Title
Gender, racial-ethnic, and socioeconomic disparities in the development of social-emotional competence among elementary school students

Authors
Lee, Juyeon
Shapiro, Valerie B
Robitaille, Jennifer L
et al.

Publication Date
2024-06-01

DOI
10.1016/j.jsp.2024.101311

Peer reviewed

Citation: Lee, J., Shapiro, V. B., Robitaille, J. L., & LeBuffe, P. A. (2024). Gender, racial-ethnic, and socioeconomic disparities in the development of social-emotional competence among elementary school students. Journal of School Psychology, 104. 101311.

https://doi.org/10.1016/j.jsp.2024.101311

Gender, Racial-Ethnic, and Socioeconomic Disparities in the Development of Social-Emotional Competence Among Elementary School Students

Abstract

Social-emotional competence (SEC) has been demonstrated to be a crucial factor for student mental health and is malleable through the high-quality implementation of effective school-based social and emotional learning (SEL) programs. SEL is now widely practiced in the United States as a Tier 1 strategy for the entire student body, yet it remains unclear whether disparities exist in the development of SEC across socioculturally classified subgroups of students. Also, despite the field's widespread concern about teacher bias in assessing SEC within diverse student bodies, little evidence is available on the measurement invariance of the SEC assessment tools used to explore and facilitate SEC development. Based on a sociocultural view of student SEC development, this study aimed to measure and examine the extent to which gender, racial-ethnic, and socioeconomic disparities exist in SEC developmental trajectories during elementary school years. Specifically, using 3 years of SEC assessment data collected from a districtwide SEL initiative ($N = 5452$; Grades K–2 at baseline; nine measurement occasions), this study (a) tested the measurement invariance of a widely-used, teacher-rated SEC assessment tool (DESSA-Mini) across student gender, race and ethnicity, and socioeconomic status (SES); and (b) examined the extent to which multiyear SEC growth trajectories differed across these subgroups under a routine SEL practice condition. The invariance testing results supported strict factorial invariance of the DESSA-Mini across all the examined subgroups, thereby providing a foundation for valid cross-group comparisons of student SEC growth. The piecewise latent growth modeling results indicated that boys (vs. girls), Black students (vs. White students), Hispanic students (vs. White students), and low-income students (vs. middle-to-high-income students) started with a lower level of SEC, with these gaps being sustained or slightly widened throughout 3 elementary school years. Based on these findings, this study calls for future research that can inform practice efforts to ensure equitable SEC assessments and produce more equitable SEL outcomes, thereby promoting equity in school mental health.

Keywords: Social-Emotional Competence; Disparities; Growth Trajectory; Social and Emotional Learning (SEL); Assessment; Measurement Invariance

Introduction

Social-emotional competence (SEC) has been defined as the knowledge, attitudes, and skills necessary to know and manage the self, understand and relate to others, and make responsible personal and social decisions (Elias et al., 1997; Weissberg et al., 2015). Children's SEC is a reliable protective and promotive factor of mental health and many other positive developmental outcomes, including behavioral health and educational success (Denham, 2006; Domitrovich et al., 2017; Zins et al., 2004). The development of SEC occurs at various developmental stages and through various socialization contexts. In infancy and toddlerhood, children typically develop a sense of self and others, emotional and behavioral control, prosocial emotion and behavior, and an understanding of social conventions and norms (Shulman, 2016; Wittmer et al., 1996). In this early stage of life, the family environment plays a critical role in children's social and emotional development (Shulman, 2016). In middle childhood, school becomes another important socialization context. School entry (i.e., starting kindergarten in the US context) provides new opportunities and challenges for children to develop SEC as they enter a social world distinct from earlier socialization experiences and spend significant time there (Collins, 1984). Empirical evidence suggests that individual children enter school with different levels of SEC and the variability in SEC is not random across diverse student subgroups (Raver & Knitzer, 2002; Yates et al., 2008).

Gender, race and ethnicity, and socioeconomic status (SES) have been conceptualized as important factors associated with individual variability in SEC that manifest in student behaviors in the school context (Garner et al., 2014). These characteristics are attributed to individuals, but the ways in which these developmental differences in SEC emerge involve the influence of various environmental and structural factors differentially experienced by diverse subgroups. Using a sociocultural perspective, Garner and colleagues (2014) provided a heuristic model depicting the interrelations between individual characteristics and social-emotional practices in family, community, and school contexts as related to the development of student SEC. In explaining how gender, race and ethnicity, and SES can each be linked to SEC development, we allude to at least two mechanisms: one related to differential socialization processes and the other related to structural inequalities.

A first explanation (which we call *differential socialization*) postulates that children of diverse genders, races, ethnicities, and socioeconomic classes may experience different socialization processes, oriented to the emotional and behavioral norms and expectations specific to the subgroup, which may or may not be aligned with the dominant norms and expectations of schools (Garner et al., 2014). In this view, some students' behaviors, manifested in ways consistent with the sociocultural norms and expectations of the subgroup, may be inconsistent with the norms and expectations of the school system. In the broader literature, this has been conceptualized and examined as cultural mismatch (e.g., Stephens & Townsend, 2015; Villegas, 1988), cultural misalignment (e.g., Boykin et al., 2005), cultural capital mismatch (e.g., Kozlowski, 2015), or home-school dissonance (e.g., Arunkumar et al., 1999; Kumar, 2006), with empirical evidence supporting its presence and effects, especially among racially and ethnically minoritized students. In the American education system, which is emergent from White, middle-class values and norms (Boykin et al., 2005; Leonardo, 2012) and is sustained by a mostly White, female, professionalized workforce (Institute of Education Sciences, 2020), certain behaviors that are more frequently shown among male students, students of color, and lower-class students could be viewed as less socially and emotionally competent than those of female, White, and upper-class students. To be clear, this explanation would not imply actual deficits in SEC among these subgroups assessed to have lower SEC, but rather a cultural misalignment between the behavioral expectations of diverse contexts, whereas strengths that are adaptive or celebrated in some contexts are overlooked or undervalued in schools, and a lack of conformance to the explicit or implicit social and emotional expectations of adults in schools is understood as a general lack of SEC.

A second explanation (which we call *structural inequalities*) postulates that children of diverse races, ethnicities, and socioeconomic classes may experience different sets of opportunities and barriers that differentially promote or hinder their SEC development due to structural inequalities (Garner et al., 2014). This view aligns with a sociological theoretical strand emphasizing structural forces contributing to chronic hardship among disadvantaged populations (Wilson, 2009). Given the history and present realities of racism and classism in the

United States that have created different developmental environments for White children versus children of color, and for children from upper versus lower socioeconomic classes (e.g., parents' working conditions and neighborhood environments; Dixon-Román, 2017; Lareau, 2011; Leonardo, 2004), students from marginalized backgrounds may be deprived of various opportunities and resources to develop SEC (e.g., high-quality childcare, extracurricular or out-of-school activities, child-friendly spaces in neighborhoods; Bennett et al., 2012; Bruner, 2017; Chin & Phillips, 2004). Also, the literature suggests that, on average, students of color and students from lower socioeconomic classes are at greater risk of encountering implicit bias, microaggressions, overt discrimination, and victimization at schools or other public spaces, which all have adverse effects on their social-emotional development (Sarcedo et al., 2015; Sue, 2010; Zoric, 2014). Of course, many students who experience adversity and disadvantage are resilient; some scholars even describe post-traumatic growth and navigating disadvantage as a means to developing social and emotional strengths (Tedeschi & Calhoun, 2004). Yet, when adversities (e.g., parent incarceration) and protective factors (e.g., a stable relationship with a trusted adult) are not distributed equally in society, we may expect that structural inequalities induce population-level differences in developmental outcomes, whereas racially, ethnically, and socioeconomically marginalized students may consequently, on average, have lower SEC.

These two potential mechanisms (i.e., differential socialization and structural inequalities), which are not deterministic on an individual level or mutually exclusive of each other in their explanatory power, provide a useful conceptual framework to understand how gender, racial-ethnic, and socioeconomic disparities in student SEC may exist in the United States. Conversely, as reviewed in the following section, empirical evidence on disparities in SEC among elementary school students has been limited and inconsistent, especially regarding whether and to what extent racial-ethnic and socioeconomic disparities manifest. Furthermore, little evidence is available to inform whether any observed SEC disparities are likely to increase over time (e.g., Matthew effects leading to growing inequalities; Merton, 1968), decrease over time (e.g., compensation effects leading to reduced inequalities; Baumert et al., 2012), or remain consistent. Using this sociocultural view on student SEC development as a motivation and guiding framework for interpretation and discussion, the goal of this study was to understand the extent to which gender, racial-ethnic, and socioeconomic disparities in student SEC trajectories exist and develop throughout the elementary school years in the United States. This study sets the stage for research on understanding and reducing disparities in student SEC to promote equity in school mental health (Kilbourne et al., 2006).

Existing Empirical Evidence on Disparities in Social-Emotional Competence Development

The literature lacks studies that investigate subgroup disparities in SEC developmental trajectories among school-aged children, particularly when compared to ample evidence on disparities in overall mental, emotional, and behavioral health problems (e.g., Alegría et al., 2015). Existing research on subgroup differences in social and emotional development has mainly focused on gender differences (Eisenberg et al., 2006; Maguire et al., 2016; Nakajima et al., 2020; Zahn-Waxler & Smith, 1992; Zimmermann & Iwanski, 2014). Evidence is scarce as to differences associated with other sociocultural subgroups and how these subgroup differences unfold over the school years.

The literature on school-based social and emotional learning (SEL) programs may provide insights into these issues. SEL refers to the process of acquiring and developing SEC (Weissberg et al., 2015). An SEL program evaluation typically involves repeated measures of SEC, the primary target outcome of many SEL programs, within or across school years, and among entire student populations that may consist of diverse student subgroups. Despite a growing body of SEL research, many evaluation studies have not described their samples regarding student gender, race and ethnicity, or SES. A review of the 117 peer-reviewed articles included in Durlak et al.'s (2011) influential meta-analysis of SEL program effectiveness revealed that 31% did not provide sample descriptions for gender, 36% for race and ethnicity, and 55% for SES (Rowe & Trickett, 2018). Even in the most updated SEL meta-analysis that reviewed 424 studies from 2008 to 2020, 18% of the reviewed studies did not provide sample descriptions for gender, 49% for race and ethnicity, and 59% for SES (Cipriano et al., 2023).

In the available literature, including both observational and intervention research, the most reported subgroup difference in SEC involves gender. Kindergarten and elementary school boys have been observed to have lower SEC than girls across diverse measurement methods such as teacher ratings (Aber et al., 2003; Bierman et al., 2010; Frey et al., 2005; Hutchison et al., 2020; Jones S. M. et al., 2011; Krishnan, 2011; Lee, Shapiro, & Kim, 2023; Maguire et al., 2016; Malti et al., 2011), caregiver ratings (Nakajima et al., 2020), self-reports (Holsen et al., 2009; Malti et al., 2011; Washburn et al., 2021; West et al., 2020), peer nominations (Bierman et al., 2010), and task-based assessments (Colle & Del Giudice, 2011). Yet, findings about gender differences in the rate of SEC growth are quite rare and mixed. In the context of SEL research, boys showed less growth than girls through teacher ratings of prosocial behavior in one study (Aber et al., 2003). Still, they had a similar rate of growth to girls in teacher ratings of SEC in other studies (Jones S. M. et al., 2011; Lee, Shapiro, & Kim, 2023).

Evidence regarding racial-ethnic differences in SEC is more limited in volume and more inconclusive. In some large-scale observational studies of elementary through high school students, students of color (e.g., Asian, Black, Latinx) rated themselves, on average, to have lower SEC than White peers (Jones T. M. et al., 2020; West et al., 2020). Some SEL intervention studies using teacher ratings also indicated that Black/African American students had lower SEC scores at baseline compared to White students (Aber et al., 2003; Jones S. M. et al., 2011; Lee, Shapiro, & Kim, 2023) or to non-Black students (Elias & Haynes, 2008), and Hispanic students had lower baseline SEC than non-Hispanic White students (Aber et al., 2003). Other studies found that teacher ratings of SEC did not differ by race (Chain et al., 2017) or Hispanic origin (Jones S. M. et al., 2011; Lee, Shapiro, & Kim, 2023). Regarding racial-ethnic differences in the rate of SEC growth in the context of SEL research, some studies have found that Black students and Hispanic students each had non-differential growth rates compared to White or other students (Elias & Haynes, 2008; Jones S. M. et al., 2011; Lee, Shapiro, & Kim, 2023), whereas Aber et al. (2003) indicated that Black students showed a slower increase in growth than White peers as assessed through teacher ratings of prosocial behavior.

Findings on the differences in SEC by family SES are also limited in volume and inconsistent in their findings. Students from lower SES backgrounds had lower SEC than others in a large-scale self-report survey (West et al., 2020). Some SEL studies have similarly reported lower levels of SEC among students of lower SES as assessed by teacher ratings (e.g., Aber et al., 2003) or child self-reports (e.g., Holsen et al., 2009), whereas others found no baseline differences in SEC using teacher ratings (Jones S. M. et al., 2011; Lee, Shapiro, & Kim, 2023). In a study that used a multi-modal assessment strategy, Malti et al. (2011) reported that teacher ratings indicated lower prosocial behavior among lower SES students, but no such differences were observed with student self-reports. Specific to socioeconomic subgroup differences in SEC growth rates, available evidence suggests that the rate of SEC growth trajectories, as rated by teachers, did not differ by student SES (e.g., Aber et al., 2003; Jones S. M. et al., 2011; Lee, Shapiro, & Kim, 2023).

In summary, evidence is inconclusive as to the extent to which gender, racial-ethnic, and socioeconomic disparities in student SEC development exist and change over time. Yet, one could tentatively say that (a) more consistent findings have been reported on gender differences in the overall level of SEC, and (b) when significant subgroup differences were observed, the results seem to suggest lower levels of SEC for boys, students of color, and low-SES students.

However, these observed differences can only be interpreted as real subgroup disparities in SEC if one assumes that the measures of SEC are not biased in their assessment of specific subgroups. For example, some scholars have questioned whether the gender differences in prosocial behavior found in the literature may be due, in part, to biases in measures (e.g., consisting of more feminine than masculine items; Eisenberg et al., 2006; Zahn-Waxler & Smith, 1992; Zarbatany et al., 1985). In fact, a recent SEC measurement study revealed that some of the items (mostly related to emotional awareness and management) in a self-reported SEC assessment tool persistently showed differential item functioning by gender, race, and ethnicity across 2 survey years (Crowder et al., 2019). However, little evidence is available regarding the subgroup comparability of the measures used in the studies reviewed above regarding subgroup differences in SEC. Below, we further describe the issue of measurement bias and the importance of testing measurement invariance, particularly for teacher ratings of student SEC. Along with more evidence relating to measurement invariance, the field needs more studies describing the presence and extent of subgroup disparities in student SEC that will ultimately drive questions about the mechanisms underlying any existing disparities and the impact of SEL programs on these gaps.

