School Mental Health (2023) 15:1076–1089
https://doi.org/10.1007/s12310-023-09603-z

ORIGINAL PAPER

Student Demographics as Predictors of Risk Placements via Universal Behavioral Screening

Heather E. Ormiston[1] · **Tyler L. Renshaw**[2]

Accepted: 24 July 2023 / Published online: 18 August 2023

Abstract
Universal screening for social, emotional, and behavioral risk is an important method for identifying students in need of additional or targeted support (Eklund and Dowdy in School Mental Health 6:40–49, 2014). Research is needed to explore how potential bias may be implicated in universal screening. We investigated student demographics as predictors of being placed at risk via a teacher-report measure: the Social, Academic, and Emotional Behavior Risk Screener as reported by Kilgus et al. (in: Theodore J. Christ et al. (eds) Social, academic, and emotional behavior risk screener (SAEBRS), 2014). Results indicated student demographics, including sex, special education status, free/reduced price lunch status, and identification as a student of color, were statistically significant predictors across multiple SAEBRS risk placements. The predictive power of student demographics was meaningful when evaluated independently (i.e., when each characteristic was considered separately with each risk placement) as well as when evaluated relatively or dependently (i.e., when all characteristics were taken together as a set to predict each risk placement). We discuss findings in the context of implications for implementation of universal behavioral screening amidst potential bias and serving students with identified levels of social, emotional, and behavioral risk.

Keywords Assessment · Social–emotional · Mental health services · Diversity · Social justice

Introduction

Universal screening for social, emotional, and behavioral (SEB) risk is becoming increasingly prevalent in today's schools (Romer et al., 2020) and is considered a more effective method for identifying students in need of socioemotional and behavioral support compared to traditional identification models (Eklund & Dowdy, 2014). Within a multi-tiered system of support (MTSS), universal screening serves to identify students at risk for SEB concerns, who may benefit from targeted or intensive supports. Screening also helps to establish base rates of SEB concerns in a school, which "represent the proportion of at-risk students across the entire school and within each classroom (respectively) that could be reasonably served via individual- or group-level Tier 2 interventions" (Kilgus & Eklund, 2016, p. 124). Although there is documented importance of utilizing universal screening measures through an MTSS framework (Splett et al., 2018; Verlenden et al., 2021), the practice is only utilized in approximately 5–15% of schools (Bruhn et al., 2014b; Dineen et al., 2022; Wood & Ellis, 2022) and is implemented more frequently at the elementary level compared to the secondary level (Dineen et al., 2022). When universal screening measures are utilized compared to traditional "wait-to-fail" methods (e.g., teacher referral for evaluation due to the display of significant behavioral problems over time), more students are identified as at-risk for SEB problems, increasing the likelihood of providing early intervention services sooner (Eklund & Dowdy, 2014; Splett et al., 2018).

Potential Rater Bias in Universal Screening

Despite the importance and utility of universal SEB screening (Dever et al., 2015), differences exist in the ratings provided by teachers compared to other raters (Dowdy & Kim, 2012). For instance, students rate themselves at-risk for

✉ Heather E. Ormiston
ormiston@indiana.edu

1 Department of Counseling and Educational Psychology, Indiana University Bloomington, Bloomington, IN, USA

2 Psychology Department, Utah State University, Logan, UT, USA

internalizing problems at higher rates than teachers rate their students (Margherio et al., 2019). One study found student self-report measures were employed in approximately 5% of schools (Wood & Ellis, 2022), even with empirical evidence suggesting students can provide valuable information during the screening process via self-report measures (Romer et al., 2020). Compared with student self-report screeners, teachers' referrals and ratings often end up identifying more externalizing behaviors (Bruhn et al., 2014a; Dowdy et al., 2016). This is concerning, considering youth with internalizing behaviors are often under-referred for mental health services in school, yet may demonstrate significant difficulties related to social isolation, low self-esteem, and suicidal ideation (Allen et al., 2019).

Although differences in ratings would ideally be attributable to *actual* differences in student SEB functioning, it appears teacher bias may indeed be contributing, at least in part, to universal screening ratings (Zakszeski et al., 2023). Teacher-report screeners were initially deemed a less biased measure of student functioning (e.g., Dever et al., 2016); however, recent evidence suggests teacher ratings on universal screeners may be biased by specific demographic factors of the students they are rating (Fallon et al., 2023; Margherio et al., 2019) as well as by the teachers' demographics (Splett et al., 2018; Zakszeski et al., 2023). Additionally, teacher burn-out and teacher self-efficacy have been found to account for a significant proportion of variance in teacher-reported universal screening ratings (McLean et al., 2019). Such findings suggest several forms of implicit bias may be influencing risk placement via screening. Like the examination of implicit bias in relation to disproportionality in discipline and special education referrals (e.g., Girvan et al., 2017; Shi & Zhu, 2022; Skiba et al., 2006), a better understanding of the ways in which implicit bias may influence teacher-rated universal screeners is needed. Thus, in this study, we investigated how several key student demographics—student sex, race/ethnicity, socioeconomic status, and disability status—may bias (or differentially predict) risk placements via teacher-rated universal screening.

Implicit bias has been defined as "the automatic and unconscious stereotypes that drive people to behave and make decisions in certain ways" (Gilliam et al., 2016, p. 3). In other words, implicit bias may affect the way an individual makes decisions without being consciously aware it is happening (McIntosh et al., 2014) due to the demographic characteristics (e.g., sex, race/ethnicity) an individual presents or is perceived to present (Gilliam et al., 2016). These unconscious and spontaneous associations (Marcucci, 2020), then, lead individuals to respond in ways that cause them to rate or evaluate an individual's behavior differently based on certain characteristics (Gilliam et al., 2016; Girvan et al., 2017; Marcucci, 2020), oftentimes perpetuating institutional and societal injustices within the school setting (Malone et al., 2022). For instance, in a study of early childhood educators, Gilliam et al. (2016) examined implicit biases in educator perceptions of behavior and discipline. Utilizing both eye tracking techniques along with racial priming via standardized vignettes, they found that educators were more likely to visually track Black boys' behaviors (compared to Black girls and white children) when primed to expect challenging behavior in the learning environment. Implicit bias may be especially likely to influence decision making in situations that call for more subjective interpretation of behaviors and require teachers to make "snap judgments," such as when making a disciplinary referral to the office (Girvan et al., 2017).

Racial Biases in Screening

Racial and ethnic minoritized youth, particularly Black boys, experience school differently than their majorized peers (Malone et al., 2022). Implicit bias has long been implicated in racial disproportionality of discipline referrals and exclusionary discipline practices (e.g., Gilliam et al., 2016; Girvan et al., 2017; Shi & Zhu, 2022; Skiba et al., 2002, 2011) and special education referrals and identification (i.e., overidentification of males and racial minority youth, and placement in more restrictive settings; Arms et al., 2008; Dever et al., 2016; Skiba et al., 2006; Wehmeyer et al., 2001). White individuals, including teachers, have been socialized to hold anti-Black biases, leading some educators to hold lower expectations for their marginalized students (Marcucci, 2020). The outcomes of these biases lead to educational inequities for students of color, including lower student achievement, increased rates of dropout, and increased contact with the juvenile justice system (McIntosh et al., 2014).

There are longstanding mental health disparities for racially and ethnically minoritized youth (Moore et al., 2023), and within this population, Black youth in particular have been found to be rated at higher levels of risk on universal screening measures compared to their majoritized peers (e.g., Fallon et al., 2022, 2023; Izumi, 2020). For instance, a recent study found Black students had higher levels of social risk as indicated by teacher ratings on the Social Behavior subscale of the Social, Academic, and Emotional Behavior Risk Screener (SAEBRS) compared with a sample of white and Latinx students (Fallon et al., 2023). Another study examined teacher ratings on the Emotional Behavior subscale of the SAEBRS for over 11,000 students, with results revealing risk ratios >1 for Black and Native American students in the sample and a risk ratio <1 for white students (Izumi, 2020). In other words, Black and Native American students were rated at higher levels of risk on the Emotional Behavior subscale compared to their representation in the sample.

