, low-burden, and inexpensive. Accomplishing this meant giving special attention to balancing feasibility and rigor. The first step in the process involved holding discussions with district leaders about what data were meaningful to them. To maximize the value of data collection, we discussed which data would not only be valuable for assessing the implementation progress, quality, and signals of impact but could also be useful for other related initiatives in the district. For example, we worked with the district to develop a school climate

survey that could inform the districts' Safe Schools Healthy Students project as well as the SEL initiative. In these discussions, it was crucial to define how the data would be collected, who would use it, for what purposes it would and would not be used, and how interim results would be disseminated. This issue is crucial because data that is collected without being used in meaningful ways by the district is a burden without adequate benefit to the district. Ideally, any SEL-related data collected would be useful at multiple levels—by district leaders, school leaders, school SEL teams, and potentially, by teachers, as each group plays an important role in improving implementation. District leaders can identify resources and supports that schools need for high-quality SEL implementation. School leaders must provide the supports and vision to implement SEL practices. School SEL teams can serve as resources for one another, especially those that have consistently strong implementation. And teachers are essential as the main implementers of SEL practices in the classroom.

In terms of developing practical measures, one important consideration is the type and nature of the measures used. This is essential because, depending on the type of measure, data collection can be very time-consuming. Because observational measures require a great deal of time and resources to use, SEL implementation data is often collected through teacher self-report measures that ask teachers about the SEL practices that they are engaged in. In selecting these measures, it is important to consider the measures' sensitivity to change. That is, the ability of the measure to pick up on change of what it is capturing over the period of time that the measure is being used. The same measure that is intended to capture change over a 6-month period may not be sensitive enough to pick up on change over a 2-month period. Thus, it is valuable to use theory to develop or select items that have a chance of changing over the time being assessed. While some established measures provide psychometric information regarding sensitivity, typically this needs to be assessed following the collection of the data by inspecting change scores. If there is very little change, there could be two possibilities—that the measure genuinely did not change or that the measure was not sensitive enough to pick up on the change. If data are collected over multiple time points it provides various time points to examine such change. We used a variety of implementation measures and other measures over the course of the Yale-BPS SEL Partnership to balance rigor, feasibility, and sensitivity to change, and to account for potential bias. Figure 1 depicts the initiative's theory of change and the measures used at each stage (Strambler and Meyer, 2018). As shown in Figure 1, the partnership developed two types of SEL implementation measures: SEL implementation logs and SEL implementation surveys, which are discussed below, along with other SEL implementation measurement approaches that we considered but did not use. Note that this paper does not discuss the leadership development surveys and RULER training surveys that were used to assess the readiness and training phases, nor does it describe the school climate and SEL student survey and administrative data that we used to assess student outcomes.

SEL implementation logs

The logs were designed to track SEL-related trainings and activities at the school level in the first 2 years after RULER training. During the final day of RULER training, we met with the school teams to explain the purpose of the SEL implementation log and asked each school's SEL team to identify a contact person who would be responsible for completing and submitting the school's SEL log each month. The researchers worked with the SEL coordinators to design this measure to be brief and easy to complete, with the final version of the measure consisting of only six items, two of which were optional open-ended questions. The first item asked the team to estimate how much whole-staff meeting or professional development time the school had spent on each of six RULER topics over the past month, with responses ranging from none to over 60 min. The second question asked the SEL team to rate the knowledge of their school staff regarding the same six RULER topics, with response categories of *Beginning, Progressing*, and *Advanced*. The third and fourth items asked the SEL team to describe any RULER events held for students and parents, respectively, during the past month. The last two questions asked the SEL team to share any comments about how implementation was progressing and what additional supports they needed to support implementation. A copy of the 2016–2017 SEL log is available on Open Science Framework at https://osf.io/h86am. The SEL log was mailed to school contact people monthly as an online survey on the Qualtrics platform, with additional reminders sent to those schools who did not complete the survey on time. At the end of the month, we compiled the results and shared them with the SEL coordinators, so that these district leaders could follow up with individual schools, as needed.