Using Teacher-completed Behavioral Rating Scales: The Importance of Testing Measurement Invariance

We have thus far introduced two potential mechanisms of the hypothesized gender, racial-ethnic, and socioeconomic disparities in student SEC (i.e., differential socialization and structural inequalities). A third explanation might be the *systematic bias* of the rater. In other words, any observed differences in ratings of student SEC may reflect measurement bias rather than the existence of real disparities in SEC. Measurement bias can have different meanings in various contexts, but the present study focused on a statistical phenomenon where observed scores are systematically influenced by factors other than the actual amount of the construct being measured. With respect to cross-group comparisons, a measure is said to be *biased* if it systematically functions in a different way across subgroups, making an accurate assessment and comparison impossible (Millsap & Everson, 1993). Thus, before examining subgroup disparities in SEC, it is crucial to make sure that the SEC measurement tool in use is comparably valid (i.e., measures the same construct) across subgroups (Meredith & Teresi, 2006; Putnick & Bornstein, 2016).

With increasing and widespread adoption of school-based SEL initiatives across the United States, researchers and practitioners have described the need for SEL assessment tools that can be equitably applied across diverse student populations (Assessment Work Group, 2019; Garner et al., 2014; Mahoney et al., 2022). In elementary education, teacher-completed behavioral rating scales have become one of the most used tools for universal screening and progress monitoring in the domain of SEL (Shapiro et al., 2024). Although no assessment method is free from the risk of measurement bias, many have questioned the cross-group validity of teacher ratings given the persistent history of the overrepresentation of marginalized students in school discipline and special education programs, which often is predicated on teacher referrals (Peters et al., 2014; Skiba et al., 2011; Zhang & Katsiyannis, 2002). A body of literature has reported the existence of teacher bias in reporting student problem behaviors, resulting in discriminatory disciplinary practices against marginalized students, particularly based on their race and ethnicity (Bradshaw et al., 2010; Gregory & Roberts, 2017; Skiba et al., 2002, 2011; Ura & d'Abreu, 2022). With a growing critique of SEL as being colorblind and reflecting White, middle-class values (Gregory & Fergus, 2017; Hoffman, 2009; Mahfouz & Anthony-Stevens, 2020), there is an emerging consensus that we must use SEL measures that are relevant and fair across diverse student populations to avoid reinforcement or reproduction of structural inequalities (Assessment Work Group, 2019).

The potential risk of measurement bias associated with assessing diverse student subgroups with teacher behavioral ratings establishes the need to empirically examine the extent to which such bias exists. Alternatives to the empirical interrogation of bias, such as just assuming that the measure is unbiased (and therefore using it recklessly) or biased (and advocating against accurately understanding the presence of disparities and addressing them) are each unacceptable. A test of measurement invariance (i.e., measurement equivalence) across student subgroups has been proposed as one way to quantitatively explore measurement bias in SEL assessments (Assessment Work Group, 2019; Gehlbach & Hough, 2018). *Measurement invariance* means that the relation between observed scores and a latent construct is the same across different groups, whereas *measurement non-invariance* means that the construct measured has a different structure or meaning as applied to different groups (Putnick & Bornstein, 2016). Although evidence of measurement invariance does not guarantee that a measure is bias-free, it is one of the initial steps to ensure that it is relevant and fairly applied across diverse subgroups (Pendergast et al., 2017).

Research has demonstrated that measurement non-invariance can lead to erroneous conclusions about cross-group comparisons (e.g., gender, racial-ethnic, socioeconomic differences) in growth trajectories, both in understanding the initial level of a construct and its rate of growth (Chen, 2008; Kim & Willson, 2014). Despite this risk, the measurement invariance testing of SEC assessment tools across these sociocultural subgroups has only recently emerged. Various self-report instruments have demonstrated invariance across gender, race and ethnicity, or SES (e.g., Anderson-Butcher et al., 2016; Anthony et al., 2021; Basting et al., 2022; Carpendale et al., 2023; Crowder et al., 2019; Davis, 2020; Gehlbach & Hough, 2018; Jones T. M. et al., 2020; Mantz et al., 2018; McKown, 2019; von der Embse et al., 2017). Relatively less evidence is available on the measurement invariance of teacher-completed SEC measures: however, the Student Risk Screening Scale (SRSS; Drummond, 1994) was found to be invariant between boys and girls (Fredrick et al., 2019) and the Social, Academic, and Emotional Behavior Risk Screener (SAEBRS; Kilgus et al., 2013) was also found to be invariant between boys and girls (von der Embse et al., 2019) and between Black and White students (Pendergast et al., 2017). Although these studies provide some promising findings, more work is needed to establish measurement invariance of teacher-rated SEC measures across diverse student subgroups and ultimately to advance our knowledge base on gender, racial-ethnic, and socioeconomic disparities in student SEC development.

Research Objectives

The objectives of this study were twofold: (a) to test the measurement invariance of a widely used teacher-completed behavioral rating scale for SEC (i.e., Devereux Student Strengths

Assessment Mini [DESSA-Mini]; Naglieri et al., 2011) across student gender, race and ethnicity, and SES; and (b) to examine the extent to which student SEC growth trajectories differ across these student subgroups during the elementary school years in the context of a district-wide SEL program implementation. Based on the reviewed literature, we hypothesized that gender, racial-ethnic, and socioeconomic disparities in student SEC would be observed at baseline by favoring girls, White students, and upper-class students, whereas no specific hypotheses were made regarding the disparities in the SEC growth rates.

Using available school administrative data, this study was primarily focused on comparisons between (a) boys and girls; (b) Black/African American, Hispanic/Latinx of any race, and non-Hispanic White students; and (c) low-income students eligible for free or reduced-price meals and middle-to-high-income students not eligible for free or reduced-price meals. We recognize that these aspects of individual identities are socially constructed and the above categorizations may not accurately represent how students identify themselves, may exclude some students from analysis, and can obscure a great deal of within-group heterogeneity. Acknowledging these limitations, we used available data to illuminate similarities and differences in SEC development. To be clear, this study does not test differential intervention effects across subgroups in an experimental sense. Rather, it compares student SEC growth trajectories across subgroups under a single SEL implementation condition.

Method

Data and Sample

Data used in this study came from a district-wide SEL initiative that implemented the Promoting Alternative Thinking Strategies (PATHS; Kusché & Greenberg, 1994) curriculum in a large urban district for 3 academic years (i.e., 2011–2012, 2012-2013, and 2013–2014). This district served a racially and ethnically diverse, low-income student body in Pennsylvania with a high mobility rate. During years comprising the present study, the district's average elementary school enrollment was approximately 9000 students, 65% of whom were identified as Hispanic of any race, 15% non-Hispanic Black, and 13% non-Hispanic White, with approximately 86% of students receiving free or reduced-price meals and 42% of students transferring in and out of schools (Pennsylvania Department of Education, 2021). The PATHS curriculum is a classroom-based universal SEL program found to be effective in promoting social-emotional, behavioral, and academic outcomes among school-aged children in several randomized controlled trials (e.g., Conduct Problems Prevention Research Group, 1999; Fishbein et al., 2016; Greenberg et al., 1995; Schonfeld et al., 2015). The PATHS curriculum was implemented with K–2 students across 15 elementary schools in Year 1 (2011–2012) and with all K–5 students across these 15 elementary schools in Year 2 (2012–2013) and Year 3 (2013–2014), except for three schools that elected to discontinue implementing PATHS starting Year 2. Teachers were provided with initial and booster training sessions, along with ongoing support and classroom visits from coaches providing technical assistance. Program implementation in this larger project is described elsewhere, suggesting that the program was implemented as planned in Year 1 with higher-than-typical implementation supports provided in the following years (e.g., Lee, Shapiro, Robitaille, & LeBuffe, 2023; Shapiro et al., 2018, 2024). Study protocols were approved by the Institutional Review Board for Research, Ethics, and Human Rights at the Institute of Clinical Training and Research at the Devereux Foundation. De-identified data (no longer regarded as human subjects' data) was subsequently shared with the first author, whose analytic activities were acknowledged by the Committee on Protection of Human Subjects at the University of California, Berkeley.

The sample for the present study consisted of all students who were in Grades K–2 in

Year 1 (Grades 2–4 in Year 3; *N* = 5452) when PATHS implementation began. In this sample, grade and gender were quite evenly distributed: 34.37% Kindergartners, 33.07% first graders, and 32.56% second graders at baseline; and was 47.95% girls and 52.02% boys. The sample was representative of the district's student population in terms of the distribution of race, ethnicity, and SES (as indicated by free or reduced-price meals eligibility): 63.92% identified as Hispanic, 15.79% Black, 13.65% White, and 87.07% low-income. Most students (77.71%) were assessed exclusively while affiliated with PATHS-implementing schools, whereas the rest of the students (22.29%) were assessed at some point in Year 2 and Year 3 when they were affiliated with one of the schools that discontinued the implementation of PATHS starting Year 2. A small number of students (< 0.2%) with missing information on gender, race and ethnicity, or SES were excluded from the analysis. Table 1 shows the sociodemographic characteristics of the sample.

Measurement

Student SEC was repeatedly measured using the DESSA-Mini (Naglieri et al., 2011) three times per year for 3 sequential school years. Consistent with the way this tool is typically used in practice, Fall ratings were collected using the DESSA-Mini Form 1 in October, Winter ratings using Form 2 in January and February, and Spring ratings using Form 3 in June. As an eight item version of the full 72-item DESSA (LeBuffe et al., 2009), the DESSA-Mini is a brief, standardized, and norm-referenced behavioral rating scale that assesses K–8 students' overall level of SEC (Hwang et al., 2023; Naglieri et al., 2013). Teachers rate the frequency of a student's behavior observed during the past 4 weeks on a 5-point Likert scale (0 = Never, 1 = *Rarely*, 2 = *Occasionally*, 3 = *Frequently*, 4 = *Very frequently*). All items are strengths-based and indicate positive rather than maladaptive behaviors (Simmons et al., 2016). Example items include "Keep trying when unsuccessful", "Respect another person's opinion", "Do something nice for somebody", and "Show good judgment." In practice, item raw scores are transformed based on a single norms table into *T*-Scores based on a nationally representative standardization sample (*M* = 50, *SD* = 10) to assess individual students' relative social-emotional strengths and needs. *T*-scores of 60 and above are classified as *Strengths*, *T*-scores between 41 and 59 represent *Typical* SEC, and *T*-scores of 40 and below indicate a *Need for Instruction*.

The DESSA-Mini has demonstrated strong psychometric qualities, including reliability (e.g., internal consistency, test-retest reliability, inter-rater reliability, alternate form reliability) and validity (e.g., concurrent and predictive criterion validity) as evidenced in several studies (Lee et al., 2018; Shapiro, Accomazzo, & Robitaille, 2017; Shapiro, Kim, et al., 2017). In the present study, the internal consistency coefficient estimates (Cronbach's alpha) ranged from 0.95 to 0.98 across the nine measurement times; these values are slightly higher than those reported in the original development study (range = 0.919–0.924; Naglieri et al., 2011). In Lee, Shapiro, Robitaille, & LeBuffe (2023), the three DESSA-Mini forms indicated high alternate form reliability (correlation coefficients ranging from 0.95 to 0.98) and negligible cross-form mean differences (ES ranging from 0.01 to 0.08). Additionally, evidence of longitudinal measurement invariance of each form (i.e., strong factorial invariance across seasons and years) was demonstrated in Lee, Shapiro, Robitaille, & LeBuffe (2023) and replicated within the present study sample in which over 90% of the students were rated by the same teachers within a single year and over 96% were rated by different teachers across years.

Student sociodemographic characteristics were collected from an administrative database. In the original source of data, student gender was measured as a binary variable (girl or boy), race and ethnicity included seven possible categories as presented in Table 1, and SES was indicated by whether the student was eligible for free or reduced-price meals, which was determined as a function of family annual income and federal poverty guidelines (e.g., the annual income eligibility criteria in 2011–2012 was $41,348 for a household size of four). In the analyses that involved the race and ethnicity variable, we had to limit our sample to the three largest subgroups (i.e., Black, Hispanic, and White) that together comprised the majority (93.4%) of the sample because the inclusion of the other subgroups in measurement invariance testing, either each as a unique subgroup or all together as a fourth group, yielded model convergence problems that were likely due to small group sizes. When comparing subgroup differences in SEC growth trajectories using conditional growth trajectory modeling, student gender and SES were dummy coded in ways that girls and middle-to-high income subgroups, respectively, had the value of 0, whereas boys and low-income subgroups, respectively, had the value of 1.

Analysis Plan

Factorial Invariance Testing Across Subgroups.

To test the measurement invariance of the DESSA-Mini across student subgroups, we conducted multigroup confirmatory factor analysis (MG-CFA), which is a structural equation modeling approach to measurement invariance tests. MG-CFA is one of the most used approaches in educational and psychological research and has been demonstrated to have relative strengths compared to other methods of testing measurement invariance, such as item response theory (IRT) or multiple indicator multiple cause (MIMIC) modeling (see Pendergast et al., 2017, for a review). One of the more significant advantages of using MG-CFA within this study is that it can be directly incorporated into second-order latent growth modeling to estimate and predict SEC growth trajectories (Kim & Willson, 2014).

Using the MG-CFA approach, four levels of factorial invariance (i.e., configural invariance, weak factorial invariance, strong factorial invariance, and strict factorial invariance) were tested across student gender (i.e., girl vs. boy), race and ethnicity (i.e., Black vs. Hispanic vs. White), and SES (i.e., low-income vs. middle-to-high income). In MG-CFA, the *configural invariance* model specifies the identical configuration or structure of the factor-indicator relations across subgroups while placing only minimal constraints for model identification. Evidence of configural non-invariance suggests that the structural relations between each item and the SEC construct is different across subgroups and that no subgroup comparisons should be made. Configural invariance is a prerequisite for all the following invariance tests.