Recent research has explicitly examined how implicit bias may be playing a role in teacher ratings of minoritized students on universal screening measures. Specifically, in a study by Fallon et al. (2022), 30 teachers completed the Assessment of Culturally and Contextually Relevant Supports (ACCReS), a self-report measure of teacher cultural responsiveness in the classroom. The association of ACCReS scores with teacher ratings of students' SEB risk via the SAEBRS was examined. Results indicated higher ACCReS scores (i.e., higher levels of perceived cultural responsiveness) were associated with significantly lower levels of students' Social Behavior risk. Yet, interestingly, there were no significant associations between ACCReS scores and teacher ratings of student risk across the Academic Behavior, Emotional Behavior, and Total Behavior scores. When SAEBRS risk levels were disaggregated by disability status and race, Black students were rated at significantly higher levels of Social Behavior risk on the SAEBRS compared to other students in the sample, and eligibility for special education significantly predicted risk across all SAEBRS scales (Fallon et al., 2022).

Other Demographic Biases in Screening

Examining SEB risk level by sex (e.g., male, female) has also been identified as an important consideration, as sex has been identified as a risk factor for developing emotional and behavioral problems over time (Murrieta & Eklund, 2022). Traditional school-based methods of identifying students in need of SEB support consistently over-represent males (Splett et al., 2018), most likely because males tend to demonstrate more externalizing behaviors, whereas females exhibit more internalizing behaviors (Murrieta & Eklund, 2022; Young et al., 2010). When universal screening procedures are utilized, compared to more traditional models of identification, more female students may be identified as at-risk, albeit still at lower rates compared to males (Splett et al., 2018). Moreover, informant discrepancies on teacher and student ratings of universal screening measures have found females are consistently rated by teachers at lower levels of risk (von der Embse et al., 2019) than when they rate themselves (Margherio et al., 2019; Zakaszeski et al., 2023). Given the socioemotional needs and negative outcomes of students with significant internalizing problems (Allen et al., 2019), examining the effect of sex on predicting risk placements resulting from universal SEB screening tools remains an important area for exploration.

Lower socioeconomic status (SES), as indicated by student eligibility for free and reduce-priced lunch (FRPL), has been associated with increased externalizing and internalizing problems (Letourneau et al., 2011; Murrieta & Eklund, 2022; van Oort et al., 2011). Some evidence has emerged to indicate students eligible for FRPL are more likely to be rated as at-risk by their teachers compared to those not eligible for FRPL (Dever et al., 2016; Margherio et al., 2019; Young et al., 2010). Students eligible for FRPL were also found to rate themselves at moderate and higher levels of risk on a universal screening measure compared to those not eligible for FRPL (Margherio et al., 2019). It may be that students eligible for FRPL experience additional stressors that negatively impact their behavior (Young et al., 2010), or it could be that implicit bias is again influencing teacher ratings of student risk (Margherio et al., 2019; Young et al., 2010). Little research has been done examining the effect of SES on predicting risks levels resulting from universal socioemotional screeners (Iaccarino et al., 2019), and there remains a need to further examine the association between FRPL status and risk ratings on universal SEB screening measures (Fallon et al., 2023; Iaccarino et al., 2019).

Considering continued disproportionate representation of students of color in special education, universal screening can be used not just as means to reduce disproportionality (Dever et al., 2016) but also to examine how students receiving services are faring from a socioemotional perspective. Indeed, students eligible for special education services are more likely to demonstrate poorer academic and socioemotional outcomes over time compared with students who are not eligible for special education (McCormick et al., 2019). For instance, one study found pre-kindergarten students receiving special education services were rated significantly lower on an observational measure of socioemotional competence compared with students receiving regular education (Mondi & Reynolds, 2021). Other evidence has indicated students in special education were rated significantly lower (i.e., indicating higher levels of risk) in academic and emotional functioning on a teacher-report universal screening measure compared with those not eligible for services (von der Embse et al., 2019). Akin to students experiencing lower SES, it is likely students in special education are already demonstrating deficits in one or more areas of functioning that led them to being identified with a disability. However, given the overrepresentation of Black and male students in some areas of special education—specifically eligibility areas that are more subjective in nature and implicate SEB functioning (Dever et al., 2016; Skiba et al., 2006)—it is possible that bias is again at play when examining teacher-rated SEB risk outcomes on universal screening measures for students eligible for special education.

Purpose of the Present Study

Concern about universal screening leading to false positives for minoritized youth due to teachers' implicit bias has been identified as a real concern, as it may perpetuate already existing disparities due to race/ethnicity and other marginalized demographics that lead youth to be removed from the classroom more often to receive unnecessary intervention (Fallon et al., 2022). Compounding the issue is the fact that experiences of racism and microaggressions in schools can have deleterious effects on youth mental health (Malone et al., 2022), when the mental health needs of youth are the very aspect of functioning being examined when conducting universal screening (Romer et al., 2020). Therefore, teachers may be unknowingly and unintentionally perpetuating discriminatory practices via screening because of the implicit biases inherent within conducting these screenings. As Fallon et al., (2022, p. 1950) suggest, "[s]tudents' inequitable (or overrepresented) access to supports, particularly low-quality supports, may lead to judgments about nonresponse in an MTSS context" and may ultimately contribute to the continued disproportionate representation of racial and ethnic minority youth in special education and other targeted support systems within MTSS. Researchers continue to advocate for further examination in this area to prevent against perpetuating disparities in educational programming for racially and ethnically minoritized youth (e.g., Fallon et al., 2022; von der Embse et al., 2019), yet more work is clearly needed.

Prior research has investigated increased risk for SEB problems for various subgroups of students as a function of their demographic characteristics, including sex, race/ethnicity, FRPL status, and those receiving special education services (Dever et al., 2015). However, only a handful of studies have examined student characteristics and differences that may predict risk placements across different domains of SEB functioning (e.g., social vs. academic vs. emotional; Fallon et al., 2023; Iaccarino et al., 2019; Izumi, 2020; Margherio et al., 2019). Therefore, the primary purpose of the present study was to investigate the association between student demographics and the outcome of being placed at risk via a commonly used teacher-rated universal screener: the SAEBRS (Kilgus et al., 2014). Our student demographics of interest were sex, student of color, special education, and FRPL. Our screening outcomes of interests were all possible risk placements derived via the several scales of the teacher-report SAEBRS: Social Behavior, Academic Behavior, Emotional Behavior, and Total Behavior. We were interested in both the *independent value* of student demographics as predictors of all possible SAEBRS risk placements as well as the *dependent and relative value* of student demographics when taken together as a set of predictors for all possible SAEBRS risk placements. For the independent framing, we asked "Does [insert each student demographic variable] predict being placed at risk by the SAEBRS for [insert each scale/risk type]?" For the dependent and relative framing, we asked "What is the overall predictive value of student demographics (taken together as a set) as well as the relative value of each particular student demographic (considered within the set) for predicting being placed at risk by the SAEBRS for [insert each scale/risk type]?".

Considering previous research demonstrating associations among youth demographics and mental health outcomes, we hypothesized that each demographic variable would have meaningful relationships with being placed at risk via the SAEBRS. More specifically, for most SAEBRS risk types, we predicted that being categorized as male (reference = female), a student of color (reference = white only), receiving special education services (reference = not receiving special education), and being eligible for FRPL (reference = not eligible) would be associated with increased odds of being placed at risk across all possible SAEBRS scales (i.e., Social Behavior, Academic Behavior, Emotional Behavior, and Total Behavior). The one exception to this expectation was for Emotional Behavior, as we predicted that being categorized as male (reference = female) would be associated with lower odds of being placed at risk. Although we expected the dependent predictive value of student demographics (taken together as a set) would explain a meaningful amount of variance in each of the SAEBRS risk placement outcomes, we did not have predictions about the magnitude of these effects. Moreover, although we expected the relative predictive value of each particular demographic variable (considered within the set) would be attenuated when compared to their independent associations with SAEBRS risk placements, we did not have predictions about which demographic variables might emerge as more or less substantive predictors (both independently and within the set) for each SAEBRS scale or risk type.