SEL implementation surveys

As described by Yeager et al. (2013), we needed to manage the level of burden on school personnel while collecting detailed data about on-the-ground implementation that could inform improvement. Classroom observations were impractical given the scale of the district-wide initiative and funding constraints. Instead, we took a more practical approach to measurement that focused on implementation specific related questions that would have more direct implications for service improvement strategies (Yeager et al., 2013); we chose to conduct periodic surveys of school leaders (i.e., principals and assistant principals), teachers, and other district personnel. In close consultation with the RULER developers, SEL coordinators, and other district personnel, we created surveys that asked educators about their perceptions of SEL implementation in their school and district. While developing the survey, we met with our partners several times to discuss potential survey items, in order to ensure we were prioritizing domains of interest, asking questions clearly and efficiently, and keeping the survey to a manageable length. Ultimately, we asked about *quality of implementation, support for program implementation, teacher attitudes, perceived barriers to implementation, principal factors, and professional experience*. Different respondent groups saw different sets of questions, as documented in Table 2. We continued to meet with our partners before and after every survey administration to discuss survey data and potential revisions to survey items. Although we generally sought to keep items consistent to allow for comparisons over time, our partners' input led to some revisions to improve clarity and address evolving priorities. The 2015–16 through 2017–18 versions of the survey are available at https://osf.io/n8dfy/.

As with the logs, a key decision was how often to administer the implementation surveys to staff. We consulted with our district partners about how many times per year would be feasible while taking into

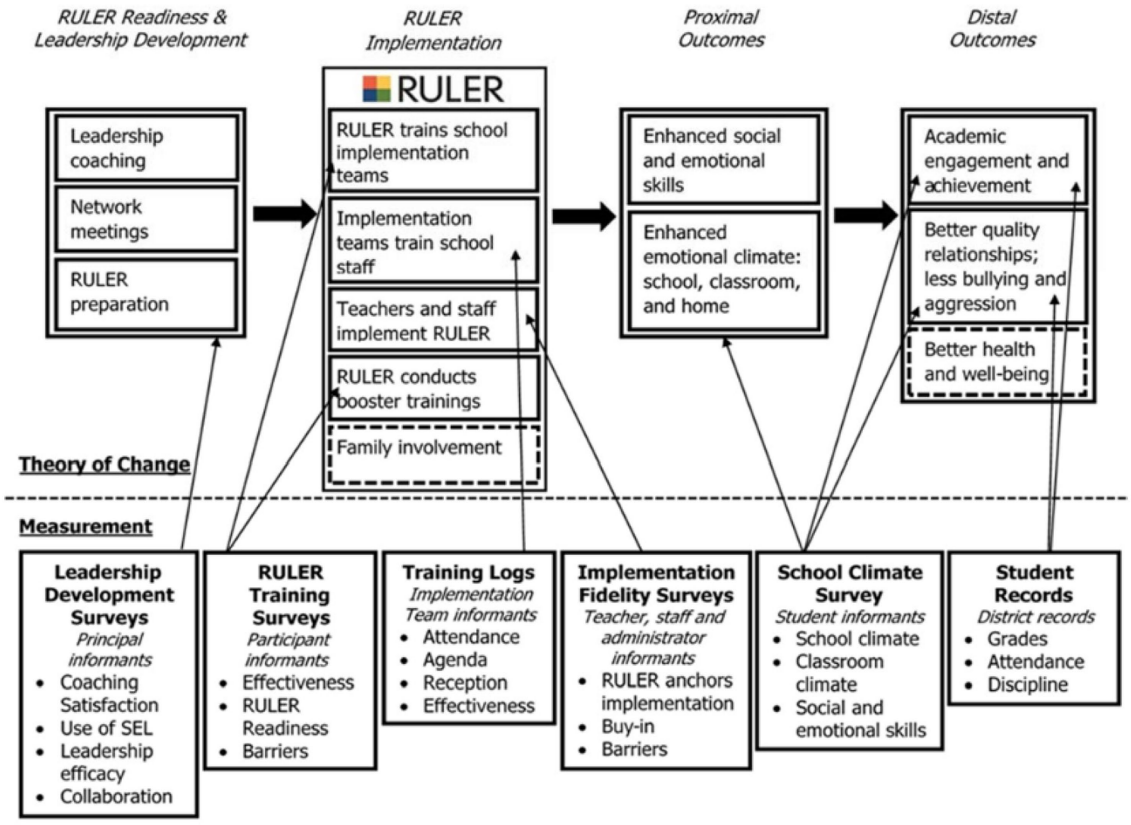


FIGURE 1
Bridgeport SEL initiative theory of change mapped to measurement tools.