The *weak factorial* invariance model adds cross-group equality constraints on the factor loading for the same indicator variable (i.e., metric invariance) to the configural invariance model. This relates to the magnitude of the relation between each item and the SEC construct. Evidence of metric *non-invariance* suggests that teacher ratings of certain items better represent overall SEC for one group rather than another and the model should be refined. Configural and metric invariance is a prerequisite for further invariance tests. Without further invariance tests, the measure is considered to have weak factorial invariance, which does not allow valid comparisons of factor means across subgroups.

The *strong factorial invariance* model adds cross-group equality constraints on the intercept of the same indicator variables (i.e., scalar invariance) to the weak factorial invariance model. This relates to the intercept, or a mean of each item compared to the mean of the construct. Evidence of *scalar non-invariance* suggests that teachers rate certain items systematically lower or higher for one group than another, independent of students' true SEC level. Scalar invariance, in addition to configural and metric invariance, implies strong factorial invariance, allowing cross-group comparisons of factor means, which is a prerequisite condition for examining subgroup disparities in growth trajectories.

Finally, the *strict factorial invariance* model adds cross-group equality constraints on unique variance of the same indicator variables (i.e., residual invariance) to the strong factorial invariance model. This relates to the residual variance of each item, or the item-unique variance remained unexplained by the SEC construct. *Residual invariance*, in addition to configural, metric, and scalar invariance, implies strict or full factorial invariance. However, many scholars suggest that strict factorial invariance is unrealistic and unnecessary when interpreting factor mean differences because the residuals are not part of the latent construct (Little, 2013; Pendergast et al., 2017; Putnick & Bornstein, 2016).

The longitudinal confirmatory factor model we compared across subgroups is presented in Figure 1. This model consists of 24 indicator variables comprising three forms of the DESSA-Mini (Item1–Item8 in Form 1, Item9–Item16 in Form 2, and Item17–Item24 in Form 3). The eight indicator variables comprising each form are hypothesized to measure a single latent variable (SEC) at each of the nine measurement times (t1–t3 in Year 1, t4–t6 in Year 2, and t7–t9 in Year 3). The SEC latent variables were allowed to covary across time as indicated by double-headed, curved arrows in Figure 1. The residuals for the same indicator (e1–e24) were also allowed to covary across occasion, although this is not shown in Figure 1 for graphical simplicity. In this model, strong longitudinal factorial invariance was first tested and confirmed by constraining the factor loadings and intercepts of the same indicator variable to be equal across occasions.

The invariance tests involved model fit comparisons of a series of nested models with increasing restrictions. Several model fit indices were compared to examine the degree to which these equality assumptions held with data. The chi-square difference test based on the likelihood ratio test statistics was used to test the statistical significance of the model fit differences.

Because the chi-square statistic is very sensitive to sample size (Bollen, 1989), alternative fit indices were also used, including the comparative fit index (CFI; Bentler, 1990), the Tucker-Lewis index (TLI; Tucker & Lewis, 1973), and the root mean squared error of approximation (RMSEA; Steiger, 1990). Given the large sample size of the present study, when the chi-square difference test suggested a significantly worse fit for a more restrictive model, but no appreciable change was found in alternative fit indices (e.g., ΔCFI smaller than or equal to $|0.01|$; Cheung & Rensvold, 2002), the more restrictive model was assessed as acceptable.

Growth Trajectory Modeling

After confirming the possibility of comparing latent means across subgroups with at least strong factorial invariance (as presented later in the Results), the growth trajectories of student SEC were estimated. Based on previous findings about the nature of changes in student DESSA-Mini scores under the same intervention condition (i.e., within-year growth, decrease over the summer, and decelerated growth rates across years; Lee, Shapiro, Robitaille, & LeBuffe, 2023), we used piecewise second-order latent growth modeling to incorporate separate growth profiles for each academic year, while simultaneously fitting a measurement model. Figure 2 illustrates the model that examined subgroup differences in student SEC growth trajectories.

The first-order component of this model is identical to the measurement model shown in Figure 1, assuming metric and scalar invariance across time and with residuals for the same indicators allowed to covary across time. The second-order component of the growth model consists of four growth factors, the intercept (baseline level), and three separate slopes (growth rates) corresponding to each academic year (Slope_{y1} to Slope_{y3}). Instead of assuming a linear or other predetermined shape of the growth curve within each year, the coefficients in the middle of the corresponding year's slope factor were freely estimated, whereas the first and third coefficients of the corresponding year's slope factor were fixed at 0 and 1, respectively. To incorporate prior findings on summer slides to be about a half of the previous year's growth (Lee, Shapiro, Robitaille, & LeBuffe, 2023), the 4th–9th coefficients of the Year 1 slope and the 7th–9th of the Year 2 slope were fixed at 0.5. The full specification of the coefficients of each growth factor are presented in Appendix Table A1. In accordance with the standardization identification method suggested by Yang et al. (2021), the mean of the intercept factor was fixed at 0 and the mean of each slope factor indicated the mean change scores between Fall and Spring within the corresponding year relative to the standard deviation of SEC scores at baseline.

The fit of the unconditional growth model was first assessed before including any covariates of the SEC growth trajectory, based on the chi-square statistic and the CFI, TLI, and RMSEA coefficients. Given the sample size ($> N = 250$) and model complexity (involving 30 or more indicator variables), a CFI > .90, TLI > .90, and RMSEA < .07 indicated an acceptable fit. Significant p-values were also expected for the chi-square test (Hair et al., 2010, p. 20). After confirming the goodness of fit of the unconditional growth model, student gender, race and ethnicity, and SES were included as a set of predictors of the intercept and three distinct slope factors, along with two other covariates: student grade level (0 = Grade K, 1 = Grade 1, 2 = Grade 2) and the exposure to PATHS implementation (0 = discontinued in Year 2 and Year 3, 1 = continued for 3 years). This analysis was conducted with the full sample and the subsample that included only three racial-ethnic subgroups (i.e., Black, Hispanic, and White) as used in the tests of factorial invariance by race and ethnicity. Although no notable difference was found in the model fit indices and parameter estimates between the two samples, we present and interpret the findings with the subsample including only three racial-ethnic subgroups, with which factorial invariance has been tested and established. This study analyzed the data primarily at the individual level due to the methodological challenges to multilevel modeling related to changing raters over time when there is only one rating at each occasion (Koch et al., 2020). All main analyses were conducted using Mplus version 8 (Muthén & Muthén, 1998–2017). The Full Information Maximum Likelihood (FIML) estimator with robust standard errors (MLR in Mplus) was used to handle missingness pragmatically and to yield robust results against any violations to normality and independence assumptions (Allison, 2003) based on the literature suggesting that Likert-type items possessing five or more levels could be reasonably treated as a continuous variable (Kline, 2015; Pendergast et al., 2017). With the use of MLR, the chi-square difference test statistics used for invariance testing were calculated by employing the Satorra-Bentler scaling correction (Satorra & Bentler, 2010) as suggested by the Mplus developers (see http://www.statmodel.com/chidiff.shtml).

Results

Descriptive Statistics

The average of eight DESSA-Mini items comprising each form increased, on average, by approximately 0.2 raw score points from Year 1 Fall to Year 3 Spring on a 5-point Likert scale from 0 (*Never*) to 4 (*Very Frequently*). When transformed into *T*-scores based on the national norm (M = 50, SD = 10), this translates into about a 4 *T*-score point increase throughout the 3 years. In Year 1 and Year 2, the mean SEC scores increased primarily between Fall and Winter of each year. In Year 3, more increases occurred between Winter and Spring. There was a slight decrease in mean SEC scores during the Year 1 summer period (by 2.7 *T*-score points) and less decrease during the Year 2 summer (by 1.6 *T*-score points). The distribution of the SEC scores was slightly left-skewed and platykurtic across all measurement time points. At the composite score level and the raw item level, no serious violations to univariate normality were identified as the absolute values of skewness and kurtosis were < 1 for all items, which meets the customary thresholds for skewness and kurtosis (i.e., 3 and 8, respectively; Kline, 2015; see Table 2).

There was a substantial amount of missing data across time. Missing data pattern analysis revealed that 686 students (12.58%) had complete data for all nine occasions, 784 (14.38%) were missing for one of the nine occasions, and 842 (15.44%) were missing for two of the nine occasions. About 9% of the students had some valid data for Year 1 but had missing data for all Year 2 and Year 3 ratings. Approximately 14% of the students had some valid data for Year 1 and Year 2 but had missing data for Year 3. Given the high student mobility rate within the district, it is likely that many of these cases represented students transferring out of the district during the study years. Conversely, about 16% of the students had missing data for Year 1 but had some valid data for Year 2 and Year 3. About 7.56% of the students had missing data for all Year 1 and Year 2 ratings but had some valid data in Year 3. It is likely that many of these cases represented students transferring into the district during the study years. Table 3 presents the distribution of missing data patterns.

Additional tests revealed that students who had complete data tended to have higher SEC at baseline than those who had some missing data for later ratings ($t = 4.22$, $p < .001$), although the difference was negligible in size (Hedges' $g = .18$). The proportions of students having any missing data did not differ significantly by student gender or SES, but differed by student race and ethnicity; compared to those who had complete data, there were slightly fewer Hispanic (by 3 percentage points) or White students (by 2 percentage points) and slightly more Black (by 3 percentage points) or multi-racial or other race/ethnicity students (by 2 percentage points) among those who had missing data ($\chi^2(3) = 8.80$, $p = .032$), but the difference was negligible in size (Cramer's V = 0.04). We used the FIML approach to missing data when estimating the rate of growth while accounting for the baseline level differences and including race and ethnicity as one of the major predictors.

Factorial Invariance Across Subgroups Testing Results

A series of model fit comparisons with increasing invariance assumptions were conducted to examine the factorial invariance of the DESSA-Mini by student gender, race and ethnicity, and SES. Across grouping variables, although the chi-square difference tests were statistically significant, inappreciable changes were observed in practical fit indices between the configural invariance model and the weak factorial invariance model (ΔCFI < |0.01|, ΔTLI < |0.01|, ΔRMSEA < |0.01|), between the weak factorial invariance model and the strong factorial invariance model (ΔCFI < |0.01|, ΔTLI < |0.01|, ΔRMSEA < |0.01|), and between the strong factorial invariance model and the strict factorial invariance model (ΔCFI ≤ |0.01|, ΔTLI < |0.01|, ΔRMSEA ≤ |0.01|). Although strict factorial invariance is not necessary for cross-group mean comparisons, these findings suggest that strict factorial invariance (i.e., metric, scalar, and residual invariance) can be assumed across gender, racial-ethnic, and socioeconomic subgroups, indicating that the entire factor model including factor loadings, intercepts, and residual variances was equal across subgroups. The strict factorial invariance model showed a good fit with each DESSA-Mini Form (CFI ≥ 0.93, TLI ≥ 0.93, RMSEA ≤ 0.04). The full invariance testing results are presented in Table 4.

SEC Growth Trajectory Modeling

After confirming factorial invariance across subgroups, the unconditional second-order piecewise latent growth model was first fitted to estimate the average SEC growth trajectory with three distinct growth curves for each year. Consistent with prior research (Lee, Shapiro, Robitaille, & LeBuffe, 2023), the model that freely estimated the slope coefficients at the middle of each year while assuming a decrease over the summer by a half of the previous year's gain showed a good fit with the data, $\chi^2(2270) = 17228.91$, $p < .001$, CFI = 0.94, TLI = 0.93, RMSEA = 0.04, 90% CI [0.034, 0.035]. This model seemed to produce the most reliable estimates with comparably good fit indices compared to a few competing models posing different assumptions about the summer loss (e.g., models assuming no decrease, different rates of decreases, decreases only for the first summer). The full specification and estimation results are presented in Appendix Table A1.

To examine the extent to which student SEC growth trajectories differed across student subgroups, we included gender, race and ethnicity, and SES as a set of predictors of the growth model examined above along with student grade level and the PATHS implementation. This conditional second-order piecewise latent growth model showed a good fit, $\chi^2(2678) = 17687.19$, $p < .001$, CFI = 0.94, TLI = 0.93, RMSEA = 0.03, 90% CI [0.033, 0.034]. Table 5 presents the unstandardized regression coefficients on the intercept and three slope factors.

At baseline, results indicated significant gender, racial-ethnic, and socioeconomic differences in SEC. Boy (vs. girl; $b = -0.43$, $p < .001$), Black (vs. White; $b = -0.17$, $p = .002$), Hispanic (vs. White; $b = -0.14$, $p = .001$), and low-income (vs. middle-to-high income; $b = -0.15$, $p = .001$) students each had lower mean intercepts in their SEC growth trajectories. Boys had marginally lower mean rates of growth than girls throughout the 3 years (Year 1: $b = -0.05$, $p < .10$; Year 2: $b = -0.05$, $p = .50$; Year 3: $b = -0.05$, $p < .10$). A difference in the mean growth rate by race and ethnicity was only found between Black and White students in Year 1. Black students showed a lower mean rate of growth than White students ($b = -0.11$, $p < .05$) in Year 1. Student SES was not significantly associated with the growth rates throughout the study years.

Figures 3a–3c indicate the estimated marginal means of the SEC factor over time across gender, racial-ethnic, and socioeconomic subgroups.

Student grade level and the exposure to continued PATHS implementation were also related to the student SEC growth trajectory. Students in higher grades had a higher mean baseline ($b = 0.18$, $p < .001$), but their mean rates of growth in the first 2 years were slower than lower-grade students (Year 1: $b = -0.17$, p < .001; Year 2: $b = -0.10$, $p < .001$). The Year 3 mean growth rate did not differ by grade levels. Students who were exposed to continued PATHS implementation for 3 years (vs. discontinued in Year 2 and Year 3) showed a larger mean growth in Year 2 ($b = 0.16$, $p < .001$) and Year 3 ($b = 0.17$, $p < .001$), whereas they had a higher mean baseline ($b = 0.16$, $p < .001$) and similar mean rate of growth in Year 1. Altogether, this conditional piecewise latent growth model explained 14% of the baseline SEC level, 19% of the Year 1 growth rate, 6% of the Year 2 growth rate, and 3% of the Year 3 growth rate.