Methods

Participants and Setting

The current study examined student universal screening data from four schools in a small, Midwestern district from the Fall 2021 benchmark period. The district enrolls approximately 2700 students in a locale classified as "suburb: midsize" (National Center for Education Statistics, n.d.). Approximately 90% of the students within the district are identified as White, approximately 33% of students are eligible for FRPL, and 17% of students are identified as eligible for special education services. The first author and the district in which the data were collected are presently engaged in a collaborative partnership via a US Department of Education funded grant. Upon receipt of grant funds, the district began implementation of a trauma-informed MTSS and corresponding universal SEB screening procedures using the SAEBRS. All study procedures adhered to the university's Institutional Review Board (IRB), and de-identified data were shared with the first author upon IRB approval.

Participant demographic characteristics are presented in Table 1. The present study consisted of 2,529 students (49% female) with 94% of the sample identifying as white, 19% of the sample eligible to receive special education services, and 29% of the sample eligible for FRPL. Student demographic variables were nominal and binary in nature, and each was based on school-reported data. *Sex* consisted of *female* or *male* classifications; *special education* and *FRPL* both consisted of *eligible* or *not eligible* classifications. *Student of color* consisted of *student of color* or *white only* classifications, which were derived from school-reported data on student race/ethnicity. Although school records contained more specific racial/ethnic classification data for students, the partnering district required that only binary data be used for research purposes to maintain student anonymity. Thus, all students with school-reported racial/ethnic data that differed from *white only* were collapsed into the *student of color* classification.

Table 1 Student demographic characteristics

Demographic/characteristic	*n*	%
School level		
Primary	623	25%
Intermediate	562	22%
Junior High	569	22%
High	775	31%
Grade level		
K	230	9%
1st	228	9%
2nd	165	7%
3rd	180	7%
4th	189	7%
5th	193	8%
6th	206	8%
7th	173	7%
8th	190	8%
9th	208	8%
10th	196	8%
11th	188	7%
12th	183	7%
Sex		
Female	1238	49%
Male	1291	51%
Student of color		
White only	2364	94%
Student of color	164	6%
Special Education		
Not eligible	2059	81%
Eligible	470	19%
Free/reduced price lunch		
Not eligible	1805	71%
Eligible	724	29%

$N = 2529$

Measure

The teacher-rated SAEBRS is a brief, 19-item, norm-referenced screening tool used to identify students at risk for SEB concerns across three subscales—Social Behavior (6 items), Academic Behavior (6 items), and Emotional Behavior (7 items)—which sum to create an overall Total Behavior score (Kilgus et al., 2014, 2018). Items are arranged along a four-point response scale: 0 = *never*, 1 = *sometimes*, 2 = *often*, 3 = *almost always*. Numerous studies have demonstrated the psychometric properties of the SAEBRS (Kilgus et al., 2018; Kilgus et al., 2014; Kilgus et al., 2013; von der Embse et al., 2016). Higher scores on each subscale and the total scale indicate more adaptive behavior, and thus, lower risk, while lower scores indicate higher levels of risk (Kilgus et al., 2018). The Social Behavior subscale captures student behaviors related to developing and maintaining relationships, the Academic Behavior subscale assesses behaviors that support one's ability to effectively participate in academic instruction, and the Emotional Behavior subscale examines behaviors related to a student's ability to regulate emotions (Illuminate Education, 2022a).

In the present study, SAEBRS risk placements were derived using a multi-step process. First, scores were calculated for the Social Behavior, Academic Behavior, Emotional Behavior, and Total Behavior scales using standard scoring rules. Following, classifications of *high risk*, *some risk*, and *low risk* were assigned to these scale scores using interpretation rules associated with the 2021 SAEBRS norms and benchmarks (Illuminate Education, 2022b; see Table 2 for classification base rates). Given the relatively low base rates of students classified as *high risk* across each of the SAEBRS scales (see Table 2) and considering the context of our intended analytic approach (see below), we decided to collapse the *some* and *high* classifications for each risk type into a more general *placed at risk* classification to facilitate better powered analyses. Thus, our resulting SAEBRS risk placements for analyses were nominal and binary in nature, consisting of *placed at risk* or *not placed at risk* classifications for Social Behavior, Academic Behavior, Emotional Behavior, and Total Behavior. Resulting base rates for being placed at risk (i.e., combining *some risk* and *high risk* classifications) across all SAEBRS scales were as follows: Social Behavior $n = 384$, 15.2%; Academic Behavior $n = 490$, 19.4%; Emotional Behavior $n = 494$, 19.5%; and Total Behavior $n = 357$, 14.1%.

Table 2 SAEBRS risk level base rates

SAEBRS scale/risk level	*n*	%
Social behavior		
Low risk	2145	85%
Some risk	307	12%
High risk	77	3%
Academic behavior		
Low risk	2039	81%
Some risk	386	15%
High risk	104	4%
Emotional behavior		
Low risk	2035	80%
Some risk	375	15%
High risk	119	5%
Total behavior		
Low risk	2172	86%
Some risk	316	12%
High risk	41	2%

$N = 2529$

Procedures

As part of standard practice in the district, classroom teachers were asked to complete the SAEBRS for each student using a secure online portal via the FastBridge data management system during the Fall 2021 benchmark screening window (i.e., early- to mid-September 2021), approximately six to seven weeks after the start of the academic year. Classroom teachers in two elementary schools (i.e., grades K–2, grades 3–5) and homeroom teachers at the middle school (i.e., grades 6–8) and high school (i.e., grades 9–12) completed the SAEBRS. Thus, only one teacher completed the SAEBRS for each student. In the secondary grades, homeroom teachers solicited input about students from grade level teams in order to complete the SAEBRS. Teachers were provided an assessment window of approximately 10 days in which to complete the ratings. Ratings took approximately 3–5 min per student to complete. Teacher ratings of students in grades K–12 were used for analysis.

Data Analyses

To evaluate the *independent value* of student demographics for predicting being placed at risk via the SAEBRS, we conducted a series of Chi-square tests of association modeling each demographic variable (i.e., sex, special education, student of color, and FRPL) with each SAEBRS risk variable (i.e., Social, Academic, Emotional, and Total). Strength of associations was evaluated via resulting odds ratios (OR) for each analysis, which were interpreted via guidelines by Chen et al. (2010) for a high base rate ($\geq 10\%$) condition: 1.46–2.49 = small, 2.50–4.13 = moderate, 4.14 + = large. For interpreting OR, reference groups for demographic predictors were set as follows: *sex* = female, *special education* = not eligible, *student of color* = white only, *FRPL* = not eligible. And reference groups for all SAEBRS risk types were set as *not placed at risk*. Thus, OR were interpreted as the odds that students classified within a particular demographic, when compared against their reference group, would be placed at risk via the SAEBRS.

To evaluate the *dependent and relative value* of student demographics (taken together as a set) for predicting being placed at risk via the SAEBRS, we conducted a series of multilevel binomial logistic regressions with each SAEBRS risk placement (i.e., Social Behavior, Academic Behavior, Emotional Behavior, and Total Behavior) modeled as the outcome variable. Level 1 (fixed effects) within each model consisted of the full set of student-level demographic variables (i.e., sex, special education, student of color, and FRPL). Level 2 (random effect) within each model accounted for the nested nature of the student data. Although ratings were completed by teachers and, thus, screening data were most proximally clustered at the teacher/classroom level, we were unable to model teachers/classrooms at Level 2 because this information was unavailable (and unobtainable) in the dataset. Given the available data from the partnering school district, we instead tested both *school level* (i.e., primary, intermediate, junior high, high) and *grade level* (K–12) at Level 2 for all models. Intraclass correlations (ICC) for grade level were universally stronger than school level for all models; thus, grade level was selected as the preferred clustering variable for the final models. Results were evaluated at both the overall model level and the individual predictor level (within each model). At the model level, we evaluated ICC values to understand the variance explained by clustering and model R^2 to understand the variance accounted for in the outcome by the set of dependent demographic predictors. At the predictor level, we evaluated OR to understand the relative predictive value of each demographic variable in relation to all other predictors. Reference groups and magnitude guidelines for interpreting OR were the same as those described for earlier analyses.