consideration the time staff had available to complete the survey and the other surveys that they were expected to complete a greater opportunity to observe change over time. It is important note that district buy-in was essential to meaningful data collection. Although the district emailed the survey link to all teachers and staff members, response rates were relatively low for the initial survey, and the superintendent and SEL coordinators expressed concern that the data may not be sufficiently representative of the experiences of district personnel as a whole. The superintendent and our team were also concerned that we would not see signals of change in SEL implementation if we only surveyed teachers and staff twice in the fall and spring. We collectively agreed to add a winter survey administration the following year, and we collaborated with our partners to increase response rates in subsequent surveys. For example, the superintendent pointed out that all schools held a monthly staff meeting on the first Wednesday of the month. She asked us to schedule each future survey to launch the day before a monthly staff meeting, and she directed principals to allocate time during those meetings for teachers and staff members to complete the survey. In addition, we agreed that during each survey administration, the research team should share weekly reports showing response rates by school so that the district could follow up with the principals of schools with low response rates to ask them to re-send the survey link to their teachers and staff. These collaborative efforts increased response rates dramatically and increased the confidence of district and school leaders in the value of the data.

Other potential data collection

While we did not have the capacity to support studying how data were used by the teachers and administrators, this can be an important process for understanding the effectiveness of the data use process. For instance, one option researchers could consider are mixed methods in which the quantitative surveys described herein are paired with qualitative interviews that focuses on how practitioners interpreted and used the data to inform decision-making. Utilizing mixed methods has multiple benefits; combining elements of qualitative and quantitative research methods and analyses allows researchers to clarify and/or develop their research approach to converge, corroborate, expand, or elaborate on research findings (Schoonenboom and Johnson, 2017). Though survey data have the potential to produce evidence that is generalizable to a larger population, the structured format limits the ability to document individuals' subjective experiences, especially when such experiences do not fit well within constructs assessed by surveys. Interviews and focus groups, however, are very useful when the objective is to understand how individuals construct meaning of what is relevant and salient to them, and descriptive details about context-specific actions within settings (Nowell et al., 2017). Conducting follow-up interviews with teachers, administrators, and district leaders could shed light on how they personally experience the implementation strategies applied. Specifically, qualitative interviews could (a) help discern possible

10.3389/fpsyg.2023.1052877

TABLE 2 SEL implementation survey domains by respondent.

Measure	Respondent		
	Teachers	Other staff	Principals/ APs
Quality of program implementation			
Anchors knowledge	X	X	X
Introduction to anchors	X		
Use of anchors	X	X	X
Integration of anchors			X
Fidelity	X		
Perceived self-efficacy	X	X	X
Support for program			
Principal support	X	X	
Internal support	X	X	
External support	X	X	X
Program receptivity			
Goodness-of-fit	X	X	X
Participant engagement	X	X	X
Perceived program effectiveness	X	X	X
Perceived barriers to implementation			
Barriers			X
Experience			
Professional experience	X	X	X

strengths and areas in need of change within implementation strategies, (b) provide multiple perspectives across the leadership hierarchy, which can identify areas of miscommunication and converging and diverging opinions about actions taken; and (c) create discussion of improvement recommendations that are grounded in the practitioners' experiences.

Teacher privacy and linking of implementation data

One necessary decision to make when collecting implementation data from school staff is whether to collect the data in a confidential, but identifiable way, or anonymously. From a research perspective, it is advantageous to collect the data in an identifiable way since it allows for individual teachers' implementation practices to be examined over time; it also allows for teacher-reported implementation data to be linked to student outcomes (assuming these data are accessible). However, educators' concerns about privacy need to be taken seriously in school-based research, to ensure that educators feel comfortable sharing their perspectives. It is not uncommon for teachers to feel uncomfortable with providing identifiable implementation due to concerns about it being used in an evaluative way rather than a supportive one. Even if the data collector were to use methods to ensure confidentiality, teachers may be understandably skeptical about whether their privacy is protected adequately. Therefore, one is often faced with a tradeoff. If data are collected anonymously, it protects privacy but limits the ability to link implementation data to students' outcomes. Yet, if data are collected in an identifiable manner, it allows for linking and other data analytic options, but runs the risk of losing the trust of the school staff and potentially biasing educators' responses toward reporting in ways that they view as more favorable. Especially in a partnership context, if one suspects that a substantial portion of the teaching body is concerned about privacy, the most prudent choice is to collect data anonymously given that trust among partners is essential for all aspects of the work. For the reasons noted above, for the Yale-BPSSEL Partnership, we opted to collect survey data anonymously, where the only identifying characteristics were the teachers' school and role. While this prevented us from linking teachers' responses over time and from linking teachers to students to analyze implementation data at the classroom level, we were able to create school-level implementation measures and to link them to student outcome data.