Discussion

The goal of the present study was to contribute to advancing research on disparities in student mental health by measuring and identifying any gender, racial-ethnic, and socioeconomic disparities in SEC developmental trajectories during the elementary school years. Specifically, this study first tested measurement invariance of scores from a widely used teacher-completed SEC rating scale across student gender, race and ethnicity, and SES and then examined the extent to which students' multiyear SEC growth trajectories differed across these subgroups under a routine, district-wide SEL implementation condition. The invariance testing results supported strict factorial invariance of DESSA-Mini scores across all examined subgroups throughout nine occasions during 3 elementary school years, providing a foundation for valid cross-group comparisons of student SEC growth trajectories. The growth modeling results indicated that boys (vs. girls), Black students (vs. White students), Hispanic students (vs. White students), and lower-income students (vs. higher-income) started with a lower level of SEC; these gaps were either sustained over time (between Hispanic and White, and between lower-income and higher-income students) or slightly widened (between boys and girls throughout all 3 years, and between Black and White students during the first year). Taken together, these findings provide robust evidence of differences in the development of SEC, which is one of the critical protective and promotive factors of student mental health among other important outcomes.

Toward Equitable SEC Assessments with Diverse Students

Measurement invariance across subgroups is an essential but often overlooked criteria for studying subgroup health disparities (Meredith & Teresi, 2006). This gap is also evident in the literature on SEL. Although teacher-completed rating scales are now widely used as universal screening and progress monitoring tools in the domain of SEL, there is a lack of studies that test measurement invariance of these scales across socioculturally classified student subgroups. Along with our recent study with a different sample (Lee, Shapiro, & Kim, 2023), this is one of the first studies that explicitly tested the factorial invariance of the DESSA-Mini across student subgroups and, to our knowledge, any teacher-completed rating scales currently used for universal SEC assessment in educational practice as identified in the SEL Assessment Guide (Assessment Work Group, 2021). The evidence supporting the highest level of factorial invariance suggests that teachers rate all the DESSA-Mini items in a similar way for students who have the same level of SEC, regardless of their gender, race and ethnicity, and SES. Therefore, the observed subgroup differences in the DESSA-Mini scores can be interpreted as reflecting differences in the construct measured rather than teachers' differential rating behaviors based on student identities.

This finding not only allows researchers to make more valid cross-group comparisons to further investigate subgroup disparities in SEC, but also contributes to reducing the field's widespread concern about the potential of measurement bias when assessing diverse student bodies (Shapiro et al., 2016). Given ample evidence of teacher bias in reporting problem behaviors against marginalized students (Bradshaw et al., 2010; Gregory & Roberts, 2017; Skiba et al., 2002, 2011; Ura & d'Abreu, 2022), the findings of this study may be surprising. It should be noted, however, that this study only tested one aspect of measurement bias (i.e., measurement invariance) and only examined one behavioral rating scale (i.e., DESSA-Mini) composed exclusively of strength-based items (i.e., indicating positive behavior rather than maladaptive behavior). In fact, a review of 13 studies that compared teacher-completed behavioral rating scales against a third-party criterion measure of behavior suggested that there is a lack of consistent evidence supporting the presence of racial-ethnic bias in teacher ratings of student internalizing or externalizing behavior (Mason et al., 2014), despite documented disproportionality in disciplinary referrals. Taken together, this study calls for more explicit testing of measurement invariance of various rating scales for SEC assessment, especially the ones that are actively used in practice. Instrument developers and organizations working to transfer knowledge between research and practice should make this evidence (or the lack thereof) transparently available to all stakeholders and rights-holders. With more data available, factors associated with measurement invariance versus non-invariance could be further explored across different tools. For example, as Mason et al. (2014) suggested, teachers' implicit and explicit beliefs about certain subgroups or cultures (e.g., acculturation, cultural bias) could be measured and examined as a potential source of measurement bias. Such research will also inform the development and improvement of SEC assessments that are less prone to teacher bias.

For practitioners and decision-makers involved in the selection of SEC assessment tools for universal screening and progress monitoring, this study emphasizes the importance of looking for evidence of measurement invariance across relevant student subgroups, along with other psychometric properties. The users of invariant assessment tools will have more trustworthy information to guide score interpretation and fair utilization of local data in efforts to remediate any disparities, whether through focused SEL delivery, building home-school partnerships and culturally affirming practices, investing in structural rearrangements to reduce opportunity gaps, or reducing disproportionate representation in stigmatizing, exclusionary, or punitive practices.

Nevertheless, the evidence of measurement invariance does not guarantee that the DESSA-Mini is free of bias and equally relevant to all student subgroups. Furthermore, the source of information for any decision should not be limited to a single measure. Although strengths-based, the DESSA-Mini does not intentionally capture constructs such as positive racial and ethnic identity that have been found to promote academic achievement among Black youth (Adelabu, 2008) and to be protective against experiences of discrimination (Thomas et al., 2009). The Assessment Work Group (2019) has raised the issue of construct underrepresentation, which occurs when certain manifestations of SEC that are critical to understand specific subgroups' SEC are left out in the assessment (e.g., racial-ethnic identity, code-switching skills for students of color; Duchesneau, 2020). In this case, even if teachers rate the included items in the same way across subgroups, the rating scores may not capture the full spectrum of SEC for certain subgroups. El Mallah (2022) recently proposed the potential utility of prototype analysis that combines both qualitative content validation and quantitative invariance testing to develop

SEC measures that represent diverse cultural assets. Use of a multi-informant approach is also recommended to gain a more comprehensive understanding of student SEC. The present study calls for more research that contributes to the evidence base for equitable SEC assessments. Measurement invariance testing of a rating scale is just one of the necessary steps that has been under-investigated and more work is needed to make sure that SEC assessment tools are equally relevant and fairly applied across diverse subgroups of students.

Disparities in SEC Developmental Trajectories

This study also contributes to our knowledge regarding disparities in SEC developmental trajectories during the elementary school years. The observed gender, racial and ethnic, and socioeconomic disparities at baseline are generally consistent with previous findings in the literature (e.g., Aber et al., 2003; Jones S. M. et al., 2011; West et al., 2020), with added confidence that they are not an artifact of teachers' differential ratings. This study also provides relatively robust evidence that these disparities were not reduced throughout multiple years, even in the context of a districtwide SEL initiative. The effect sizes for baseline subgroup disparities were small for gender (ES = .43, $p < .001$) and negligible for race, ethnicity, and SES (ES = [.14, .17]) per Cohen's (1969) criteria. However, the use of Cohen's standards has been criticized to be too stringent for educational data (Kraft, 2019). Given that the estimated effect sizes for yearly SEC growth in this study ranged from .12 to .17 and the total average increase was estimated to be .30, it is quite alarming that boys, on average, lagged behind girls by more than 3 years of difference and that students of color and low-income students, on average, lagged behind their White and higher-income peers by about 1 year of difference.

Among the three goals of research on disparities (i.e., detecting, understanding, and reducing disparities), the focus of the present study was on disparity detection. It is beyond the scope of this study to explain the causes and mechanisms underlying these subgroup disparities. However, it should be emphasized that there is no reason to believe that there are any inherent differences in SEC by child gender, race and ethnicity, and SES. Rather, the subgroup disparities observed in this study may be understood as a reflection of different socialization or marginalization experiences based on these characteristics. For example, the baseline differences could be indicative of the inequalities in earlier social and emotional developmental environments such as parental time spent with children or early childhood education quality (Reardon & Portilla, 2016), whereas the sustainment of these gaps over time could indicate the continued cultural mismatch or structural inequities experienced throughout the school years.

In practice settings, the evidence of subgroup disparities in SEC must not be misused to reinforce existing prejudices and systems of oppression. Just as the academic achievement gap should be reframed as the educational debts that society owes its students (Ladson-Billings, 2006), we hope this observation of SEC disparities can inform and lead to practice decisions and policy-making efforts to effectively address the underlying social inequalities and ensure more equitable contexts for promoting social-emotional development.

From this standpoint, the findings of this study raise many interesting questions as to why boys, Black or Hispanic students, or low-income students, on average, consistently showed lower levels of SEC than their counterparts during the elementary school years. Regarding the gender gap, differential norms and expectations for social-emotional behavior based on gender (e.g., “boys will be boys”) and social pressure to conform to gender stereotypes (e.g., “boys are stronger both physically and emotionally”) might be responsible for the observed gender disparities in SEC (Skipper & Fox, 2022). For instance, some studies with younger children have found that parents’ and preschool teachers’ emotion socialization behaviors differed based on child gender, contributing to girls generally having higher emotional competence than boys (Denham et al., 2010; King, 2021). Gender stereotyping may not just hinder boys' positive social-emotional development, but may also take a toll on girls in the longer term. For example, in a large-scale self-report survey study that simulated social-emotional growth trajectories of students in Grades 4–12 (West et al., 2020), female students had consistently higher levels of self-management and social awareness than male students. Yet, their self-efficacy became lower than males from Grade 6 forward and this gap widened over time. More research is needed to better understand the mechanisms underlying these gender disparities in SEC development and the implications for promoting gender equity in mental and behavioral health.

Although smaller in size than gender disparities, racial-ethnic and socioeconomic disparities observed in this study require special attention as structurally disadvantaged subgroups were found to have lower levels of SEC than structurally advantaged subgroups. These gaps might reflect the societal racial, ethnic, and economic inequities that shape unequal developmental conditions. Research has suggested that students of color and low-income students tend to be more frequently exposed to stressful events that can impede children's healthy cognitive, emotional, and behavioral development (Morsy & Rothstein, 2019). Additionally, regarding race and ethnicity, cultural misalignment or dissonance between home and school has been suggested as an important factor to consider when understanding seemingly less adaptive (or more precisely, less conforming) social-emotional behavior among students of color within a school culture endorsing White, Anglocentric values (Arunkumar et al., 1999; Boykin et al., 2005). This also is connected to the question of bias in SEL assessments and whether culturally diverse manifestations of SEC are equally promoted and fully assessed in current SEL practices. Acknowledging again that the present study findings of measurement invariance are not sufficient to fully answer this question, further research is warranted to understand the underlying causes and mechanisms behind these racial, ethnic, and socioeconomic disparities to create more equitable school environments and challenge social inequities at large.

Limitations

This study has several limitations to note. First, as in other studies using school administrative data to measure student identities, the simplistic categorization of gender, race and ethnicity, and SES does not respect students' own views of their identities and obscures within-group variability. Second, this study cannot answer how intersectional identities based on gender, race and ethnicity, SES, and other constructs are linked to student SEC development. The challenges of defining and measuring subgroups (e.g., Perez & Hirschman, 2009) and examining intersectionality (Bauer, 2014; Bowleg, 2012) in quantitative research are well documented in the literature. Also, a substantial amount of variability in student SEC growth trajectories remained unexplained by the examined model. This study generates questions about advancing quantitative research on subgroup disparities to better represent real-world complexities. Based on our data indicating the presence of differences in SEC developmental trajectories, we call for more nuanced and sophisticated research that can further advance our understanding of subgroup disparities in SEC development that lead to appropriate societal interventions.

Another important limitation to note is that, with the observational study design, it was not possible to estimate counterfactual growth trajectories under the condition that no SEL initiative was implemented. Therefore, the fact that the disparities were not reduced across school years, and even widened for some subgroups, does not reveal anything about the effects of SEL practices or development under non-intervention conditions. To understand whether an SEL intervention reduces, sustains, or increases existing disparities, experimental or quasi-experimental studies involving a reasonable comparison condition are required (e.g., Lee, Shapiro, & Kim, 2023). In addition, future research should account for varying degrees of SEL implementation to further rule out the effects of differential exposure to the program.

Also, it is important to note that three different forms of the DESSA-Mini were used in rotation, conflating time and form when measuring the growth within each year. Although these three DESSA-Mini forms have shown high alternate form reliability and similarities in descriptive statistics (such as means and standard deviations), alpha coefficients, and correlations with the full 72-item DESSA score (Naglieri et al., 2011; Lee, Shapiro, Robitaille, & LeBuffe, 2023), it is not possible to test the measurement invariance of forms composed of non-overlapping items. The use of three to four different DESSA-Mini forms throughout a year reflects the way these tools are typically used in practice to monitor student progress while minimizing practice effects; nonetheless, future research should use longitudinally invariant measures to more precisely estimate the extent to which student SEC changes within and across years.

Among other factors contributing to the limited generalizability of study findings (e.g., time, age group, location), underrepresentation of other racial-ethnic groups should also be noted. In the present study that analyzed secondary data, SEC growth trajectories of students identified as American Indian or Alaska Native, Asian, Hawaiian or Pacific Islander, and multi-racial could not be reliably estimated due to small group sizes. However, small group size should not be a perpetual excuse for not investigating minority (or minoritized) groups' outcomes. More research representing diverse student populations is needed to advance epistemic justice. We also call for more contemporary studies with various samples from different developmental stages and geographical regions to gain a more up-to-date and comprehensive understanding of disparities in SEC development.

Recommendations for Future Research and Practice

Despite these limitations, this study provides relatively robust evidence of existing subgroup disparities in SEC developmental trajectories, suggesting (a) boys (vs. girls), Black students (vs. White students), Hispanic students (vs. White students), and low-income students (vs. middle-to-high-income students) were found to have a lower level of SEC than their counterparts and (b) these gaps were not reduced throughout the elementary school years. The measurement invariance findings support that the disparities observed in this study are not an artifact of teachers' differential ratings of SEC. These findings provide the groundwork for future research agendas and practices in the field of SEL, such as how to determine whether SEL assessments are equitable and fair across diverse student bodies, how to understand the observed subgroup disparities in SEC development, and how to effectively tackle these inequities through addressing the root causes in the school system and beyond rather than blaming individual students or by trying a "quick fix" (Duchesneau, 2020).

List of funding sources

The original project (including data collection and cleaning) was enabled through the support of the Pennsylvania Departments of Education, Health, and Public Welfare and then the Hellman Foundation. The first author received the Doctoral Completion Fellowship from the University of California, Berkeley, while conducting the current research (including data analysis and manuscript writing). No sponsor engaged in decisions regarding the study design, data interpretation, the decision to publish, or the claims made in this paper. The statements made in this article do not reflect the views of any funder or host organization.

Declaration of competing interest

The first author has no competing interests to declare. Although some of the other authors were involved in the development of the DESSA while employed at the Devereux Center for Resilient Children, no author of this paper receives any direct financial remuneration from the sale of the assessments or resources mentioned within this manuscript.

Acknowledgments

The authors wish to acknowledge all of the teachers, staff, and administrators who participated in the original Social and Emotional Learning Initiative, and specifically Dr. Gerald L. Zahorchak, Dr. C. Russell Mayo, Susan Lozada, Robin Powlus, and John Monahan. The authors also wish to thank Katherine Ross and Jennifer Croner who each played an essential role in the completion of this project. The authors appreciate Dr. Sophia Rabe-Hesketh, Dr. Susan I. Stone, Dr. Bo Kyung Elizabeth Kim, Dr. Sophia Hwang, Dr. Kelly L. Ziemer, and Nehal Eldeeb, who each provided valuable feedback and incredible support in writing this manuscript.