Results

Preliminary Analyses

Chi-square tests of association were conducted among all demographic predictors to explore their interrelationships prior to inclusion within primary analyses. Results showed *sex* had a statistically significant and small association with *special education* ($\chi^2 = 51.4$, $df = 1$, $p < 0.001$; OR [95% CI] = 2.13 [1.73, 2.63]) but no meaningful association with *student of* *color* ($\chi^2 = 0.002$, $df = 1$, $p = 0.968$; OR [95% CI] = 1.01 [0.73, 1.38]) nor with *FRPL* ($\chi^2 = 1.39$, $df = 1$, $p = 0.238$; OR [95% CI] = 1.11 [0.93, 1.32]). *Special education* had a statistically significant and small association with *FRPL* ($\chi^2 = 43.7$, $df = 1$, $p < 0.001$; OR [95% CI] = 2.00 [1.63, 2.47]) but no meaningful association with *student of color* ($\chi^2 = 0.01$, $df = 1$, $p < 0.919$; OR [95% CI] = 0.98 [0.65, 1.47]). Finally, *student of color* did not have a statistically significant nor meaningful association with *FRPL* ($\chi^2 = 3.21$, $df = 1$, $p = 0.073$; OR [95% CI] = 1.36 [0.97, 1.89]). Taken together, only 2/6 analyses showed small meaningful associations among demographics, suggesting each demographic variable could be maintained as an independent predictor for the purposes of primary analyses.

We also conducted Chi-square tests of association among all SAEBRS risk placements to explore their interrelationships prior to primary analyses. Results showed risk placement for Social Behavior had a statistically significant and large association with being placed at risk for Academic Behavior ($\chi^2 = 627$, $df = 1$, $p < 0.001$; OR [95% CI] = 15.5 [12.10, 19.97]), Emotional Behavior ($\chi^2 = 399.4$, $df = 1$, $p < 0.001$; OR [95% CI] = 8.89 [7.01, 11.28]), and Total Behavior ($\chi^2 = 1125.3$, $df = 1$, $p < 0.001$; OR [95% CI] = 49.69 [36.79, 67.12]). Risk placement for Academic Behavior had a statistically significant and large association with being placed at risk for Emotional Behavior ($\chi^2 = 620.5$, $df = 1$, $p < 0.001$; OR [95% CI] = 13.41 [10.64, 16.91]) and Total Behavior ($\chi^2 = 1292.7$, $df = 1$, $p < 0.001$; OR [95% CI] = 94.81 [65.67, 136.88]). Likewise, being placed at risk for Emotional Behavior had a statistically significant and large association with risk placement for Total Behavior ($\chi^2 = 934$, $df = 1$, $p < 0.001$; OR [95% CI] = 34.76 [25.97, 46.52]). Overall, these findings suggest students placed at risk via the SAEBRS on any one scale had very strong odds of being placed at risk on the other scales. To determine if these risk placements were independent enough to use as outcome variables in separate models for the primary analyses, we also evaluated the phi (ϕ) coefficients for each association, which are interpreted similar to Pearson's *r* correlation coefficients. Results were as follows: Social–Academic $\phi = 0.50$, Social–Emotional $\phi = 0.40$, Social–Total $\phi = 0.67$, Academic–Emotional $\phi = 0.50$, Academic–Total $\phi = 0.71$, and Emotional–Total $\phi = 0.61$. Squaring these phi coefficients indicated the shared variance among SAEBRS risk placements ranged from 16–50%, suggesting each pair of placements maintained enough unique variance ($\geq 50\%$) to warrant modeling as separate outcomes within primary analyses.

Primary Analyses

Independent Value of Demographic Predictors

Results from the series of Chi-square tests of association conducted with student demographic variables and all possible SAEBRS risk placements (i.e., Social Behavior, Academic Behavior, Emotional Behavior, and Total Behavior) are presented in Table 3. Most models were statistically significant ($p < 0.05$), except for *student of color* with Emotional Behavior and Total Behavior, respectively ($p > 0.05$). Moreover, most statistically significant models (10/16) yielded small effect sizes, as indicated by OR magnitude. That said, a *negligible* OR was observed for the statistically significant association of *sex* with Emotional Behavior, whereas *moderate* OR were observed for *special education* with Academic Behavior, Emotional Behavior, and Total Behavior, respectively (see Table 3). Comparing OR magnitudes across demographics, *special education* had the strongest associations with SAEBRS risk placements, whereas *student of color* had the weakest associations with most SAEBRS risk placements (sans Emotional Behavior, where *sex* was weakest). The magnitude of OR for *sex* and *FRPL* with SAEBRS risk placements generally ranged between the OR for *special education* (higher bound) and *student of color* (lower bound; see Table 3). Taken together, findings suggest all student demographic variables predicted meaningfully increased odds of being placed at risk across SAEBRS scales.

Dependent and Relative Value of Demographic Predictors

Multilevel binomial logistic regression models were conducted to evaluate the dependent and relative value of demographic variables (taken as a set at Level 1) for predicting being placed at risk via the SAEBRS (dependent/outcome variables), while accounting for the influence of grade-level nesting (Level 2). Model level results for each model indicated both Level 1 (fixed effects) and Level 2 (random effect) variables made meaningful contributions toward predicting SAEBRS risk placements. Specifically, for Social Behavior: marginal $R^2 = 0.07$, conditional $R^2 = 0.08$, ICC = 0.02; for Academic Behavior: marginal $R^2 = 0.10$, conditional $R^2 = 0.14$, ICC = 0.04; for Emotional Behavior: marginal $R^2 = 0.05$, conditional $R^2 = 0.11$, ICC = 0.06; and for Total Behavior: marginal $R^2 = 0.10$, conditional $R^2 = 0.12$, ICC = 0.02. Evaluating the marginal R^2 values across models, student demographic predictors (taken as a set at Level 1) explained a small to moderate amount of variance in each SAEBRS risk placement, and they explained relatively more variance in the Academic Behavior and Total Behavior models compared with the Social Behavior and Emotional Behavior models, respectively. Furthermore, the grade level nesting effect (Level 2) accounted for a meaningful amount of variance in all models, yet it explained relatively more variance for the Emotional Behavior and Academic Behavior models compared with the Social Behavior and Total Behavior models, respectively.

Table 3 Independent Associations of Student Demographics with SAEBRS Risk Placements

Demographic/SAEBRS scale	χ^2	*df*	*p*	OR [95% CI]	OR Magnitude
Sex					
Social behavior	38.5	1	< .001	2.03 [1.62, 2.55]	Small
Academic behavior	37.5	1	< .001	1.88 [1.53, 2.30]	Small
Emotional behavior	6.71	1	.01	1.30 [1.07, 1.58]	Negligible
Total behavior	32.3	1	< .001	1.95 [1.54, 2.46]	Small
Special education					
Social behavior	36.9	1	< .001	2.12 [1.66, 2.72]	Small
Academic behavior	102	1	< .001	3.04 [2.43, 3.80]	Moderate
Emotional behavior	72.8	1	< .001	2.60 [2.08, 3.25]	Moderate
Total behavior	92.9	1	< .001	3.18 [2.49, 4.06]	Moderate
Student of color					
Social behavior	4.18	1	.041	1.50 [1.01, 2.23]	Small
Academic behavior	4.35	1	.037	1.47 [1.02, 2.12]	Small
Emotional behavior	3.32	1	.068	1.41 [0.97, 2.03]	Negligible
Total behavior	2.52	1	.113	1.39 [0.92, 2.11]	Negligible
Free/reduced price lunch					
Social behavior	17.4	1	< .001	1.62 [1.29, 2.03]	Small
Academic behavior	44.2	1	< .001	1.99 [1.62, 2.45]	Small
Emotional behavior	22.3	1	< .001	1.64 [1.34, 2.02]	Small
Total behavior	29.2	1	< .001	1.88 [1.49, 2.37]	Small