Summarizing and analyzing implementation data

Once SEL implementation data are collected, there are two broad ways in which the data can be summarized to use for formative purposes—descriptively and statistically. Descriptive summaries (for example, frequency tables), visualizations (for example, frequency plots), or combining items into composites using mean or sum scores, are especially useful for using data continuous improvement purposes. These data can also be organized thematically in ways that are most meaningful to staff members. For example, a self-report measure might involve a collection of items organized around the components of an SEL program. In such cases, decisions might be made about reporting single items under category headings or creating mean and/or sum scores of the items of such items. In general, it has been our experience that when using Likert-type continuous items, means are more interpretable than sum scores.

In the case of the Yale-BPS SEL Partnership, the way we presented data to our partners depended on the audience. As noted above, we provided the SEL coordinators with a tabulation of implementation log responses at the end of the month. The district was not interested in a summary or descriptive statistics for the SEL logs, because the SEL coordinators were using each school's response to guide their interactions with that school. The monthly report gave the SEL coordinators a snapshot of each school's progress that the SEL coordinators could use to start conversations and provide tailored supports. For example, if a school's SEL team reported that they had hosted their second all-staff RULER training, the SEL coordinators could ask about how it went. If a school's SEL team reported that they had not done anything in the past month, the SEL coordinators could inquire about barriers and offer their support.

For the SEL implementation surveys, which received responses from hundreds of educators, it was essential to summarize and visualize overall responses descriptively and also to share each school's results with its leaders. For this reason, we communicated results from each survey to our partners in four formats. First, we generated a

district-level summary report, which showed frequencies and means for key survey items, to share with our district administrator partners. Second, we generated a school-level report for each school, which included frequency tables and plots for each item, to share with the principal and assistant principal at each school. Third, we generated a district-level detailed report, which showed frequencies and means for key survey items broken out by school and with comparisons over time, to share with our district administrator partners. Finally, we presented survey results to all district and school leaders as part of one of the districts regularly scheduled meetings for administrators. At these meetings, we focused on a small number of key items and discussed how responses were changing over time. For example, we reported the percentage of teachers who said they had used a specific SEL practice with their students in the past week. We typically provided the district-level summary report within 10–14 days of the survey administration, so that the SEL coordinators, the superintendent, and her leadership team could see an overview of the data when it was still quite recent. Although it took more time to produce the detailed reports, we made sure to share them with district and school leaders within 1 month while the results were still relevant. We were usually invited to present at the first administrator meeting after the survey.

Although the Yale-BPS SEL Partnership did not have a process for systematically tracking how district or school leaders received and used data from the SEL survey, we believe it useful to provide some anecdotal evidence about how we built trust and buy-in around the collection and use of data. Our first presentation to the BPS administrative council was during the leadership development year. Our first goal was to explain The Consultation Center's role within the districtwide SEL initiative that the superintendent had launched the preceding summer. Our second goal was to explain what data would be collected and why. In this initial presentation, we explained that in close collaboration with district leaders, we would use data to know whether we were achieving our goals, to improve professional development programming provided by Yale Center for Emotional Intelligence, and to facilitate evidence-based decision-making. We made it clear that we were not evaluating school leaders or teachers, and we emphasized how we would protect the confidentiality of survey respondents throughout the project. These themes remained central when we presented to the administrative counsel two more times that year and in subsequent years, as well as remaining central in our meetings with the SEL coordinators and superintendent. Over time, we observed greater interest and engagement among school leaders during our presentations and more instances where they approached us with questions in person or via email. We took these interactions as signs of greater trust although we do not have data to this effect.