References

Aber, J. L., Brown, J. L., & Jones, S. M. (2003). Developmental trajectories toward violence in middle childhood: Course, demographic differences, and response to school-based intervention. *Developmental Psychology*, *39*(2), 324–348. http://doi.org/10.1037/0012-1649.39.2.324

Adelabu, D. H. (2008). Future time perspective, hope, and ethnic identity among African American adolescents. *Urban Education*, *43*(3), 347–360. https://doi.org/10.1177/0042085907311806

Alegría, M., Green, J. G., McLaughlin, K. A., & Loder, S. (2015). *Disparities in child and adolescent mental health and mental health services in the US*. William T. Grant Foundation.

Allison, P. D. (2003). Missing data techniques for structural equation modeling. *Journal of Abnormal Psychology*, *112*(4), 545–557. https://doi.org/10.1037/0021-843X.112.4.545

Anderson-Butcher, D., Amorose, A. J., Lower, L. M., Riley, A., Gibson, A., & Ruch, D. (2016). The case for the Perceived Social Competence Scale II. *Research on Social Work Practice*, *26*(4), 419–428. https://doi.org/10.1177/1049731514557362

Anthony, C. J., Elliott, S. N., DiPerna, J. C., & Lei, P.-W. (2021). Initial development and validation of the Social Skills Improvement System—Social and Emotional Learning Brief Scales-Teacher Form. *Journal of Psychoeducational Assessment*, *39*(2), 166–181. https://doi.org/10.1177/0734282920953240

Arunkumar, R., Midgley, C., & Urdan, T. (1999). Perceiving high or low home-school dissonance: Longitudinal effects on adolescent emotional and academic well-being. *Journal of Research on Adolescence*, *9*(4), 441–466.

https://doi.org/10.1207/s15327795jra0904_4

Assessment Work Group. (2019). *Student social and emotional competence assessment: The current state of the field and a vision for its future*. Collaborative for Academic, Social, and Emotional Learning (pp. 4-53). https://files.eric.ed.gov/fulltext/ED607094.pdf

Assessment Work Group. (2021). *SEL Assessment guide. Measuring SEL*. https://web.archive.org/web/20210304193245/https://measuringsel.casel.org/access-assessment-guide/

Basting, E. J., Naser, S., & Goncy, E. A. (2022). Assessing the factor structure and measurement invariance of the BASC-3 behavioral and emotional screening system student form across race/ethnicity and gender. *Assessment for Effective Intervention*, *48*(1), 43–51. https://doi.org/10.1177/15345084221095440

Bauer, G. R. (2014). Incorporating intersectionality theory into population health research methodology: Challenges and the potential to advance health equity. *Social Science & Medicine*, *110*, 10–17. https://doi.org/10.1016/j.socscimed.2014.03.022

Baumert, J., Nagy, G., & Lehmann, R. (2012). Cumulative advantages and the emergence of social and ethnic inequality: Matthew effects in reading and mathematics development within elementary schools? *Child Development*, *83*(4), 1347–1367. https://doi.org/10.1111/j.1467-8624.2012.01779.x

Bennett, P. R., Lutz, A. C., & Jayaram, L. (2012). Beyond the schoolyard: The role of parenting logics, financial resources, and social institutions in the social class gap in structured activity participation. *Sociology of Education*, *85*(2), 131–157. https://doi.org/10.1177/0038040711431585

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Bierman, K. L., Coie, J. D., Dodge, K. A., Greenberg, M. T., Lochman, J. E., McMahon, R. J., & Pinderhughes, E. (2010). The effects of a multiyear universal social–emotional learning program: The role of student and school characteristics. *Journal of Consulting and Clinical Psychology*, *78*(2), 156–168. https://doi.org/10.1037/a0018607

Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley & Sons.

Bowleg, L. (2012). The problem with the phrase women and minorities: Intersectionality—An important theoretical framework for public health. *American Journal of Public Health*, *102*(7), 1267–1273. https://doi.org/10.2105/AJPH.2012.300750

Boykin, A. W., Tyler, K. M., & Miller, O. (2005). In search of cultural themes and their expressions in the dynamics of classroom life. *Urban Education*, *40*(5), 521–549. https://doi.org/10.1177/0042085905278179

Bradshaw, C. P., Mitchell, M. M., O'Brennan, L. M., & Leaf, P. J. (2010). Multilevel exploration of factors contributing to the overrepresentation of black students in office disciplinary referrals. *Journal of Educational Psychology*, *102*(2), 508–520. https://doi.org/10.1037/a0018450

Bruner, C. (2017). ACE, place, race, and poverty: Building hope for children. *Academic Pediatrics*, *17*(7), S123–S129. https://doi.org/10.1016/j.acap.2017.05.009

Carpendale, E. J., Green, M. J., Williams, K. E., Tzoumakis, S., Carr, V. J., & Laurens, K. R. (2023). Measurement invariance across sex and language background of a self-report social-emotional learning questionnaire for middle childhood. *Child Indicators Research, 16*, 1233–1249. https://doi.org/10.1007/s12187-023-10015-3

Chain, J., Shapiro, V. B., LeBuffe, P. A., & Bryson, A. M. (2017). Academic achievement of

American Indian and Alaska native students: Does social-emotional competence reduce the impact of poverty? *American Indian and Alaska Native Mental Health Research*, *24*(1), 1–29. https://www.proquest.com/docview/1892134342?pq-origsite=gscholar&fromopenview=true

Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology*, *95*(5), 1005–1018. https://doi.org/10.1037/a0013193

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Chin, T., & Phillips, M. (2004). Social reproduction and child-rearing practices: Social class, children's agency, and the summer activity gap. *Sociology of Education*, *77*(3), 185–210. https://doi.org/10.1177/003804070407700301

Cipriano, C., Strambler, M. J., Naples, L. H., Ha, C., Kirk, M., Wood, M., Sehgal, K., Zieher, A. K., Eveleigh, A., McCarthy, M., Funaro, M., Ponnock, A., Chow, J. C., & Durlak, J. (2023). The State of Evidence for Social and Emotional Learning: A Contemporary Meta-Analysis of Universal School-Based SEL Interventions. Child Development, 94(5), 1181–1204. https://doi.org/10.1111/cdev.13968

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. Academic Press.

Colle, L., & Del Giudice, M. (2011). Patterns of attachment and emotional competence in middle childhood. *Social Development*, *20*(1), 51–72. https://doi.org/10.1111/j.1467-9507.2010.00576.x

Collins, W. A. (1984). *Development during middle childhood: The years from six to twelve*.

National Academy Press.

Conduct Problems Prevention Research Group. (1999). Initial impact of the Fast Track prevention trial for conduct problems: II. Classroom effects. *Journal of Consulting and Clinical Psychology*, *67*(5), 648–657. https://europepmc.org/article/pmc/pmc2761630

Crowder, M. K., Gordon, R. A., Brown, R. D., Davidson, L. A., & Domitrovich, C. E. (2019). Linking Social and Emotional Learning Standards to the WCSD Social-Emotional Competency Assessment: A Rasch Approach. School Psychology, 34(3), 281–295. https://doi.org/10.1037/spq0000308

Davis, K. M. (2020). Psychometric Properties and Construct Validity of a Student Survey Measuring the Casel Five Core Social and Emotional Competencies. ProQuest Dissertations Publishing.

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

Denham, S. A., Bassett, H. H., & Wyatt, T. M. (2010). Gender differences in the socialization of preschoolers' emotional competence. *New Directions for Child and Adolescent Development*, *2010*(128), 29–49. https://doi.org/10.1002/cd.267

Dixon-Román, E. J. (2017). *Inheriting possibility: Social reproduction and quantification in education*. The University of Minnesota Press.

Domitrovich, C. E., Durlak, J. A., Staley, K. C., & Weissberg, R. P. (2017). Social-emotional competence: An essential factor for promoting positive adjustment and reducing risk in school children. *Child Development*, *88*(2), 408–416. https://doi.org/10.1111/cdev.12739

Drummond, T. (1994). Student Risk Screening Scale (SRSS) [Database record]. APA PsycTests.

https://doi.org/10.1037/t27737-000

Duchesneau, N. (2020). *Social, emotional, and academic development through an equity lens.* The Education Trust (pp. 2-48). https://files.eric.ed.gov/fulltext/ED607298.pdf

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

Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial development. In N. Eisenberg (Ed.)*, Handbook of child psychology* (Vol. 1–3, pp. 646–718). John Wiley & Sons. http://onlinelibrary.wiley.com/doi/10.1002/9780470147658.chpsy0311/full

El Mallah, S. (2022). Toward equity-oriented assessment of social and emotional learning: Examining equivalence of concepts and measures. *Urban Education*, *57*(2), 289–317. https://doi.org/10.1177/0042085920933335

Elias, M. J., & Haynes, N. M. (2008). Social competence, social support, and academic achievement in minority, low-income, urban elementary school children. *School Psychology Quarterly*, *23*(4), 474–495. https://doi.org/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., Kessler, R., Schwab-Stone, M. E., & Shriver, T. P. (1997). *Promoting social and emotional learning: Guidelines for educators (pp. 1-164)*. Association for Supervision and Curriculum Development.

Fishbein, D. H., Domitrovich, C., Williams, J., Gitukui, S., Guthrie, C., Shapiro, D., & Greenberg, M. (2016). Short-term intervention effects of the PATHS curriculum in young low-income children: Capitalizing on plasticity. *The Journal of Primary Prevention*, *37*(6), 493–511. https://doi.org/10.1007/s10935-016-0452-5

Fredrick, S. S., Drevon, D. D., & Jervinsky, M. (2019). Measurement invariance of the Student Risk Screening Scale across time and gender. *School Psychology*, *34*(2), 159–167. https://doi.org/10.1037/spq0000278

Frey, K. S., Hirschstein, M. K., Snell, J. L., Edstrom, L. V. S., MacKenzie, E. P., & Broderick, C. J. (2005). Reducing playground bullying and supporting beliefs: An experimental trial of the steps to respect program. *Developmental Psychology*, *41*(3), 479–490. https://doi.org/10.1037/0012-1649.41.3.479

Garner, P. W., Mahatmya, D., Brown, E. L., & Vesely, C. K. (2014). Promoting desirable outcomes among culturally and ethnically diverse children in social emotional learning programs: A multilevel heuristic model. *Educational Psychology Review*, *26*(1), 165–189. https://doi.org/10.1007/s10648-014-9253-7

Gehlbach, H., & Hough, H. J. (2018). *Measuring social emotional learning through student surveys in the CORE districts: A pragmatic approach to validity and reliability.* Policy Analysis for California Education (pp. 1-29). https://files.eric.ed.gov/fulltext/ED591082.pdf

Greenberg, M. T., Kusche, C. A., Cook, E. T., & Quamma, J. P. (1995). Promoting emotional competence in school-aged children: The effects of the PATHS curriculum. *Development and Psychopathology*, *7*, 117–136. https://doi.org/10.1017/S0954579400006374

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

Gregory, A., & Roberts, G. (2017). Teacher beliefs and the overrepresentation of Black students in classroom discipline. *Theory Into Practice*, *56*(3), 187–194.

https://doi.org/10.1080/00405841.2017.1336035

Hair, J. K., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson.

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

Holsen, I., Iversen, A. C., & Smith, B. H. (2009). Universal social competence promotion programme in school: Does it work for children with low socio-economic background? *Advances in School Mental Health Promotion*, *2*(2), 51–60. https://doi.org/10.1080/1754730X.2009.9715704

Hutchison, M., Russell, B. S., & Wink, M. N. (2020). Social-emotional competence trajectories from a school-based child trauma symptom intervention in a disadvantaged community. *Psychology in the Schools*, *57*(8), 1257–1272. https://doi.org/10.1002/pits.22388

Hwang, S. H., Shapiro, V. B., Eldeeb, N., Lee, J., Robitaille, J. L., & Naglieri, J. A. (2023). Assessing social and emotional competencies in educational settings: Supporting resilience in young people. In S. Goldstein & R. Brooks (Eds.), *Handbook of resilience in children* (pp. 237–249). Springer International Publishing.

Institute of Education Sciences. (2020). *Characteristics of public and private elementary and secondary school teachers in the United States: Results from the 2017–18 National Teacher and Principal Survey* (NCES 2020-142rev). The National Center for Education Statistics, U.S. Department of Education. https://nces.ed.gov/pubs2020/2020142rev.pdf

Jones, S. M., Brown, J. L., & Aber, J. L. (2011). Two-year impacts of a universal school-based social-emotional and literacy intervention: An experiment in translational developmental research. *Child Development*, *82*(2), 533–554. https://doi.org/10.1111/j.1467-8624.2010.01560.x

Jones, T. M., Fleming, C., & Williford, A. (2020). Racial equity in academic success: The role of school climate and social emotional learning. Children and Youth Services Review, 119, 105623-. https://doi.org/10.1016/j.childyouth.2020.105623

Kilbourne, A. M., Switzer, G., Hyman, K., Crowley-Matoka, M., & Fine, M. J. (2006). Advancing health disparities research within the health care system: A conceptual framework. *American Journal of Public Health*, *96*(12), 2113–2121. https://doi.org/10.2105/AJPH.2005.077628

Kilgus, S. P., Chafouleas, S. M., & Riley-Tillman, T. C. (2013). Development and Initial Validation of the Social and Academic Behavior Risk Screener for Elementary Grades. School Psychology Quarterly, 28(3), 210–226. https://doi.org/10.1037/spq0000024

Kim, E. S., & Willson, V. L. (2014). Measurement invariance across groups in latent growth modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(3), 408–424. https://digitalcommons.usf.edu/edq_facpub/1

King, E. K. (2021). Fostering toddlers' social emotional competence: Considerations of teachers' emotion language by child gender. *Early Child Development and Care*, *191*(16), 2494–2507. https://doi.org/10.1080/03004430.2020.1718670

Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). Guilford publications.