OR = odds ratio. Reference groups for demographic predictors: *sex* = female, *special education* = not eligible, *student of color* = white only, *free/reduced price lunch* = not eligible. Reference group for all SAEBRS risk placements = not placed at risk

Predictor level results for the fixed effects within all models are presented in Table 4. Most student demographic predictors were statistically significant ($p < 0.05$) across SAEBRS risk placement models, except for *student of color* in the Social Behavior and Total Behavior models as well as *sex* in the Academic Behavior model. OR magnitude for all demographic predictors was consistently small for the Social Behavior model, ranged from small to moderate for the Academic Behavior model, was negligible or small for the Emotional Behavior model, and ranged from negligible to moderate in the Total Behavior model (see Table 4). Evaluating the overlap of 95% CI for the OR of each demographic predictor within each model, the relative value of each demographic for predicting increased odds of risk placements via the SAEBRS was practically equivalent for the Social Behavior and Academic Behavior models. Although the relative value of most predictors was also practically equivalent for the OR yielded by the other models, non-overlapping 95% CI suggest that *special education* was a practically stronger predictor than both *sex* in the Emotional Behavior model and *FRPL* in the Total Behavior model. Like results for the independent predictions (reported above), *special education* remained the most robust predictor of SAEBRS risk placements across dependent models (see Table 4), although effects were somewhat attenuated compared with those observed in the independent models (see Table 3).

Discussion

Universal SEB screening is an essential component of an MTSS framework (Eklund & Dowdy, 2014; Romer et al., 2020) beyond just identifying students at risk for SEB concerns. Screening data can be continuously used as a data-based decision-making tool (National Practitioner Advisory Group, 2019) to determine if students are responding to universal SEL instruction (Kilgus & Eklund, 2016). To this end, screening data can be used to evaluate the strengths of universal classroom- or school-level SEL curriculum and/or to determine if intensification of Tier 1 instruction is needed to reduce SEB risk rates (Center on Multi-Tiered Systems of Support at the American Institutes for Research, 2019; Kilgus & Eklund, 2016). Additionally, given the potential for implicit bias implicated by our findings in this study, universal screening data could be disaggregated by relevant subgroups to determine if all demographics are responding to universal curriculum or if systemic practices, biases, or structural barriers are contributing to disparities (National Practitioner Advisory Group, 2019).

Recently, the literature has emerged to suggest teacher ratings on universal screeners may be biased as a function of specific demographic factors of the students they are rating (Fallon et al., 2023; Margherio et al., 2019), indicating teacher ratings may be contributing, in part, to differences in

Table 4 Fixed effects for binomial logistic regression models

Model/Parameter	*b*	*SE*	*z*	*p*	OR [95% CI]	OR Magnitude
Social Behavior						
Intercept	− 1.38	0.13	− 10.72	< .001	–	–
Sex	0.63	0.12	5.36	< .001	1.88 [1.49, 2.38]	Small
SoC	0.40	0.21	1.93	.053	1.49 [0.99, 2.23]	Small
SPED	0.61	0.13	4.66	< .001	1.84 [1.42, 2.37]	Small
FRPL	0.40	0.12	3.35	< .001	1.49 [1.18, 1.89]	Small
Academic Behavior						
Intercept	− 0.93	0.15	− 6.20	< .001	–	–
Sex	0.55	0.11	5.04	< .001	1.73 [1.40, 2.14]	Small
SoC	0.45	0.20	2.31	0.021	1.57 [1.07, 2.31]	Small
SPED	0.98	0.12	8.20	< .001	2.66 [2.11, 3.37]	Moderate
FRPL	0.59	0.11	5.40	< .001	1.81 [1.46, 2.25]	Small
Emotional behavior						
Intercept	− 0.97	0.16	− 5.98	< .001	–	–
Sex	0.18	0.11	1.67	0.095	1.19 [0.97, 1.47]	Negligible
SoC	0.41	0.19	2.08	0.037	1.50 [1.02, 2.20]	Small
SPED	0.88	0.12	7.35	< .001	2.42 [1.91, 3.07]	Small
FRPL	0.43	0.11	3.90	< .001	1.54 [1.24, 1.91]	Small
Total behavior						
Intercept	− 1.37	0.14	− 10.00	< .001	–	–
Sex	0.57	0.12	4.64	< .001	1.77 [1.39, 2.25]	Small
SoC	0.35	0.22	1.61	0.108	1.42 [0.93, 2.18]	Negligible
SPED	1.02	0.13	7.86	< .001	2.76 [2.14, 3.56]	Moderate
FRPL	0.52	0.12	4.20	< .001	1.68 [1.32, 2.13]	Small

OR = odds ratio. SPED = special education, SoC = student of color, FRPL = free/reduced price lunch. Reference group for all SAEBRS risk outcomes = not placed at risk. Reference groups for demographic predictors: *sex* = female, *special education* = not eligible, *student of color* = white only, *free/reduced price lunch* = not eligible

subgroups and identified levels of risk found across populations screened (Zakszeski et al., 2023). Thus, the purpose of the present study was to examine student demographic characteristics (i.e., sex, student of color, FRPL, and special education status) as predictors for being placed at risk on a universal screening teacher-report measure, the SAEBRS. Across each of the demographic predictors, results showed students in the targeted demographic group (i.e., male, eligible for special education, identified as a student of color, and eligible for FRPL) had greater odds of being placed at risk on the Social Behavior and Academic Behavior scales of the SAEBRS compared to students in the reference demographic groups (i.e., female, not eligible for special education, identified as white only, and not eligible for FRPL). On the Emotional Behavior scale, most demographic predictors were again significant and contributed to higher odds of risk placement for the targeted demographic groups, with the exception of sex, where OR indicated the odds for males and females were practically equivalent. Similarly, on the Total Behavior scale, most of the demographic predictors were significant predictors of increased odds of risk placement, except for being identified as a *student of color*, which had practically equivalent odds as being identified as white only. We should note that while special education status was found to be a moderate predictor of being placed at risk for Academic Behavior and Total Behavior, the 95% confidence intervals are relatively wide and share enough overlap with other demographic predictors to conclude that the relative value of each OR is practically equivalent.

Findings from our study align with emerging research examining student demographic characteristics and levels of SEB risk. Universal screening procedures have consistently identified female students at lower levels of SEB risk (Margherio et al., 2019; von der Embse et al., 2019; Zakszeski et al., 2023) and students eligible for FRPL at higher levels of SEB risk (Dever et al., 2016; Margherio et al., 2019; Young et al., 2010). Additionally, racially and ethnically minoritized youth have been found to be rated at higher levels of risk in the domains of social and emotional functioning (e.g., Fallon et al., 2022, 2023; Izumi, 2020), while students in special education have been shown to have more elevated levels of risk in academic and emotional domains of functioning (von der Embse et al., 2019). While excellent work has been done in recent years to examine relationships between demographic factors and teacher-rated levels of risk, our study is the first, to our knowledge, to concurrently examine four specific demographic factors from one sample as predictors of being placed at risk on a teacher-reported universal screening measure. Overall, our findings add value to existing literature by demonstrating that, when considered both independently and when taken as a set, student sex, race/ethnicity, socioeconomic status, and disability status all contribute to increased odds of being placed at risk across multiple SEB domains (i.e., academic, social, emotional, and total). Thus, multiple teacher-rated SEB risk placements appear to be biased (or differentially predicted) by multiple student demographic factors.

Implications for Practice

Considering students eligible for special education have been found to have poorer academic and socioemotional outcomes over time (McCormick et al., 2019), universal screening measures can provide valuable information when determining a child's individualized programming (Romer et al., 2020). Even though students in special education may demonstrate mental health concerns (Poppen et al., 2016), mental health services may not be provided due to limited personnel and resources (Atkins et al., 2010). To account for the limited resources associated with special education programming, there is some emerging evidence indicating students in special education may be positively impacted by receiving evidence-based, universal curriculum targeting behavioral and socioemotional outcomes (Hart et al., 2021). It is worthwhile to consider the importance of this—not just because we know the benefits of SEL curriculum for all students (Mahoney et al., 2018/2019) but because the provision of related mental health services may not always be provided in a student's individualized education program (Poppen et al., 2016). Because results from our current study indicate a greater likelihood of students eligible for special education were rated at risk on the SAEBRS in social and emotional domains, schools should consider how programming for SEB support may be provided to students receiving special education.