Meeting with these partners over time also allowed us to build interest in the data we were sharing, especially when we were able to build curiosity. Initially, the SEL coordinators valued qualitative data from the SEL logs more than quantitative data from the SEL survey. We suspected that part of the problem was that while it was challenging to consider each survey item separately, the SEL coordinators found it challenging to interpret or use the reported scale scores. We also suspected that while it was overwhelming for them to review 30 school-level reports, the SEL coordinators were interested in school-level results. To address these perceived concerns and to promote their interest and investment in the data, we designed an experience to help the coordinators interpret and connect with the data. Specifically, we brought three simple bar charts to a meeting with the SEL coordinators, each of which showed the median value by school for one of three items in the "teacher self-efficacy" scale, but the schools were not labeled by name. This approach piqued SEL coordinators' curiosity as they began looking for patterns to try to guess which school had produced which values. This practice of observing patterns with real data provided a basis for us to discuss the basis for applying these skills more broadly to the full reports we provided. The meeting is memorable because it marked a shift in the SEL coordinators' investment in the SEL survey as a source of meaningful data. We also aimed to build curiosity in school leaders about their schools' SEL survey data by presenting district-level results to them at a meeting before they received their school-level reports. We found that sharing the district-level results at these meetings got school leaders excited to receive their individual reports and to see how their school-level data would compare to the district as a whole.

While descriptive summaries can be useful for research purposes, sometimes summarizing the data requires using more sophisticated techniques. For example, when the data are intended to be used for predicting student outcomes from implementation measures. Reporting such findings can be challenging when sharing them with practitioners who usually do not have the research background to interpret technical statistical findings. In such cases like ours, it is necessary to translate findings in a way that is interpretable to practitioners. When conducting analyses focused predicting student outcomes from implementation, we used multilevel confirmatory factor analysis to create school-level measures of implementation, and then used multilevel modeling to examine the relationships between these measures and outcomes. It would have been inappropriate to report such findings as one would for a scientific journal. Instead, in a brief report, we described the goal of these statistical techniques in lay terms and summarized results visually. As shown in the example in Table 3, we use symbols to indicate whether effects were effects were present (positive or negative) or absent (blank) and color-coded these findings to indicate whether they were in the expected direction (green if yes, yellow if in the opposite of the predicted effect). Researchers can increase the level of detail in these types of depictions, such as including regression coefficients and other relevant statistics, depending on the background of the audience to which they are presenting.

Approaches for building sustainable implementation data use

Anytime a partnership is established between researchers and practitioner careful attention needs to paid to sustaining the practices implemented. However, a sustainability practice that is often underappreciated is considering the sustainability of data use. Also. while it is common for sustainability to become the focal point toward the end of implementation-supporting resources such as grant funding coming to an end, we argue for the importance of building sustainable data practices from the start. We have found that regular and meaningful opportunities to promote the engagement of partners with data can go a long way for deepening the roots of the partnership and increasing the chances for sustainability—from the development of measures to the collection of data to exploring what the data are

saying. It is often the case that education practitioners find the collection and use of data as detached from the "real work" of teaching students. This is in part because researchers are not commonly trained on the nuances of managing the factors associated with building enduring data procedures and practices that work for *practitioners* as opposed to other researchers. In the dissemination space, there are four stakeholders that are important to consider engaging around data: (1) key district leaders; (2) school leaders and their SEL implementation teams; (3) teachers and other school staff; and (4) community partners. Engaging such a broad "web" of stakeholders is especially valuable in urban settings where there is a higher rate of leadership transition at the school and district levels. Sharing valuable information about the progress of SEL implementation with various stakeholders can help keep the partnership engaged and motivated to continue their mission, even in the midst of top-level leadership changes such superintendents.

In Yale BPS SEL Partnership, there were four partnership structures that were developed and utilized to support the dissemination of implementation data. One structure was the establishment of SEL teams at each school as noted above. As depicted in Figure 1, the school-based SEL teams were established prior to SEL implementation during a "RULER-readiness" phase of the initiative's rollout. These teams, consisting of 4–5 school staffs supported RULER trainings and the monitoring of SEL implementation at the school level. The second structure was an SEL task force, which consisted of district-level members and representatives from community-based organizations and universities. The SEL task force met quarterly to provide updates about the initiative's progress and opportunities for input from members. The Task Force served as a valuable venue for the evaluation team to provide status reports about SEL implementation across the district and to receive input about improving implementation.