Koch, T., Holtmann, J., Eid, M., & West, S. G. (2020). Analyzing longitudinal multirater data with changing and stable raters. *Structural Equation Modeling: A Multidisciplinary Journal*, *27*(1), 73–87. https://doi.org/10.1080/10705511.2019.1638784

Kozlowski, K. P. (2015). Culture or teacher bias? Racial and ethnic variation in student–teacher effort assessment match/mismatch. *Race and Social Problems*, *7*(1), 43–59. https://doi.org/10.1007/s12552-014-9138-x

Kraft, M. A. (2019). *Interpreting effect sizes of education interventions* (EdWorkingPaper No. 19-10). Annenberg Institute at Brown University. https://scholar.harvard.edu/files/mkraft/files/kraft_2019_effect_sizes.pdf

Krishnan, V. (2011). Teachers' assessment of preschoolers' social and emotional competence: Does sex of children matter in developmental outcomes? *Proceedings of the World Academy of Science, Engineering and Technology, Year 7, Issue 79*, 1622–1632. https://www.ualberta.ca/community-university-partnership/media-library/community-university-partnership/research/ecmap-reports/teacherassessment-of-preschoolers-sei.pdf

Kumar, R. (2006). Students' experiences of home–school dissonance: The role of school academic culture and perceptions of classroom goal structures. *Contemporary Educational Psychology*, *31*(3), 253–279. https://doi.org/10.1016/j.cedpsych.2005.08.002

Kusché, C. A., & Greenberg, M. T. (1994). *The PATHS Curriculum: Promoting alternative thinking strategies*. Developmental Research & Programs.

Ladson-Billings, G. (2006). From the achievement gap to the education debt: Understanding achievement in US schools. *Educational Researcher*, *35*(7), 3–12. https://www.jstor.org/stable/3876731

Lareau, A. (2011). *Unequal childhoods: Class, race, and family life, with an update a decade later*. The University of California Press.

LeBuffe, P. A., Shapiro, V. B., & Naglieri, J. A. (2009). *The Devereux Student Strengths Assessment (DESSA) Assessment, technical manual, and user's guide*. Lewisville, NC:

Kaplan Press.

Lee, J., Shapiro, V. B., & Kim, B. K. E. (2023). Universal school-based social and emotional learning (SEL) for diverse student subgroups: Implications for enhancing equity through SEL. *Prevention Science, 24*, 1011–1022. https://doi.org/10.1007/s11121-023-01552-y

Lee, J., Shapiro, V. B., Kim, B. E., & Yoo, J. P. (2018). Multilevel structural equation modeling for social work researchers: An introduction and application to healthy youth development. *Journal of the Society for Social Work and Research*, *9*(4), 689–719. https://doi.org/10.1086/701526

Lee, J., Shapiro, V. B., Robitaille, J. L., & LeBuffe, P. (2023). Measuring the development of social-emotional competence using behavioral rating scales in the context of school-based social and emotional learning. *Social and Emotional Learning: Research, Practice, and Policy*, *2*, 100015. https://doi.org/10.1016/j.sel.2023.100015

Leonardo, Z. (2004). The color of supremacy: Beyond the discourse of 'white privilege.' *Educational Philosophy and Theory*, *36*(2), 137–152. https://doi.org/10.1111/j.1469-5812.2004.00057.x

Leonardo, Z. (2012). The race for class: Reflections on a critical raceclass theory of education. *Educational Studies*, *48*(5), 427–449. https://doi.org/10.1080/00131946.2012.715831

Little, T. D. (2013). *Longitudinal structural equation modeling*. The Guilford Press.

Mahfouz, J., & Anthony-Stevens, V. (2020, April). Why trouble SEL? The need for cultural relevance in SEL. *Occasional Paper Series*, (43), 58–70. https://doi.org/10.58295/2375-3668.1354

Maguire, L. K., Niens, U., McCann, M., & Connolly, P. (2016). Emotional development among early school-age children: gender differences in the role of problem behaviours.

Educational Psychology (Dorchester-on-Thames), 36(8), 1408–1428. https://doi.org/10.1080/01443410.2015.1034090

Mahoney, J. L., LeBuffe, P. A., Shapiro, V. B., Robitaille, J. L., Johnson, E. S., & Adamson, J. L. (2022). *What is assessment bias and how is Aperture Education working to reduce it?* Aperture Education (pp. 3-16). https://f.hubspotusercontent30.net/hubfs/2613587/White%20Papers/Assessment%20Bias%20White%20Paper.pdf

Malti, T., Ribeaud, D., & Eisner, M. P. (2011). The effectiveness of two universal preventive interventions in reducing children's externalizing behavior: A cluster randomized controlled trial. *Journal of Clinical Child & Adolescent Psychology*, *40*(5), 677–692. https://doi.org/10.1080/15374416.2011.597084

Mantz, L. S., Bear, G. G., Yang, C., & Harris, A. (2018). The Delaware Social-Emotional Competency Scale (DSECS-S): Evidence of validity and reliability. *Child Indicators Research*, *11*(1), 137–157. https://doi.org/10.1007/s12187-016-9427-6

Mason, B. A., Gunersel, A. B., & Ney, E. A. (2014). Cultural and ethnic bias in teacher ratings of behavior: A criterion-focused review. *Psychology in the Schools*, *51*(10), 1017–1030. https://doi.org/10.1002/pits.21800

McKown, C. (2019). Reliability, factor structure, and measurement invariance of a web-based assessment of children's social-emotional comprehension. *Journal of Psychoeducational Assessment*, *37*(4), 435–449. https://doi.org/10.1177/0734282917749682

Meredith, W., & Teresi, J. A. (2006). An essay on measurement and factorial invariance. *Medical Care*, S69–S77. https://www.jstor.org/stable/41219507

Merton, R. K. (1968). The Matthew Effect in Science: The reward and communication systems of science are considered. *Science*, *159*(3810), 56–63. https://doi.org/10.1126/science.159.3810.56

Millsap, R. E., & Everson, H. T. (1993). Methodology review: Statistical approaches for assessing measurement bias. *Applied Psychological Measurement*, *17*(4), 297–334. https://doi.org/10.1177/014662169301700401

Morsy, L., & Rothstein, R. (2019). Toxic stress and children's outcomes: African American children growing up poor are at greater risk of disrupted physiological functioning and depressed academic achievement. *Economic Policy Institute*. https://files.eric.ed.gov/fulltext/ED598149.pdf

Muthén, L. K., & Muthén, B. O. (1998-2017). Mplus statistical analysis with latent variables: User's Guide. Muthén & Muthén.

Naglieri, J. A., LeBuffe, P. A., & Shapiro, V. B. (2011). *The Devereux Student Strengths Assessment—Mini (DESSA-Mini): Assessment, technical manual, and user's guide*. Lewisville, NC: Kaplan Press.

Naglieri, J. A., LeBuffe, P. A., & Shapiro, V. B. (2013). Assessment of social-emotional competencies related to resilience. In S. Goldstein & R. B. Brooks (Eds.), *Handbook of resilience in children* (2nd ed.; pp. 261–272). Springer. https://doi.org/10.1007/978-1-4614-3661-4_15

Nakajima, N., Jung, H., Pradhan, M., Hasan, A., Kinnell, A., & Brinkman, S. (2020). Gender gaps in cognitive and social-emotional skills in early primary grades: Evidence from rural Indonesia. Developmental Science, 23(5), e12931-n/a. https://doi.org/10.1111/desc.12931

Pendergast, L. L., von der Embse, N., Kilgus, S. P., & Eklund, K. R. (2017). Measurement equivalence: A non-technical primer on categorical multi-group confirmatory factor analysis in school psychology. *Journal of School Psychology*, *60*, 65–82. https://doi.org/10.1016/j.jsp.2016.11.002

Pennsylvania Department of Education. (2021). *Public school enrollments*. https://www.education.pa.gov/DataAndReporting/Enrollment/Pages/PublicSchEnrReports.aspx

Perez, A. D., & Hirschman, C. (2009). The changing racial and ethnic composition of the US population: Emerging American identities. *Population and Development Review*, *35*(1), 1–51. https://doi.org/10.1111/j.1728-4457.2009.00260.x

Peters, C. D., Kranzler, J. H., Algina, J., Smith, S. W., & Daunic, A. P. (2014). Understanding disproportionate representation in special education by examining group differences in behavior ratings. *Psychology in the Schools*, *51*(5), 452–465. https://doi.org/10.1002/pits.21761

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Raver, C. C., & Knitzer, J. (2002). *Ready to enter: What research tells policymakers about strategies to promote social and emotional school readiness among three-and four-year-old children*. National Center for Children in Poverty, Columbia University (pp. 2-24). https://doi.org/10.7916/D82V2QVX

Reardon, S. F., & Portilla, X. A. (2016). Recent trends in income, racial, and ethnic school readiness gaps at kindergarten entry. *AERA Open*, *2*(3), 1–18. https://doi.org/10.1177/2332858416657343

Rowe, H. L., & Trickett, E. J. (2018). Student Diversity Representation and Reporting in

Universal School-Based Social and Emotional Learning Programs: Implications for Generalizability. Educational Psychology Review, 30(2), 559–583. https://doi.org/10.1007/s10648-017-9425-3

Sarcedo, G. L., Matias, C. E., Montoya, R., & Nishi, N. (2015). Dirty dancing with race and class: Microaggressions toward first-generation and low income college students of color. *Journal of Critical Scholarship on Higher Education and Student Affairs*, *2*(1), Article 1. https://ecommons.luc.edu/jcshesa/vol2/iss1/1

Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika*, *75*(2), 243–248. https://doi.org/10.1007/s11336-009-9135-y

Schonfeld, D. J., Adams, R. E., Fredstrom, B. K., Weissberg, R. P., Gilman, R., Voyce, C., Tomlin, R., & Speese-Linehan, D. (2015). Cluster-randomized trial demonstrating impact on academic achievement of elementary social-emotional learning. *School Psychology Quarterly*, *30*(3), 406–420. https://doi.org/10.1037/spq0000099

Shapiro, V. B., Accomazzo, S., & Robitaille, J. L. (2017). In the same ballpark or a whole new ball game? Staff as raters of youth behavior. *Journal of Child and Family Studies*, *26*(4), 1051–1055. https://doi.org/10.1007/s10826-016-0632-1

Shapiro, V. B., Kim, B. K. E., Accomazzo, S., & Roscoe, J. N. (2016). Predictors of rater bias in the assessment of social-emotional competence. *International Journal of Emotional Education*, *8*(2), 25–44. https://files.eric.ed.gov/fulltext/EJ1121269.pdf

Shapiro, V. B., Kim, B. K., Robitaille, J. L., & LeBuffe, P. A. (2017). Protective factor screening for prevention practice: Sensitivity and specificity of the DESSA-Mini. *School Psychology Quarterly*, *32*(4), 449–464. https://doi.org/10.1037/spq0000181

Shapiro, V. B., Kim, B. E., Robitaille, J. L., LeBuffe, P. A., & Ziemer, K. L. (2018). Efficient implementation monitoring in routine prevention practice: A grand challenge for schools. *Journal of the Society for Social Work and Research*, *9*(3), 377–394. https://doi.org/10.1086/699153

Shapiro, V. B., Kim, B. K. E., Robitaille, J. F., Mahoney, J. F., Lee, J., & LeBuffe, P. A. (2024). Monitoring the growth of social and emotional competence to guide practice. Social and Emotional Learning: Research, Practice, and Policy. Advance Online Publication. https://doi.org/10.1016/j.sel.2024.100032

Shulman, C. (2016). Social and emotional development in infant and early childhood mental health. In *Research and Practice in Infant and Early Childhood Mental Health* (pp. 23–42). Springer International Publishing. https://link.springer.com/chapter/10.1007/978-3-319-31181-4_2

Simmons, C. A., Shapiro, V. B., Accomazzo, S., & Manthey, T. J. (2016). Strengths-based social work: A meta-theory to guide social work research and practice. In N. Coady & P. Lehmann (Eds.), *Theoretical perspectives for direct social work practice* (3rd ed.; pp. 131–154). Springer. https://escholarship.org/uc/item/7bn5d1w1

Skiba, R. J., Horner, R. H., Chung, C.-G., Rausch, M. K., May, S. L., & Tobin, T. (2011). Race is not neutral: A national investigation of African American and Latino disproportionality in school discipline. *School Psychology Review*, *40*(1), 85–107. https://doi.org/10.1080/02796015.2011.12087730

Skiba, R. J., Michael, R. S., Nardo, A. C., & Peterson, R. L. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *The Urban Review*, *34*(4), 317–342. https://doi.org/10.1023/A:1021320817372

Skipper, Y., & Fox, C. (2022). Boys will be boys: Young people's perceptions and experiences of gender within education. *Pastoral Care in Education*, 40(4), 391–409. https://doi.org/10.1080/02643944.2021.1977986

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, *25*(2), 173–180. https://doi.org/10.1207/s15327906mbr2502_4

Stephens, N. M., & Townsend, S. S. (2015). The norms that drive behavior: Implications for cultural mismatch theory. *Journal of Cross-Cultural Psychology*, *46*(10), 1304–1306. https://doi.org/10.1177/0022022115600264

Sue, D. W. (2010). *Microaggressions and Marginality: Manifestation, Dynamics, and Impact*. John Wiley & Sons.