Next, it is worthwhile to highlight the results indicating increased odds of being placed at risk for students eligible for FRPL or being identified as a student of color. The importance of identifying economically or racially/ethnically marginalized students in need of support should not be understated, as they may lack access to and face increased barriers to receiving mental health services outside of school (Malone et al., 2022; Verlenden et al., 2021). However, because implicit bias may be operating when teachers complete ratings for universal screening (Fallon et al., 2022), practitioners should be cognizant of the potential role of bias so that universal screening data are not used to reinforce stereotypes or inadvertently negatively portray marginalized students (Verlenden et al., 2021).

Finally, it is noteworthy that student sex significantly predicted increased odds of being placed at risk for social, academic, and total scales, yet did not significantly predict risk placement on the emotional scale. Because females have been found to be rated at-risk at lower rates compared to males (Splett et al., 2018), and given the context of existing research documenting teacher under-reporting of student emotional risk compared to student self-report (Margherio et al., 2019), particularly for females (Margherio et al., 2019; Zakszeski et al., 2023), this finding should be carefully considered. Internalizing behaviors that go undetected, and untreated, can lead to interpersonal conflicts, reduced academic performance, and potential engagement in risky behaviors such as suicide (Allen et al., 2019; Splett et al., 2019). It is possible that female students' internalizing symptomatology may go undetected by teachers due to a lack of overt, observable symptoms in the classroom or because teachers perceive internalizing symptoms to be less concerning than externalizing problems (Splett et al., 2019). Teachers have expressed a desire for increased training in youth mental health (Reinke et al., 2011), and additional training in recognizing internalizing symptomatology specifically may be especially important as youth with these problems are under-referred for additional mental health supports (Splett et al., 2019).

Taken together, results from the current study emphasize the potential of implicit bias in universal screening and, therefore, the importance of using an equity-focused lens when implementing an MTSS framework (Malone et al., 2022). Practitioners must work to ensure data regarding risk prediction, such as results from the current study are not used to reinforce stereotypes or inadvertently negatively portray marginalized students (Verlenden et al., 2021). Malone et al. (2022) argue that MTSS can serve as an important catalyst for reducing mental health disparities, particularly for minoritized youth, but caution "this framework is only as effective as the interventions and assessments used (Fabiano & Evans, 2019)," warning further that "schools may adopt a 'one size fits all' approach to selecting mental health interventions and not consider students' cultural context" (p. 2442).

Similarly, Moore et al. (2023) advance a notion of equity-focused mental health screening which requires "a shift from individual- and deficit-focused approaches to systems- and holistic-focused approaches that (a) identify strengths and stressors among individuals, groups, and communities; (b) dismantle structural forms of oppression; and (c) promote positive mental health outcomes for minoritized youth" (p. 57). They delineate three core principles that should serve as a foundation to all work related to equitable practices in schools. First, schools must work to disrupt unjust systems rather than perpetuating predictable outcomes of students based on demographic characteristics (e.g., sex, race, ethnicity, SES). Second, schools must work to dismantle the "biased and oppressive school policies, programs, practices, and interactions" (p. 60) that perpetuate and exacerbate systemic inequality. Third, schools must work with minoritized youth and families to develop, implement, and sustain "learning environments that are intellectually and socially safe and affirming for all students and their families" (p. 60). Ultimately, Moore et al. (2023) argue practitioners should be utilizing universal screening measures in a way that considers how student demographic characteristics, such as those examined in the present study, may also be contributing to student strengths and resilience, rather than only considering them from a deficit perspective.

While a comprehensive review of evidence-based strategies for mitigating bias is beyond the scope of this article, we encourage readers to examine the following resources to identify relevant practices that can be implemented in their schools and districts to reduce implicit bias. Specifically, Malone et al. (2022) advocate for equity-focused MTSS, describing Tier 1 interventions for promoting a positive school racial climate and presenting culturally relevant elements to integrate into Tier 2 and Tier 3 interventions. Additionally, Romero et al. (2020) provide a review of three different interventions that have demonstrated evidence to alleviate implicit bias as part of a decision-making process: (a) an empathic mindset intervention aimed at enabling teachers to see the value of students' diverse perspectives and experiences while building and maintaining positive relationships with students (Okonofua et al., 2016); (b) training related to recognizing and replacing stereotyped reactions (i.e., counterstereotyping; Burns et al., 2017); and (c) a "habit-breaking intervention" designed to increase participant awareness and knowledge of bias while also teaching different bias-reduction strategies (Forscher et al., 2017, p.133). However, Romero et al. (2020) caution against widespread implementation of these interventions until additional research can be conducted. Despite the promise of the strategies presented here, "[m]ore work is necessary to establish the conditions under which individuals consciously or unconsciously discriminate in the K–12 setting [in order] to inform policies aimed at curbing these behaviors" (Shi & Zhu, 2022, p.10).

Limitations and Directions for Future Research

The current study must be considered in the context of its limitations. First, the current sample is limited in size and representation, including only one small district in the Midwest with a predominantly white student population. This sample size necessarily limited the number of students identified at the highest level of risk on the SAEBRS, thus leading us to collapse risk placement into a dichotomous variable (i.e., *not placed at risk* vs. *placed at risk*). We therefore do not know if student demographics may perform differently when predicting increasing levels of risk placement, such as *low risk* vs. *some risk* vs. *high risk*. Moreover, the limited representation of racially and ethnically diverse students within the sample, as well as the requirements of the data-sharing agreement with the local district, required that we consolidate students' diverse identities into *student of color* vs. *white only* categories. We therefore could not explore if different racial/ethnic identifies may have differential value in predicting risk placements. To account for these limitations, future research should examine student characteristics as predictors of teacher-rated SEB risk placements across larger, more racially/ethnically diverse samples of student populations. Such large, diverse samples would also have the benefit of allowing for the exploring the effects of intersectionality on screening outcomes by analyzing interactions among demographics (e.g., student of color × FRPL), which may provide a more nuanced understanding of the variables at play.

Next, teacher ratings at the secondary level were limited to homeroom teacher ratings with input from and collaboration with other academic-area teachers. Because of this, the homeroom teachers may have rated the students differently than core teachers, who interact with students in a more academically oriented capacity (e.g., checking for assignment completion across classes vs. delivering direct instruction of course content). The choice of optimal informants for completing screeners at the secondary level is a known challenge (von der Embse et al., 2021) and schools may opt to utilize self-report measures to screen for SEB risk in conjunction with teacher report (Dowdy & Kim, 2012). However, simply adding an informant to any benchmark screening window may not address the discrepancies (e.g., see Margherio et al., 2019). Additional research warrants examination of these predictors across time, informant (e.g., teacher report, self-report), as well as instrument to determine if the use of multiple informants mitigates these discrepancies. Relatedly, future research could benefit from parsing the effects of teacher informants from grade-level or school-building effects on student screening outcomes. Given we modeled grade level as a nesting variable in this study, we could not then explore its value as a predictor. Yet future studies could make different modeling choices in order to intentionally compare the predictive value of educational contexts with student demographics.

Further, generalization of our findings may be limited considering the larger context of our study, wherein SAEBRS ratings were taken at one point in time from a district that received a sizable grant to implement trauma-informed MTSS. It is possible that teachers rated students differently, perhaps more favorably or perhaps with more awareness/

discernment of signs of mental health challenges, after having undergone grant-related training related to student mental health and trauma-informed practices. Moreover, due to the small number of students eligible to receive special education services in the sample, as well as the uneven distribution of special education eligibility classifications within the sample (i.e., far more students were classified with specific learning disability compared with all other classifications), we chose to investigate special education status as a dichotomous variable (i.e., *not eligible* vs. *eligible*) and thus, did not examine potential differences in eligibility categories for predicting being placed at risk. Future research may account for this limitation by collecting much larger samples or by intentionally over-sampling for students with disabilities that have lower base rates within special education (e.g., emotional and intellectual disabilities).