The third structure was the establishment of monthly meetings between the university partners and the district partners, specifically the researchers, the RULER developers, the SEL coordinators, and the superintendent. These meetings provided a crucial opportunity for communication and strategic planning and demonstrated the district leadership's deep commitment to the partnership. The fourth structure for supporting sustainable implementation and data practices was establishing a SEL coordinator position at the district level—a person who is responsible for overseeing and supporting SEL implementation. This role greatly facilitated the use of implementation data at both the district and school levels. The coordinator proved to be instrumental in using the implementation data in ways that were palpably useful. In Figure 2, we depict the flow of data to the coordinator and how the coordinator used it. As shown, the evaluation team would compile implementation data from the SEL team logs and educator surveys in addition to student outcome data and share it with the SEL coordinator, who in turn would use the data to identify where implementation was going well and where it needed more improvement. These data would be used to inform her regular visits to the schools focused on supporting and strengthening implementation.

TABLE 3 Practitioner-oriented reporting of the statistical association between SEL implementation (as reported by teachers) and school climate outcomes (as reported by students).

	Student-teacher trust		Rules and norms		Emotional climate		Peer support	
Grade level	3–5	6–8	3–5	6–8	3–5	6–8	3–5	6–8
Fidelity of Implementation		+		+	+		+	
Support for Implementation					–			
Perceived Effectiveness							–	
n	1,941	2,081	1,946	2,085	1,938	2,067	1,941	2,080

Note: Plus signs indicate positive correlations, and minus signs indicate negative correlations. Green shading indicate that correlations are in the expected/favorable direction, whereas yellow boxes indicate counterintuitive/unfavorable effects.

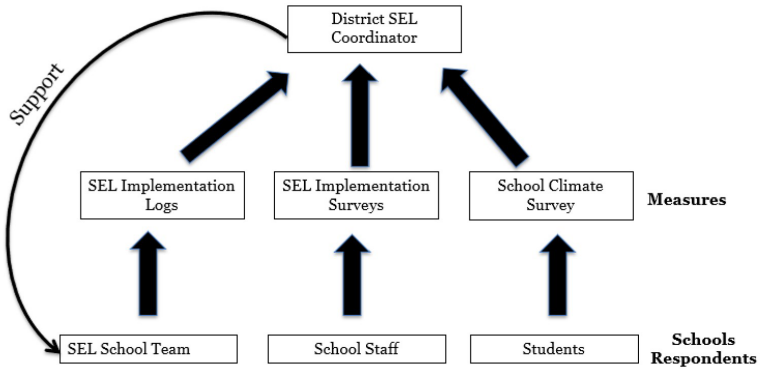


FIGURE 2
Illustration of data flow and data use for school supports.

Conclusion

Whenever one is interested in studying the effects of SEL, it is also important to consider coupling outcome measures with measures of SEL implementation. Doing so allows one to move beyond understanding whether SEL programming works to understanding why and how it works. Although the development of measures of SEL implementation is lagging measures of SEL outcomes, the field is rapidly growing in this area with implementation being afforded a greater deal of attention. In this paper, we focus on strategies for advancing SEL implementation. First, developing useful measures of SEL implementation that are feasible to use and capture meaningful indicators, provides valuable information to district and school leaders about the progress of SEL implementation. This information is especially helpful for understanding where implementation progress is being made, where more supports are required, and how to make use of effective implementation happening in schools to support the less effective ones. As we discuss in this paper, to make these measures as useful as possible, researchers should be in regular consultation with district and school leaders during the development/selection of measures and the methods for administering them. To ensure that the measures are aligned with the theory of the program that is being implemented, it is also essential that one consult the program's theory of change, and/or the program developers if possible. When SEL practices are "home grown" by districts or schools, the developers should create a theory of change or logic model that articulates a clear process about the key elements of the practices and how they are anticipated to effect outcomes. In short, the aim of these practice recommendations is to make measures that are useful, practical, and reflective of the theory and mechanisms expected to change outcomes.

We also emphasize the importance of building and maintaining relationships with practitioners in the development and administration of implementation measures. While this is important to do in any context, an especially effective way of doing this is through research-practice partnerships (RPPs) that create opportunities for researchers and practitioners to have ongoing collaborative interactions with each other that are mutually beneficial. In such partnerships, practitioners benefit by expanding their capacity to conduct implementation evaluation and research—a capacity that is often very limited in schools and districts. Practitioners also have opportunities to make valuable contributions to the work such that it reflects what measures that are important to them. For researchers, RPPs can help make the research they care about more relevant and applicable in the "real world." RPPs can also help advance scientific knowledge by improving our understanding of the nuances of setting features that act to enhance and hinder high quality implementation. Finally, by advancing our knowledge of implementation and connecting them to outcomes, we can improve our understanding of the "active" ingredients most important to impacting outcomes.