Tedeschi, R. G., & Calhoun, L. G. (2004). Posttraumatic growth: Conceptual foundations and empirical evidence. *Psychological Inquiry*, *15*(1), 1–18. https://doi.org/10.1207/s15327965pli1501_01

Thomas, O. N., Caldwell, C. H., Faison, N., & Jackson, J. S. (2009). Promoting academic achievement: The role of racial identity in buffering perceptions of teacher discrimination on academic achievement among African American and Caribbean Black adolescents. *Journal of Educational Psychology*, *101*(2), 420–431. https://doi.org/10.1037/a0014578

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*(1), 1–10. https://doi.org/10.1007/BF02291170

Ura, S. K., & d'Abreu, A. (2022). Racial bias, social–emotional competence, and teachers' evaluation of student behavior. *Children & Schools*, *44*(1), 17–26. https://doi.org/10.1093/cs/cdab028

Villegas, A. M. (1988). School failure and cultural mismatch: Another view. *The Urban Review*, *20*(4), 253–265. https://doi.org/10.1007/BF01120137

von der Embse, N., Kilgus, S. P., Iaccarino, S., & Levi-Nielsen, S. (2017). Screening for student mental health risk: Diagnostic accuracy, measurement invariance, and predictive validity of the Social, Academic, and Emotional Behavior Risk Screener-Student Rating Scale (SAEBRS-SRS). *School Mental Health*, *9*(3), 273–283. https://doi.org/10.1007/s12310-017-9214-7

von der Embse, N., Kim, E. S., Kilgus, S., Dedrick, R., & Sanchez, A. (2019). Multi-informant universal screening: Evaluation of rater, item, and construct variance using a trifactor model. *Journal of School Psychology*, *77*, 52–66. https://doi.org/10.1016/j.jsp.2019.09.005

Washburn, I. J., Acock, A., Vuchinich, S., Snyder, F., Li, K.-K., Ji, P., Day, J., DuBois, D., & Flay, B. R. (2011). Effects of a Social-Emotional and Character Development Program on the Trajectory of Behaviors Associated with Social-Emotional and Character Development: Findings from Three Randomized Trials. Prevention Science, 12(3), 314–323. https://doi.org/10.1007/s11121-011-0230-9

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

West, M. R., Pier, L., Fricke, H., Hough, H., Loeb, S., Meyer, R. H., & Rice, A. B. (2020). Trends in student social-emotional learning: Evidence from the first large-scale panel student survey. *Educational Evaluation and Policy Analysis*, *42*(2), 279–303. http://doi.org/10.3102/0162373720912236

Wilson, W. J. (2009). Toward a framework for understanding forces that contribute to or reinforce racial inequality. *Race and Social Problems*, *1*(1), 3–11. http://nrs.harvard.edu/urn-3:HUL.InstRepos:4778369

Wittmer, D., Doll, B., & Strain, P. (1996). Social and emotional development in early childhood: The identification of competence and disabilities. *Journal of Early Intervention*, *20*(4), 299–318. https://doi.org/10.1177/105381519602000404

Yang, Y., Luo, Y., & Zhang, Q. (2021). A cautionary note on identification and scaling issues in second-order latent growth models. *Structural Equation Modeling: A Multidisciplinary Journal*, 28(2), 302–313. https://doi.org/10.1080/10705511.2020.1747938

Yates, T., Ostrosky, M. M., Cheatham, G. A., Fettig, A., Shaffer, L., & Santos, R. M. (2008). *Research synthesis on screening and assessing social-emotional competence*. The Center on the Social and Emotional Foundations for Early Learning. https://csefel.vanderbilt.edu/documents/rs_screening_assessment.pdf

Zahn-Waxler, C., & Smith, K. D. (1992). The development of prosocial behavior. In V. B. Van Hasselt & M. Hersen (Eds.), *Handbook of social development: A lifespan perspective*, (pp. 229–256). Springer.

Zarbatany, L., Hartmann, D. P., Gelfand, D. M., & Vinciguerra, P. (1985). Gender differences in altruistic reputation: Are they artifactual? *Developmental Psychology*, *21*(1), 97–101. https://doi.org/10.1037/0012-1649.21.1.97

Zhang, D., & Katsiyannis, A. (2002). Minority representation in special education: A persistent challenge. *Remedial and Special Education*, *23*(3), 180–187. https://doi.org/10.1177/07419325020230030601

Zimmermann, P., & Iwanski, A. (2014). Emotion regulation from early adolescence to emerging adulthood and middle adulthood: Age differences, gender differences, and emotion-specific developmental variations. International Journal of Behavioral Development, 38(2), 182–194. https://doi.org/10.1177/0165025413515405

Zins, J. E., Bloodworth, M. R., Weissberg, R. P., & Walberg, H. J. (2004). The scientific base linking social and emotional learning to school success. In J. E. Zins, R. P. Weissberg, M. C. Wang, & H. J. Walberg (Eds.), *Building academic success on social and emotional learning: What does the research say?* (pp. 3–22). Teachers College Press.

Zoric, T. (2014). Poverty and class bias in schools. In E. L. Brown, P. C. Gorski, & G. Lazaridis (Eds.), *Poverty, class, and schooling: Global perspectives on economic justice and educational equity* (pp. 347–364). Information Age Publishing.

Table 1

Sample Characteristics

Variable	Category	Frequency	Percent
Grade at Baseline	K	1874	34.37%
	1	1803	33.07%
	2	1775	32.56%
Gender	Girl	2614	47.95%
	Boy	2836	52.02%
	Missing	2	0.04%
Race and Ethnicity	American Indian or Alaska Native	9	0.17%
	Asian	67	1.23%
	Black (non-Hispanic)	861	15.79%
	Hawaiian or Pacific Islander	9	0.17%
	Hispanic (of any race)	3485	63.92%
	White (non-Hispanic)	744	13.65%
	Multi-race	275	5.04%
	Missing	2	0.04%
Socioeconomic Status	Low-income	4747	87.07%
	Middle-to-high income	697	12.78%
	Missing	8	0.15%
Exposure to PATHS	Continued for 3 years	4237	77.71%
	Discontinued in Years 2–3	1215	22.29%
Total		5452	100.00%

Table 2

Descriptive Statistics of DESSA-Mini Ratings

		Average of eight items				*T*-Scores		Cronbach's
Time	*N*	*M*	*SD*	Skewness	Kurtosis	*M*	*SD*	α
Year 1 Fall	3697	2.69	0.96	-0.43	-0.28	50.16	11.14	0.95
Year 1 Winter	3149	2.85	0.95	-0.50	-0.33	53.18	11.72	0.96
Year 1 Spring	3476	2.86	1.00	-0.56	-0.40	54.01	12.05	0.97
Year 2 Fall	3744	2.76	0.97	-0.44	-0.39	51.29	11.75	0.96
Year 2 Winter	2829	2.87	0.96	-0.47	-0.48	53.52	12.13	0.97
Year 2 Spring	3656	2.84	1.02	-0.52	-0.55	53.87	12.49	0.98
Year 3 Fall	3493	2.84	0.94	-0.49	-0.29	52.29	11.69	0.96
Year 3 Winter	2263	2.82	0.98	-0.47	-0.45	52.91	12.24	0.97
Year 3 Spring	2640	2.90	0.97	-0.58	-0.30	54.42	12.04	0.98

Note. The descriptive statistics were presented only at the composite score level for simplicity. There was negligible item-level variation in these statistics. The number of valid cases did not differ at the item level at each occasion.

Table 3

Missing Data Patterns

Missing Pattern	Frequency	Percent
None missing across all 3 years	686	12.58%
Missing earlier ratings		
Missing data for Year 1 + some valid data for Year 2 and Year 3	852	15.63%
Missing data for Year 1 and Year 2 + some valid data for Year 3	412	7.56%
Missing later ratings		
Some valid data for Year 1 + missing data for Year 1 and Year 2	512	9.39%
Some valid data for Year 1 and Year 2 + missing data for Year 1	758	13.90%
Missing inconsistently		
Missing at one time point across 3 years	784	14.38%
Missing at two time points across 3 years	842	15.44%
Other inconsistent missing patterns	606	11.12
Missing all ratings across 3 years	0	0.00
Total	5452	100.00

Table 4

Fit Indices of a Series of Models with Increasing Invariance

Model [Comparison Model]	$\chi^2(df)$ [$\Delta\chi^2(\Delta df)$][a]	CFI [ΔCFI]	TLI [ΔTLI]	RMSEA [ΔRMSEA]
By Gender (girl vs. boy)				
Model 1: Configural invariance	18889.95(4494)***	0.94	0.93	0.03
[—]	[—]	[—]	[—]	[—]
Model 2: Weak factorial invariance	18987.93(4518)***	0.94	0.93	0.03
[vs. Model 1]	[97.56(24)***]	[.00]	[.00]	[.00]
Model 3: Strong factorial invariance	19509.47(4542)***	0.94	0.93	0.03
[vs. Model 2]	[574.77(24)***]	[.00]	[.00]	[.00]
Model 4: Strict factorial invariance	20308.60(4614)***	0.93	0.93	0.04
[vs. Model 3]	[609.39(72)***]	[-.01]	[.00]	[.01]
By Race/Ethnicity (Hispanic vs. non-Hispanic Black vs. non-Hispanic White)				
Model 1: Configural invariance	24330.89(6732)***	0.94	0.93	0.04
[—]	[—]	[—]	[—]	[—]
Model 2: Weak factorial invariance	24375.75(6780)***	0.94	0.93	0.04
[vs. Model 1]	[44.86(48)]	[.00]	[.00]	[.00]
Model 3: Strong factorial invariance	24629.55(6828)***	0.94	0.93	0.04
[vs. Model 2]	[253.80(48)***]	[.00]	[.00]	[.00]
Model 4: Strict factorial invariance	25099.37(6872)***	0.94	0.93	0.04
[vs. Model 3]	[469.82(144)***]	[.00]	[.00]	[.00]
By Socioeconomic Status (low-SES vs. middle-to-high-SES)				
Model 1: Configural invariance	19356.52(4494)***	0.94	0.93	0.04
[—]	[—]	[—]	[—]	[—]

Model 2: Weak factorial invariance	19417.25(4518)***	0.94	0.93	0.04
[vs. Model 1]	[52.51(24)**]	[.00]	[.00]	[.00]
Model 3: Strong factorial invariance	19563.06(4542)***	0.94	0.93	0.04
[vs. Model 2]	[152.19(24)***]	[.00]	[.00]	[.00]
Model 4: Strict factorial invariance	19552.63(4614)***	0.94	0.93	0.03
[vs. Model 3]	[225.09(72)**]	[.00]	[.00]	[-.00]

Note. Strong factorial invariance over time was assumed for each model.
[a] The chi-square difference test statistics used the Satorra-Bentler scaling correction.
*p < .05. **p < .01. ***p < .001.

Table 5

Estimated Unstandardized Regression Coefficients of the Conditional Piecewise SEC Growth Model

	Est.	*SE*	*p*	Est.	SE	*p*	Est.	*SE*	*p*
		Intercept (R^2 = .14)							
Boy	-0.43	0.03	< .001						
Black (vs. White)	-0.17	0.05	.002						
Hispanic (vs. White)	-0.14	0.04	.001						
Low-income	-0.15	0.05	.001						
Grade	0.18	0.02	< .001						
PATHS-continued	0.16	0.04	< .001						
		Slope_{Y1} (R^2=.19)			Slope_{Y2} (R^2=.06)			Slope_{Y3} (R^2=.03)	
Boy	-0.05	0.03	.085	-0.05	0.03	.050	-0.05	0.03	.075
Black (vs. White)	-0.11	0.05	.028	0.03	0.05	.523	-0.04	0.06	.488
Hispanic (vs. White)	0.01	0.04	.888	0.05	0.04	.165	0.06	0.05	.215
Low-SES	0.05	0.04	.221	-0.01	0.04	.821	-0.05	0.05	.215
Grade	-0.17	0.02	< .001	-0.10	0.02	<.001	-0.02	0.02	.239
PATHS-continued	0.00	0.04	.958	0.16	0.03	<.001	0.17	0.04	< .001

Figure 1

Longitudinal Confirmatory Factor Model of Student Social-Emotional Competence (SEC) Measured by Three DESSA-Mini Forms Across Nine Times of Measurement

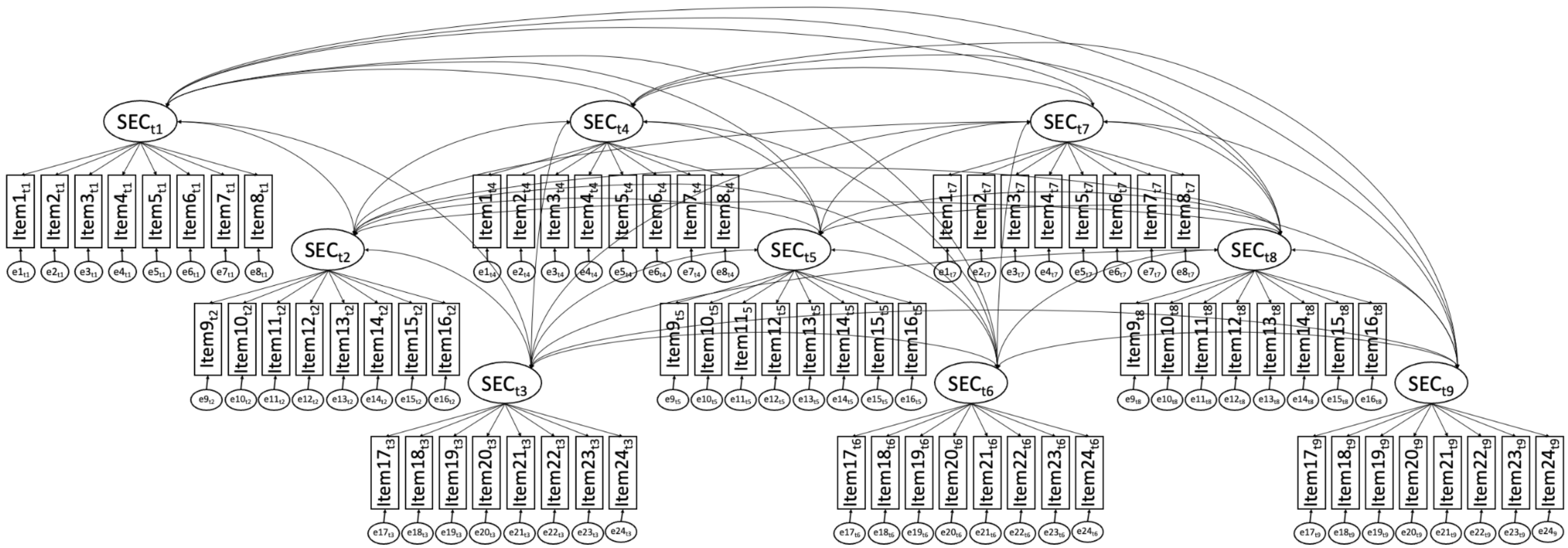


Figure 2

Conditional Piecewise Second-order Latent Growth Model of Student Social-Emotional Competence (SEC) Measured by Three DESSA-Mini Forms Across Nine Times of Measurement