Finally, we acknowledge that the most significant limitation of our study is that our research design did not allow for directly validating our assumption of implicit bias; rather, we infer teacher bias as a function of the differences in SEB risk placements we observed based on student demographic factors. Thus, when considering our results, we wonder: Are teachers really operating under implicit biases that are influencing their ratings of students' SEB risk? Or are we potentially seeing true differences across SEB functioning for various subgroups of students, which may be a function of educational disparities experienced by marginalized youth? Or, as a third alternative, is the screening instrument (SAEBRS) perhaps biased itself, lacking cultural relevancy or invariance across the demographics screened? Future research is needed to examine these questions and tease apart competing explanations that may more fully explain our results. Implicit bias can be a construct that is difficult to measure (Marcucci, 2020) and this line of research related to universal screening is certainly in its infancy (e.g., Fallon et al., 2023; Margherio et al., 2019). We encourage interested researchers to continue this line of work by examining the implications of biases, disparities, and inequities as they manifest in screening.

Conclusion

Schools play an important role in supporting students' positive socioemotional and behavioral development, particularly for marginalized youth (Malone et al., 2022). Examining students' levels of SEB functioning via universal screening is an important component to the provision of effective prevention and promotion services in an MTSS framework (Eklund & Dowdy, 2014; Splett et al., 2018), and understanding the potential role of bias and inequities in this process is critical (Moore et al., 2023). Our study furthered existing research regarding the contributions of student demographics in predicting SEB risk placements via teacher-rated universal screeners. Consistent with Fallon et al. (2023), we found that student demographics may be contributing to teacher-rated levels of risk on universal screening measures. Overall, our findings add value to the existing literature by demonstrating that, when considered both independently and when taken as a set, student sex, race/ethnicity, socioeconomic status, and disability status all contribute to increased odds of being placed at risk across multiple SEB domains (i.e., academic, social, emotional, and total). Thus, multiple teacher-rated SEB risk placements appear to be biased (or differentially predicted) by multiple student demographic factors. We recognize that much more research is needed in this area, and we especially highlight the need for work that considers student demographics in a manner that incorporates these variables as strengths and predictors of resilience to advance implementation of equity-focused MTSS (Malone et al., 2022; Moore et al., 2023).

Acknowledgements Financial support was provided by the US Office of Elementary and Secondary Education Mental Health Professional Demonstration Grant #S184X190033, but the funding source had no such involvement in the study design, data collection, analysis andinterpretation of data, in writing the report, or in deciding to submit for publication.

Funding This work was supported by the US Office of Elementary and Secondary Education Mental Health Professional Demonstration Grant #S184X190033.

Declarations

Conflict of interest The authors do not disclose a financial interest in this work.

References

Allen, A., Kilgus, S. P., Burns, M. K., & Hodgson, C. (2019). Surveillance of internalizing behaviors: a reliability and validity generalization study of universal screening evidence. *School Mental Health, 11*, 194–209.

Arms, E., Bickett, J., & Graf, V. (2008). Gender bias and imbalance: girls in US special education programmes. *Gender and Education, 20*(4), 349–359.

Atkins, M. S., Hoagwood, K. E., Kutash, K., & Seidman, E. (2010). Toward the integration of education and mental health in schools. *Administration and Policy in Mental Health and Mental Health Services Research, 37*, 40–47.

Bruhn, A. L., Lane, K. L., & Hirsch, S. E. (2014a). A review of tier 2 interventions conducted within multitiered models of behavioral prevention. *Journal of Emotional and Behavioral Disorders, 22*(3), 171–189.

Bruhn, A. L., Woods-Groves, S., & Huddle, S. (2014b). A preliminary investigation of emotional and behavioral screening practices in K-12 schools. *Education & Treatment of Children, 37*(4), 611–634.

Burns, M. D., Monteith, M. J., & Parker, L. R. (2017). Training away bias: the differential effects of counterstereotype training and self-regulation on stereotype activation and application. *Journal of Experimental Social Psychology, 73*, 97–110.

Center on Multi-Tiered Systems of Support at the American Institutes for Research. (2019). *Tips for intensifying instruction at Tier 1.* American Institutes for Research.

Chen, H., Cohen, P., & Chen, S. (2010). How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Communications in Statistics—Simulation and Computation*, *39*(4), 860–864.

Dever, B. V., Dowdy, E., Raines, T. C., & Carnazzo, K. (2015). Stability and change of behavioral and emotional screening scores. *Psychology in the Schools, 52*(6), 618–629.

Dever, B. V., Raines, T. C., Dowdy, E., & Hostutler, C. (2016). Addressing disproportionality in special education using a universal screening approach. *The Journal of Negro Education, 85*(1), 59–71.

Dineen, J. N., Chafouleas, S. M., Briesch, A. M., McCoach, D. B., Newton, S. D., & Cintron, D. W. (2022). Exploring social, emotional, and behavioral screening approaches in US public school districts. *American Educational Research Journal, 59*(1), 146–179.

Dowdy, E., Dever, B. V., Raines, T. C., & Moffa, K. (2016). A preliminary investigation into the added value of multiple gates and informants in universal screening for behavioral and emotional risk. *Journal of Applied School Psychology, 32*(2), 178–198.

Dowdy, E., & Kim, E. (2012). Choosing informants when conducting a universal screening for behavioral and emotional risk. *School Psychology Forum, 6*(4), 1–10.

Illuminate Education. (2022a). *FastBridge SAEBRS.* Retrieved from https://www.illuminateed.com/products/fastbridge/social-emotional-behavior-assessment/saebrs/.

Illuminate Education. (2022b). *SAEBRS and mySAEBRS norms and benchmarks.* Retrieved from https://fastbridge.illuminateed.com/hc/en-us/articles/1260802344370-SAEBRS-and-mySAEBRS-Norms-and-Benchmarks.

Eklund, K., & Dowdy, E. (2014). Screening for behavioral and emotional risk versus traditional school identification methods. *School Mental Health, 6*, 40–49.

Fabiano, G. A., & Evans, S. W. (2019). Introduction to the special issue of school mental health on best practices in effective multi-tiered intervention frameworks. *School Mental Health, 11*, 1–3.

Fallon, L. M., Veiga, M. B., Susilo, A., & Kilgus, S. P. (2023). Do teachers' perceptions of high cultural responsiveness predict better behavioral outcomes for students? *Behavioral Disorders, 48*(2), 97–105.

Fallon, L. M., Veiga, M. B., Susilo, A., Robinson-Link, P., Berkman, T. S., Minami, T., & Kilgus, S. P. (2022). Exploring the relationship between teachers' perceptions of cultural responsiveness, student risk, and classroom behavior. *Psychology in the Schools, 59*, 1948–1964.

Forscher, P. S., Mitamura, C., Dix, E. L., Cox, W. T. L., & Devine, P. G. (2017). Breaking the prejudice habit: Mechanisms, timecourse, and longevity. *Journal of Experimental Social Psychology, 72*, 133–146. https://doi.org/10.1016/j.jesp.2017.04.009

Gilliam, W. S., Maupin, A. N., Reyes, C. R., Accavitti, M., & Shic, F. (2016). *Do early educators' implicit biases regarding ex and race relate to behavior expectations and recommendations of preschool expulsions and suspensions?* Yale University Child Study Center.

Girvan, E. J., Gion, C., McIntosh, K., & Smolkowski, K. (2017). The relative contribution of subjective office referrals to racial disproportionality in school discipline. *School Psychology Quarterly, 32*(3), 392–404.

Hart, S. R., Domitrovich, C., Embry, D. D., Becker, K., Lawson, A., & Ialongo, N. (2021). The effects of two elementary school-based universal prevention interventions on special education students' socioemotional outcomes. *Remedial and Special Education, 42*(1), 31–43.

Iaccarino, S., von der Embse, N., & Kilgus, S. P. (2019). Interpretation and use of the social, academic, and emotional behavior risk screener: A latent transition approach. *Journal of Psychoeducational Assessment, 37*(4), 486–503.

Izumi, J. T. (2020). Detecting and explaining differential item functioning on the social, academic, and emotional behavior risk screener. (Doctoral dissertation, University of Missouri-Columbia).

Kilgus, S. P., Chafouleas, S. M., Riley-Tillman, T. C., & von der Embse, N. P. (2014). *Social, Academic, and Emotional Behavior Risk Screener (SAEBRS)*. Theodore J. Christ & Colleagues.

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*, 210–226.

Kilgus, S. P., & Eklund, K. (2016). Consideration of base rates within universal screening for behavioral and emotional risk: A novel procedural framework. *School Psychology Forum: Research in Practice, 10*(1), 120–130.

Kilgus, S. P., Taylor, C. N., & von der Embse, N. P. (2018). Screening for behavioral risk: Identification of high risk cut scores within the social, academic, and emotional behavior risk screener (SAEBRS). *School Psychology Quarterly, 33*(1), 155–159.

Letourneau, N. L., Duffett-Lelger, L., Levac, L., Watson, B., & Young-Morris, C. (2011). Socioeconomic status and child development: A meta-analysis. *Journal of Emotional and Behavioral Disorders, 21*(3), 211–224.

Mahoney, J. L., Durlak, J. A., & Weissberg, R. P. (2018). An update on social and emotional learning outcome research. *Phi Delta Kappan, 100*(4), 18–23.

Malone, C. M., Wycoff, K., & Turner, E. A. (2022). Applying a MTSS framework to address racism and promote mental health for racial/ethnic minoritized youth. *Psychology in the Schools, 59*(12), 2438–2452.

Marcucci, O. (2020). Implicit bias in the era of social desirability: Understanding antiblackness in rehabilitative and punitive school discipline. *The Urban Review, 52*, 47–74.

Margherio, S. M., Evans, S. W., & Owens, J. S. (2019). Universal screening in middle and high schools: Who falls through the cracks? *School Psychology, 34*(6), 591–602.

McCormick, M. P., Neuhaus, R., Horn, E. P., O'Connor, E. E., White, H. I., Harding, S., Cappella, E., & McClowry, S. (2019). Long-term effects of social–emotional learning on receipt of special education and grade retention: Evidence from a randomized trial of insights. *AERA Open, 5*(3), 1–21.

McIntosh, K., Girvan, E. J., Horner, R. H., & Smolkowski, K. (2014). Education not incarceration: A conceptual model for reducing racial and ethnic disproportionality in school discipline. *The Journal of Applied Research on Children, 5*, 1–22.

McLean, D., Eklund, K., Kilgus, S. P., & Burns, M. K. (2019). Influence of teacher burnout and self-efficacy on teacher-related variance in social-emotional and behavioral screening scores. *School Psychology, 34*(5), 503–511.

Mondi, C. F., & Reynolds, A. J. (2021). Socio-emotional learning among low-income prekindergarteners: The roles of individual factors and early intervention. *Early Education and Development, 32*(3), 360–384.

Moore, S., Long, A. C. J., Coyle, S., Cooper, J. M., Mayworm, A. M., Amirazizi, S., Edyburn, K. L., Pannozzo, P., Choe, D., Miller, F. G., Eklund, K., Bohnenkamp, J., Whitcomb, S., Raines, T. C., & Dowdy, E. (2023). A roadmap to equitable school mental health screening. *Journal of School Psychology, 96*, 57–74.

Murrieta, I., & Eklund, K. (2022). Universal screening to detect emotional and behavior risk among English language learners. *School Psychology Review, 51*(4), 441–453.

National Practitioner Advisory Group. (2019). *Making SEL assessment work: Ten practitioner beliefs*. Collaborative for Academic, Social, and Emotional Learning and the American Institutes for Research.

Okonofua, J. A., Paunesku, D., & Walton, G. M. (2016). Brief intervention to encourage empathic discipline cuts suspension rates in half among adolescents. *Proceedings of the National Academy of Sciences, 113*(19), 5221–5226.

Poppen, M., Sinclair, J., Hirano, K., Lindstrom, L., & Unruh, D. (2016). Perceptions of mental health concerns for secondary students with disabilities during transition to adulthood. *Education and Treatment of Children, 39*(2), 221–246.

Reinke, W. M., Stormont, M., Herman, K. C., Puri, R., & Goel, N. (2011). Supporting children's mental health in schools: Teacher perceptions of needs, roles, and barriers. *School Psychology Quarterly, 26*(1), 1–13.

Romer, N., von der Embse, N., Eklund, K., Kilgus, S., Perales, K., Splett, J. W., Suldo, S., & Wheeler, D. (2020). *Best practices in social, emotional, and behavioral screening: An implementation guide. Version 2.0.* Retrieved from smhcollaborative.org/universalscreening

Romero, L. S., Scahill, V., & Charles, S. R. (2020). Restorative approaches to discipline and implicit bias: Looking for ways forward. *Contemporary School Psychology, 24*, 309–317.

Shi, Y., & Zhu, M. (2022). Equal time for equal crime? Racial bias in school discipline. *Economics of Education Review, 88*, 1–12.

Skiba, R. J., Horner, R. H., Chung, C., 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*, 85–107.

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*, 317–342.

Skiba, R. J., Poloni-Staudinger, L., Gallini, S., Simmons, A. B., & Feggins-Azziz, R. (2006). Disparate access: The disproportionality of African American students with disabilities across educational environments. *Exceptional Children, 72*(4), 411–424.

Splett, J. W., Garzona, M., Gibson, N., Wojtalewicz, D., Raborn, A., & Reinke, W. M. (2019). Teacher recognition, concern, and referral of children's internalizing and externalizing behavior problems. *School Mental Health, 11*, 228–239.

Splett, J. W., Trainor, K. M., Raborn, A., Halliday-Boykins, C. A., Garzona, M. E., Dongo, M. D., & Weist, M. D. (2018). Comparison of universal mental health screening to students already receiving intervention in a multitiered system of support. *Behavioral Disorders, 43*(3), 344–356.

van Oort, F. V. A., van der Ende, J., Wadsworth, M. E., Verhulst, F. C., & Achenbach, T. M. (2011). Cross-national comparison of the link between socioeconomic status and emotional and behavioral problems in youths. *Social Psychiatry and Psychiatric Epidemiology, 46*(2), 167–172.

Verlenden, J., Naser, S., & Brown, J. (2021). Steps in the implementation of universal screening for behavioral and emotional risk to support multi-tiered systems of support: Two case studies. *Journal of Applied School Psychology, 37*(1), 69–107.

von der Embse, N., Kim, E. S., Jenkins, A., Sanchez, A., Kilgus, S. P., & Eklund, K. (2021). Profiles of rater dis/agreement within universal screening in predicting distal outcomes. *Journal of Psychopathology and Behavioral Assessment, 43*, 632–645.

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.

von der Embse, N. P., Pendergast, L. L., Kilgus, S. P., & Eklund, K. R. (2016). Evaluating the applied use of a mental health screener: Structural validity of the social, academic, and emotional behavior risk screener. *Psychological Assessment, 28*, 1265–1275.

Wehmeyer, M. L., & Schwartz, M. (2001). Disproportionate representation of males in special education services: Biology, behavior, or bias? *Education & Treatment of Children, 24*, 28–45.

Wood, B. J., & Ellis, F. (2022). Universal mental health screening practices in Midwestern schools: A window of opportunity for school psychologist leadership and role expansion? *Contemporary School Psychology*. https://doi.org/10.1007/s40688-022-00430-8

Young, E. L., Sabbah, H. Y., Young, B. J., Reiser, M. L., & Richardson, M. J. (2010). Sex differences and similarities in a screening process for emotional and behavioral risks in secondary schools. *Journal of Emotional and Behavioral Disorders, 18*(4), 225–235.

Zakszeski, B., Ormiston, H. E., Nygaard, M. A., & Carlock, K. (2023). *Informant discrepancies in universal screening as a function of student and teacher characteristics*. Unpublished manuscript. Education & Human Development, University of Delaware.