Author contributions

NG contributed to conceptualization of the manuscript, writing of the manuscript, and reviewing. JM contributed to conceptualization of the research and manuscript, carrying out the described research, writing of the manuscript, and reviewing. MS contributed to conceptualization of the research and manuscript, carrying out the research, writing of the manuscript, and acquiring funding. All authors contributed to the article and approved the submitted version.

Funding

The Tauck Family Foundation supported this research.

Acknowledgments

We thank all of the Bridgeport Public Schools leaders, teachers, and staff who committed many hours to the work described in this article, including Fran Rabinowitz, Aresta Johnson, Michal Testani, Alana Callahan, Helen Moran, and Carrie Ramanauskas. We also appreciate our partners at the Bridgeport Child Advocacy Coalition at RYASAP, including Mary Pat Healy, Mory Hernandez, Ashley Blanchard-Miller, and Marc Donald. We thank our past and present partners at the Yale Center for Emotional Intelligence, especially Marc Brackett, Susan Rivers, Bonnie Brown, and Dena Simmons. We are grateful to Mirellise Vasquez, Kim Hein, and the rest of the Tauck Family Foundation for funding this work and providing many other resources beyond financial supports.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Ackerman, C., and Skoog-Hoffman, A. (2020). School-Level Learnings from the Field: Insights in Establishing a Collaborative Research-Practice Partnership [Learning Series on Research-Practice Partnerships (RPPs). Brief 2]. Collaborative for Academic, Social, and Emotional Learning.

Bauer, M. S., Damschroder, L., Hagedorn, H., Smith, J., and Kilbourne, A. M. (2015). An introduction to implementation science for the non-specialist. *BMC Psychology* 3, 1–12. doi: 10.1186/s40359-015-0089-9

Bauer, M. S., and Kirchner, J. (2020). Implementation science: what is it and why should I care? *Psychiatry Res.* 283:112376. doi: 10.1016/j.psychres.2019.04.025

Blyth, D. A., Jones, S., and Borowski, T. (2018). SEL frameworks–what are they and why are they important. *Measur SEL Data Pract* 1, 1–9.

Brackett, M. A., Bailey, C. S., Hoffmann, J. D., and Simmons, D. N. (2019). RULER: a theory-driven, systemic approach to social, emotional, and academic learning. *Educ. Psychol.* 54, 144–161. doi: 10.1080/00461520.2019.1614447

Chambers, D. A. (2012). The interactive systems framework for dissemination and implementation: enhancing the opportunity for implementation science. *Am. J. Community Psychol.* 50, 282–284. doi: 10.1007/s10464-012-9528-4

Cipriano, C., Strambler, M. J., Naples, L. H., Ha, C., Kirk, M., Wood, M., et al. (2022). The state of evidence for social and emotional learning: A contemporary meta-analysis of universal school-based SEL interventions. Accessed December 2022.

Coburn, C. E., Penuel, W. R., and Geil, K. E. (2013). *Research practice partnerships: A strategy for leveraging research for educational improvement in school districts*. New York: William T. Grant Foundation

Connecticut Data Collaborative. (2023). Demographics and Population Data. Available at: https://www.ctdata.org/demographics (Accessed March 21, 2023).

Connecticut State Department of Education (CSDE). (2023). EdSight Interactive Data Portal: Enrollment Dashboard. Available at: https://public-edsight.ct.gov/Students/Enrollment-Dashboard (Accessed March 21, 2023).

Denner, J., Bean, S., Campe, S., Martinez, J., and Torres, D. (2019). Negotiating trust, power, and culture in a research practice partnership. *AERA Open* 5, 233285841985863–233285841985811. doi: 10.1177/2332858419858635

Durlak, J. A. (2016). Programme implementation in social and emotional learning: basic issues and research findings. *Camb. J. Educ.* 46, 333–345. doi: 10.1080/0305764X.2016.1142504

Durlak, J. A., and DuPre, E. P. (2008). Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am. J. Community Psychol.* 41, 327–350. doi: 10.1007/s10464-008-9165-0

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

Eccles, M. P., and Mittman, B. S. (2006). Welcome to implementation science. *Implement. Sci.* 1. doi: 10.1186/1748-5908-1-1

Farrell, C. C., Penuel, W. R., Coburn, C. E., Daniel, J., and Steup, L. (2021). Research-Practice Partnerships in Education: The State of the Field. William T. Grant Foundation.