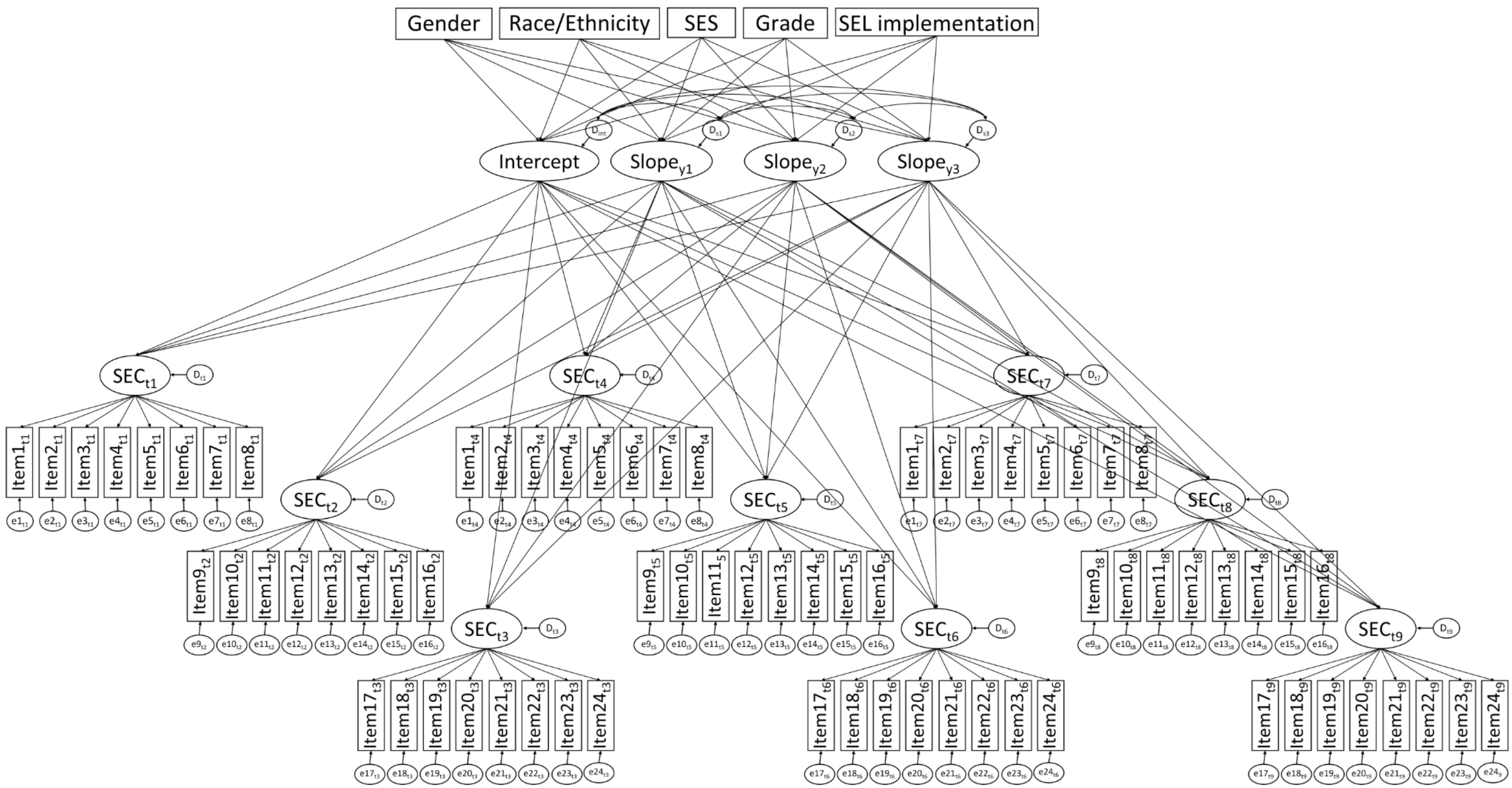
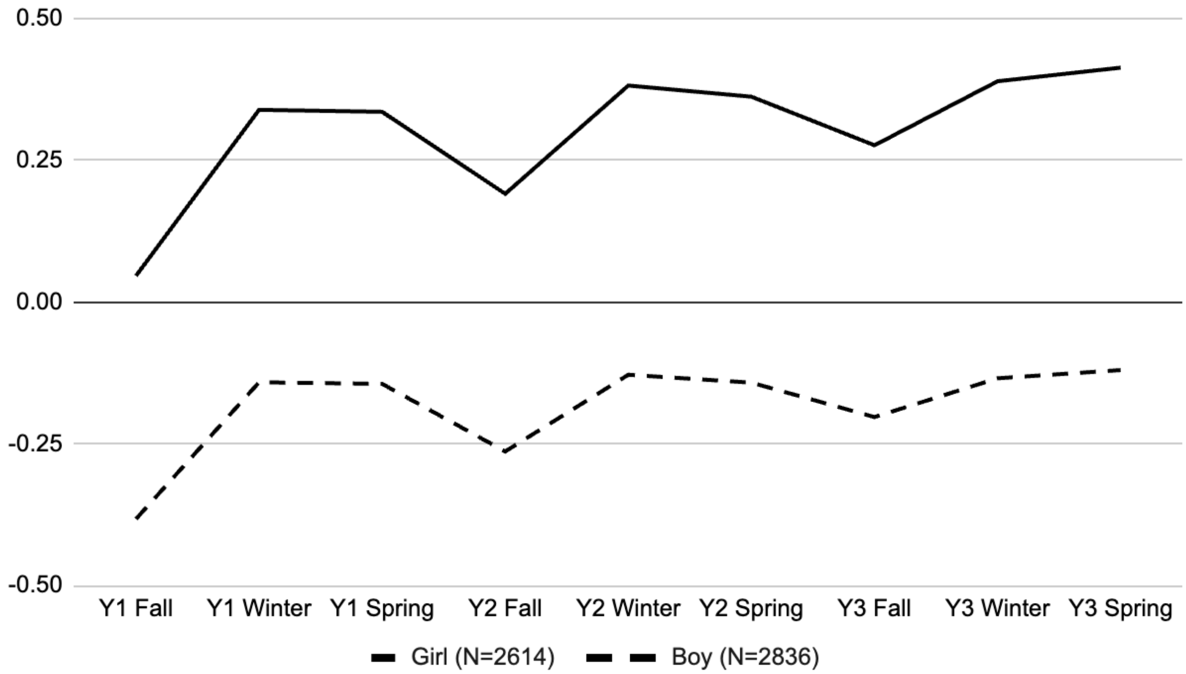


Figure 3a

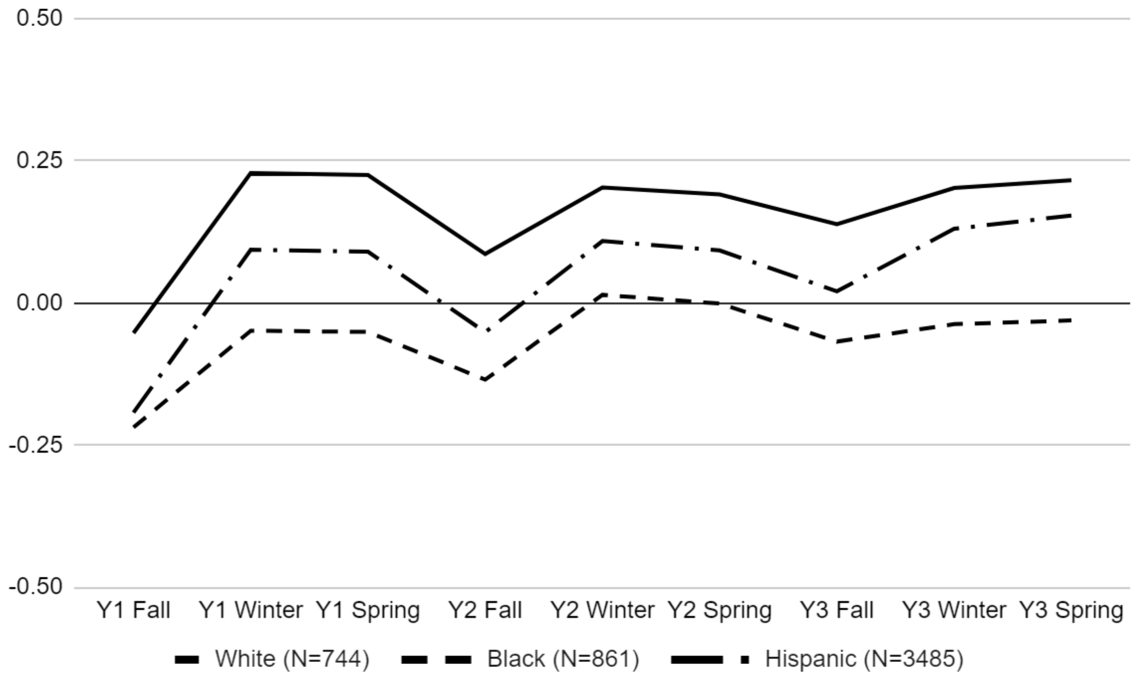
Marginal Means of SEC Over Time Across Gender Subgroups



Note. Y-axis indicates the relative group difference in SD units.

Figure 3b

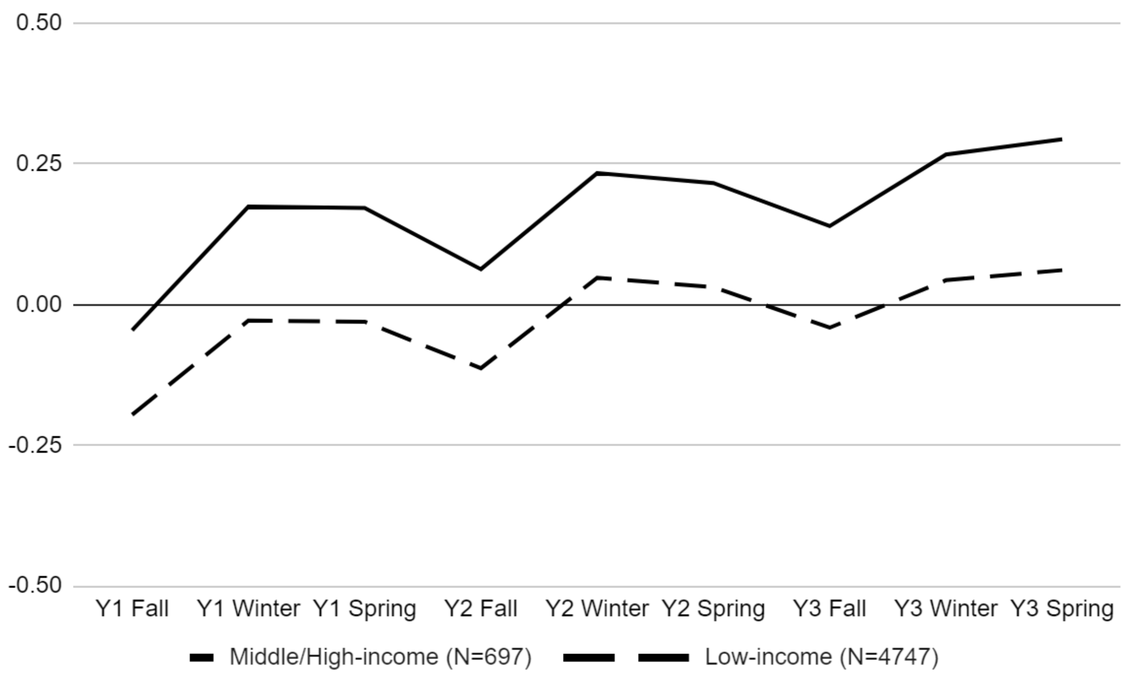
Marginal Means of SEC Over Time Across Racial-Ethnic Subgroups



Note. Y-axis indicates the relative group difference in SD units.

Figure 3c

Marginal Means of SEC Over Time Across Socioeconomic Subgroups



Note. Y-axis indicates the relative group difference in SD units.

Appendix A

Table A1

Model Specification and Estimation Results of the Unconditional Piecewise Growth Model

	Estimate	*SE*	*p*	Estimate	SE	*p*	Estimate	*SE*	*p*
Coefficients of the Slope factor									
	$Slope_{Y1}$			$Slope_{Y2}$			$Slope_{Y3}$		
SEC_{Y1Fall}	0.00	—	—	0.00	—	—	0.00	—	—
$SEC_{Y1Winter}$	1.04	0.17	< .001	0.00	—	—	0.00	—	—
$SEC_{Y1Spring}$	1.00	—	—	0.00	—	—	0.00	—	—
SEC_{Y2Fall}	0.50	—	—	0.00	—	—	0.00	—	—
$SEC_{Y2Winter}$	0.50	—	—	1.07	0.12	< .001	0.00	—	—
$SEC_{Y2Spring}$	0.50	—	—	1.00	—	—	0.00	—	—
SEC_{Y3Fall}	0.50	—	—	0.50	—	—	0.00	—	—
$SEC_{Y3Winter}$	0.50	—	—	0.50	—	—	0.58	0.16	< .001
$SEC_{Y3Spring}$	0.50	—	—	0.50	—	—	1.00	—	—
Growth Factor Parameter Estimates									
	M			Variances					
Intercept	0.00	—	—	0.56	0.03	< .001			
$Slope_{Y1}$	0.23	0.03	< .001	0.10	0.00	< .001			
$Slope_{Y2}$	0.14	0.02	< .001	0.24	0.05	< .001			
$Slope_{Y3}$	0.12	0.02	< .001	0.50	0.14	< .001			
Covariances									
	w/ $Slope_{Y1}$			w/ $Slope_{Y2}$			w/ $Slope_{Y3}$		
Intercept	0.01	0.02	0.66	-0.00	0.03	0.98	-0.08	0.02	< .001
$Slope_{Y1}$				-0.07	0.03	0.04	-0.01	0.02	0.81
$Slope_{Y2}$							-0.02	0.02	0.31

Residual Variances of SEC Latent Variables				*Expected Values of SEC Latent Variables*			
SEC_{Y1Fall}	0.46	0.03	< .001	SEC_{Y1Fall}	0.00	—	—
$SEC_{Y1Winter}$	0.37	0.03	< .001	$SEC_{Y1Winter}$	0.24	0.04	< .001
$SEC_{Y1Spring}$	0.28	0.02	< .001	$SEC_{Y1Spring}$	0.23	0.03	< .001
SEC_{Y2Fall}	0.55	0.02	< .001	SEC_{Y2Fall}	0.12	0.01	< .001
$SEC_{Y2Winter}$	0.30	0.03	< .001	$SEC_{Y2Winter}$	0.26	0.03	< .001
$SEC_{Y2Spring}$	0.27	0.02	< .001	$SEC_{Y2Spring}$	0.25	0.03	< .001
SEC_{Y3Fall}	0.52	0.02	< .001	SEC_{Y3Fall}	0.18	0.02	< .001
$SEC_{Y3Winter}$	0.37	0.04	< .001	$SEC_{Y3Winter}$	0.25	0.04	< .001
$SEC_{Y3Spring}$	0.06	0.13	.660	$SEC_{Y3Spring}$	0.30	0.03	< .001

Note. The coefficients for the intercept factor were fixed at 1. The mean of the intercept factor was fixed at 0 and the mean of each slope factor indicates the mean change scores between Fall and Spring within the corresponding year relative to the *SD* of SEC scores at baseline. For simplicity, the results for the first-order components of the model are not presented here.