Herrenkohl, T. I., Jones, T. M., Lea, C. H. III, and Malorni, A. (2020). Leading with data: using an impact-driven research consortium model for the advancement of social emotional learning in schools. *Am. J. Orthop.* 90, 283–287. doi: 10.1037/ort0000435

Meyers, D. C., Durlak, J. A., and Wandersman, A. (2012). The quality implementation framework: a synthesis of critical steps in the implementation process. *Am. J. Community Psychol.* 50, 462–480. doi: 10.1007/s10464-012-9522-x

Nowell, L. S., Norris, J. M., White, D. E., and Moules, N. J. (2017). Thematic analysis: striving to meet the trustworthiness criteria. *Int J Qual Methods* 16, 1–13. doi: 10.1177/1609406917733847

Oberle, E., Domitrovich, C. E., Meyers, D. C., and Weissberg, R. P. (2016). Establishing systemic social and emotional learning approaches in schools: a framework for schoolwide implementation. *Camb. J. Educ.* 46, 277–297. doi: 10.1080/0305764X.2015.1125450

Phelps, D. (2019). The challenges of bridging the research–practice gap through insider–outsider partnerships in education. *Teach. Coll. Rec.* 121, 1–28. doi: 10.1177/016146811912101202

Schoonenboom, J., and Johnson, R. B. (2017). How to construct a mixed methods research design. *Kolner Zeitschrift fur Soziologie und Sozialpsychologie* 69, 107–131. doi: 10.1007/s11577-017-0454-1

Strambler, M. J., and Meyer, J. L. (2018). Bridgeport SEL Evaluation. Open Science Framework. Available at: https://doi.org/10.17605/osf.io/nwzrs (Accessed March 21, 2023).

Wandersman, A., Duffy, J., Flaspohler, P., Noonan, R., Lubell, K., Stillman, L., et al. (2008). Bridging the gap between prevention research and practice: the interactive systems framework for dissemination and implementation. *Am. J. Community Psychol.* 41, 171–181. doi: 10.1007/s10464-008-9174-z

Yamashiro, K., Wentworth, L., and Kim, M. (2023). Politics at the boundary: exploring politics in education research-practice partnerships. *Educ. Policy* 37, 3–30. doi: 10.1177/08959048221134916

Yeager, D., Bryk, A., Muhich, J., Hausman, H., and Morales, L. (2013). *Practical measurement*. Palo Alto, CA: Carnegie Foundation for the Advancement of Teaching, 78712

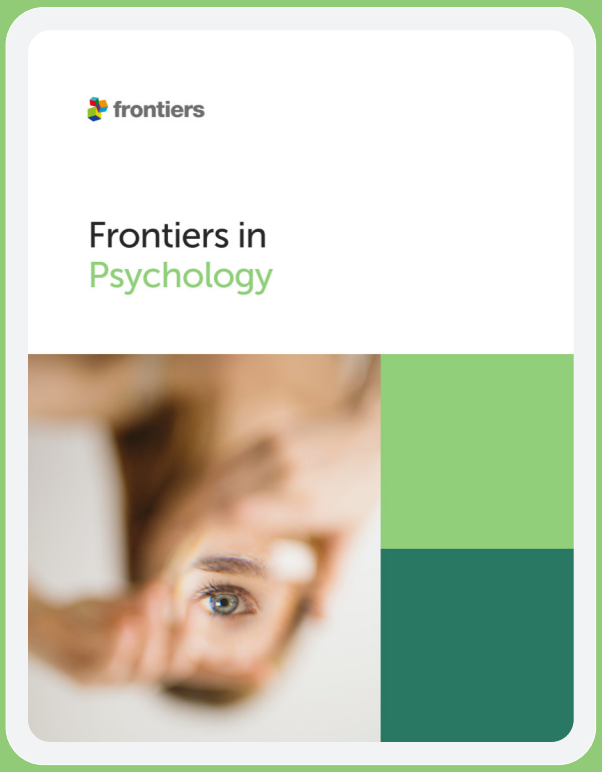
Frontiers in Psychology

Paving the way for a greater understanding of human behavior

The most cited journal in its field, exploring psychological sciences - from clinical research to cognitive science, from imaging studies to human factors, and from animal cognition to social psychology.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact