IN PRESS AT ***School Psychology Review***

Investigating the Longitudinal Association between Fidelity to a Large-Scale Comprehensive School Mental Health Prevention and Intervention Model and Student Outcomes

Abstract

Many youth experience mental health problems. Schools are an ideal setting to identify, prevent, and intervene in these problems. The purpose of this study was to investigate patterns of student social, emotional, and behavioral risk over time among a community sample of 3rd through 12th grade students and the association of these risk patterns with fidelity to a school-based mental health model. Overall growth of social, emotional, and behavioral problems declined over a three year period. Four classes of students were identified using Growth Mixture Modeling (GMM): 1) students with high levels of problems, 2) students with decreasing problems, 3) students with increasing problems, and 4) students with stable, low levels of problems. These growth trajectories were associated fidelity to the model, in that trajectories where students with higher or increasing problems were more likely to be from schools with lower fidelity. Implications for practice and policy are provided.

Key words: prevention, school mental health, fidelity, evidence-based intervention

Impact Statement: Mental health issues interfere with the ability of children and youth to learn in the school setting. Schools can implement comprehensive mental health models that include universal screening, prevention, and evidence-based intervention. Doing so with high fidelity can lead to prevention of newly developing risk and in decreasing risk for students over time.

Investigating the Longitudinal Association between Fidelity to a Large-Scale Comprehensive School Mental Health Prevention and Intervention Model and Student Outcomes

Children of all ages and abilities attend school every day, each with different social, emotional, and behavioral strengths and needs. As school psychologists, our top priority is ensuring that each of these students is given the best opportunity to have successful educational experiences. Many interrelated variables can impact these experiences, however, such as a child's externalizing, internalizing, social relationships, and school participation behaviors (Côté, Vaillancourt, LeBlanc, Nagin, & Tremblay, 2006). Atypical manifestations of these variables can have detrimental impacts on students' relationships with their peers, school connectedness and overall academic achievement (Dishion, Véronneau, & Myers, 2010). More concerning, however, is how these known variables can signify deeper mental health problems in students, and when unresolved, cause deleterious effects on their long-term health (Suldo, Gormley, DuPaul, & Anderson-Butcher, 2014). Though this information is not news to school psychologists, the mental health of students in U.S. schools appears to be declining. Two separate national surveys administered biannually have documented a steady increase in youth mental health concerns over the past decade (Center for Disease Control and Prevention [CDC], 2019; Twenge, Cooper, Joiner, Duffy, & Binau, 2019). For example, in a nationally representative sample of over 200,000 adolescents, Twenge and colleagues reported a 52% increase in youth who have met diagnostic criteria for at least one major depressive episode in the prior year from 2005-2017. Likewise, the CDC reported that 31.5% of youth reported persistent feelings of sadness on the Youth Risk Behavior Survey, a 17% increase from 2009-2017 (CDC, 2019). Even more alarming, these same data sources revealed a significant increase in youth seriously considering a suicide attempt (17.4%; CDC, 2019) and a 56% increase in deaths by suicide in the past 15 years among young adults (Twenge et al., 2019). Coupled with increasing prevalence rates of anxiety in our youth (CDC, 2019), this information alone suggests that the mental health needs of youth remain largely unmet at the national level, indicating a dire need for greater effort in this area. This is important because even our best efforts to support academic growth are crippled when underlying behavioral or emotional issues exist (Cooper, Brown & Yu, 2020).

With students spending the greatest portion of their day in the school environment, schools provide the perfect setting for both conducting evidence-based assessments to identify the degree of student need and delivering targeted social, emotional and behavioral supports to students (Herman, Reinke, Thompson, 2019; Kilgus, Reinke, & Jimerson, 2015). School-based delivery for mental health supports has shown great success, and one could argue that this is largely due to the reduction in barriers associated with accessing these services within the community (e.g., health insurance, transportation; Pullmann et al., 2013). Additionally, school-based services have the added benefit of being able to support students who have a wide variety of needs, through a continuum of supports that may already exist within the school setting (Bruns et al., 2016).

Although school-based delivery is arguably one of the most effective methods for providing mental health supports to youth, it does not come without its challenges. Schools are responsible for serving a large group of youth and unfortunately, even the most well-intentioned efforts are ineffective in the absence of a formalized structure for doing so (Reinke, Herman, & Tucker, 2006; Reinke, Stormont, Herman, Puri, & Goel, 2011). In many schools across the country, proper screening and assessment methods to identify the more than 15% of students who present with additional social, emotional and behavioral needs (CDC, 2019), is practically nonexistent, leading to the under-identification of these youth (McIntosh, Ty, & Miller, 2014). Less than 14% of schools report conducting universal screening of student social and emotional needs (Bruhn, Woods-Groves, & Huddle, 2014).

What's more, the "Refer-Test-Place" model and teacher referral methods (i.e., Office Discipline Referrals) that public education continues to use for student problem identification, despite recommendations against it, are entirely reactive and occur after precursors to these issues could have already been identified, and have repeatedly been associated with student failure (Cash & Nealis, 2004; Bruns, et al, 2006). These methods of identification create unnecessary lags in services for students—with services often being offered after the problem is already in place (Shinn & Walker, 2010; Bruns et al, 2006). Most importantly, continuing to utilize these methods of identification for social, emotional and behavioral issues frequently results in school professionals missing the window of opportunity to properly intervene before an issue becomes ingrained or gets progressively worse (Oakes, Lane, Cox, & Messenger, 2014).

To address this problem, many schools utilize Multi-Tiered Systems of Support (MTSS; McIntosh & Goodman, 2016) to more effectively provide the appropriate level of services to each student. Through the use of MTSS, schools can more efficiently utilize school resources and provide targeted services to students with both more and less-intensive needs. Additionally, MTSS provides a framework for taking a more proactive approach to remedying potential mental health concerns in youth by employing more preventive methods at the universal level, to address issues before they occur. This approach is very different from the typical reactive approach that is typically taken in the area of social, emotional and behavioral student concerns.

Often, school professionals believe that identifying students with the most need is of the utmost priority when considering service delivery in the school setting. However, we would argue that the group of students who might be classified as “at risk” of developing more serious social, emotional and behavioral concerns are equally as important to identify and treat, as they often go undetected using traditional reporting measures (Burns et al., 2004). Equally prioritizing early intervention can not only prevent present risk factors from worsening, thus being a more effective solution, but can be done using much less-intensive resources than if the issues were more severe (Dowdy, 2015). Identifying students who require these early intervention services is most effectively achieved through the use of universal screening methods, classified under universal supports within the MTSS framework (Cowan et al., 2013). In addition to identifying students with a greater need of support, universal screening provides data that can identify strengths and weaknesses within a school setting, including potential school or grade-wide needs regarding social, emotional and behavioral skills (see Reinke, et al., 2018).

The systematic screening of all students in a school provides educational professionals with the information needed to more efficiently determine the social, emotional, and behavioral status of all students (Dowdy et al., 2010; McIntosh, Reinke, & Herman, 2010), and adequately provide either a more or less-intensive group of services for each student (Gresham, 2005; Bruns et al, 2016). We know that the availability of reliable and effective assessment and interventions to promote students' social, emotional and behavioral well-being is not the issue. Rather, a school's success in this area heavily relies on its dedication to the implementation of prevention and early intervention efforts. To positively impact student academic outcomes, schools must first satisfy social, emotional and behavioral needs, which can be effectively achieved through the use of universal screening, and interventions that are closely aligned with the needs of the school and each individual student (Green et al., 2013; Bruns et al., 2016). The following section describes a current comprehensive tiered model of support that uses universal screening linked to evidence-based interventions in efforts to meet the prevention and intervention needs of students in schools.

Description of the Coalition Model

Since 2016, a Coalition of six school districts and one parochial school, in collaboration with researchers, have been utilizing universal screening within a comprehensive model devised to identify youth in need of supports before the issues become ingrained, to prevent social, emotional, and behavioral problems among youth, and to intervene using evidence-based interventions directly tied to the problem areas identified through valid and important data (see AUTHORS, 2018; AUTHORS, 2017). The model utilizes the Early Identification System (EIS; AUTHORS, 2019), which is a universal social, emotional, and behavioral screener that is administered three times per year. The EIS has both a teacher report and student report version. Teachers complete the measure for all students in their classrooms across Kindergarten to 12th grade. Students in 3rd through 12th grades complete the measure themselves. The EIS has a total of seven areas that are connected directly to evidence-based universal, selective, and indicated interventions, including internalizing behavior, externalizing behavior, emotion dysregulation, peer relationship problems, attention and academic problems, school disengagement, and relational aggression (AUTHORS, under review). Each of these areas are linked to evidence-based interventions across grade levels and within the tiered model of supports. All data are gathered electronically and once surveys are complete, the data are populated in reports for schools to review in a red, yellow, green format (AUTHORS, 2018). As such, for universal prevention planning, the report will indicate in red (needs attention) areas where 20% or more of students are reported to have risk, indicating that intervening one student at a time may not be the most efficient or effective practice. Instead, school teams are linked to evidence-based universal prevention interventions that can be implemented with a large number of students. School level, grade level, and individual student level reports are provided. Data from the EIS are locally normed (e.g., scores are relative to students in the same school) and student reports are generated that indicate an area of need (red) for students who are two standard deviations or higher than their peers within each target area of the screener. Yellow indicates students are one standard deviation higher than their peers, and green indicates that they are within the normative range compared to their peers.

Once universal screening is complete, school teams review the data-based reports, identify areas of concern, determine appropriate interventions that are linked directly to areas of concern, and devise a plan for training staff and implementing selected interventions across all tiers. Data, including pre-post assessments and progress monitoring data are gathered and used to evaluate the effectiveness of the interventions. A measure to assess the fidelity to the model that includes gathering and using data, selecting appropriate interventions, implementing interventions, and evaluating the effectiveness of interventions implemented across all tiers was recently developed and utilized by the Coalition. The purpose of the fidelity measure is to help guide areas in each school that can be improved so that schools are implementing the model with high levels of fidelity, which is expected to lead to improved student outcomes over time.

Fidelity to Multi-Tiered Systems of Support

It is widely agreed upon that interventions implemented as intended, thereby with fidelity, are associated with higher outcomes and increased chances of mirroring efficacy trial results (Power et al., 2005; Durlak & DuPre, 2008). Whereas tiered models of prevention and intervention are well-founded empirically, as noted earlier, the implementation of large-scale evidence-based programming and prevention frameworks, like the Coalition model, is complex and presents unique challenges to school. To be successful, schools need to adopt a systematic process for screening, data-based decision-making, and the selection and implementation of evidence-based interventions at the universal, selective, and indicated levels (Herman, Reinke, Thompson, & Hawley, 2019). Indeed, schools demonstrate difficulty adopting and translating such processes and practices with the same level of fidelity and rigor that is observed in efficacy trials (e.g., Dusenbery et al., 2005, Gottfredson & Gottfredson, 2002). Several barriers impede the implementation of these models, including administrator support, organizational structure, and school personnel certification and training (Bradshaw, Koth, Thorton, & Leaf, 2009; Bradshaw & Pas, 2011; Domitrovich et al., 2008). As a result, implementation of such large-scale models often takes approximately three to five years to achieve high implementation quality and positive outcomes (Rimm-Kauffman et al, 2007; Sugai & Horner, 2006).

Although it is common to examine the fidelity of universal, selective, and indicated interventions, research examining the association between outcomes and fidelity to large-scale tiered model processes is limited. In one notable examination, Pas and Bradshaw (2012) examined the scale up of School-Wide Positive Behavior Intervention Supports (SW-PBIS) using three fidelity measures:

Implementation Phases Inventory (IPI; Bradshaw et al., 2009), the School-wide Evaluation Tool (SET; Sugai et al., 2001), and the Benchmarks of Quality (BoQ; Kincaid et al., 2010). The fidelity measures were designed to assess key features of SW-PBIS and key implementation variables necessary to deploy the model and achieve student outcomes. Results indicated that higher levels of IPI fidelity scores were associated with student math and reading achievement and truancy but not suspensions. The SET and BoQ were unrelated to student outcomes. The authors suggested that the IPI may have been more sensitive to change because it uniquely measured core features of PBIS as well as advanced implementation phases and skills compared to the SET which focuses on critical features of early implementation quality. In a similar manner, a statewide evaluation of a comprehensive school guidance model found that school counselor ratings of their current adherence to the three core components of this model predicted student reports of their grades and school climate (Lapan et al., 1997). These results highlight the importance of measuring the fidelity of core components as well as advanced implementation skills of tiered models. Measures that assess the fidelity to tiered model processes can help to identify schools that are considered low and high implementation quality, allow for school professionals to better predict outcomes, and help to determine gaps in implementation.

Purpose of the Study

The Coalition model is currently implemented across 54 schools with varying degrees of fidelity to the model. The purpose of this study was to investigate the association between the degree of fidelity to the model and patterns of student social, emotional, and behavioral risk over time. Using data from student report on the universal screening measure, the EIS-SR, a person-centered approach was used to identify patterns of student risk over time. We were interested in determining the growth trajectories for the total student reported social, emotional, and behavior problems on the universal screening measure for this community sample of 3rd through 12th grade students. First, we examined the overall growth of social, emotional, and behavioral problems over time for the full sample. Because all schools were partially to fully implementing the comprehensive mental health model with universal screening and supports, we expected a small and steady decline in youth-reported total problems over the study period. Next, using Growth Mixture Modeling (GMM), we expected different subgroups of students to emerge with unique patterns of growth over time. Specifically, we hypothesized that (a) a small portion of students would exhibit high levels of social, emotional, and behavioral problems across time, (b) a small portion of students would demonstrate high levels of social, emotional, and behavioral problems with reductions over time, (c) a small number would demonstrate an increase in social, emotional, and behavioral problems over time, and (d) the majority of individuals would have stable, low levels of social, emotional, and behavioral problems over time. We further hypothesized that these growth trajectories would be associated with the level of fidelity to the comprehensive mental health model, in that trajectories where students with higher or increasing total social, emotional, and behavioral problems would be from schools with lower fidelity to the model.

Method

Participants

The participating students (N = 16,782) were from 54 school buildings situated in 6 school districts plus one parochial school participating in a county-wide school mental health program that included universal screening, conducted three times per year. Student participants were, 51.2% male, 70.2% White, 14.9% Black, 5.8% Multiracial, 4.6%Latinx, 4.0% Asian, 0.3% American Indian, and 0.2% were Pacific Islander. Participants were from elementary (52/5%), middle (30.3%) and high school (17.2%) settings. Thirty-six percent of the sample received Free or Reduced Meals (FRM).

Measures

Early Identification System- Student Report (EIS-SR). The EIS-SR is a universal social, emotional, and behavioral screening instrument that has seven subscales, including attention and academic problems, peer relationship problems, internalizing problems, externalizing problems, emotion dysregulation, school disengagement, relational aggression and a total problems score across all items. The seven subscales were initially identified from a review of the developmental cascades model (Patterson, Reid & Dishion, 1992). Next, the EIS development team identified salient risk factors reflective of each domain in coordination with school-based student support professionals. Subsequent exploratory factor studies revealed the individual items that were developed and reviewed by school professionals and researchers did indeed coalesce into the hypothesized seven subscale structure and a total scale score with excellent fit to the data and reliable scale alphas. For more information on the selection and development of the items and corresponding domains assessed by the EIS-SR please see previously published research reports by AUTHORS (2019), AUTHORS (2018) and AUTHORS (2017) and colleagues. Students in grades 3rd through 12th grade have completed the measure three times per year across three school years (October, January, and April, 2016-2019). Students answered a total of 37 questions, such as, "I feel left out by others" and "I am a good friend." The survey took students between 5 to 15 minutes to complete. Response options were Likert-type scales (0 = *Never*, 1 = *Sometimes*, 2 = *Often*, 3 = *Always*). For the purpose of this study the total problems score was used in the growth analyses. A total of 9 time points were included across three years. Coefficient alpha reliabilities for the total problems score ranged from .91 to .93.

Fidelity to the Coalition Model. A measure to evaluate the fidelity to the comprehensive school mental health model was developed and completed by the mental health consultants and school staff from each building. Items on this measure were directly related to specific activities that each school should be doing to be implementing the model with fidelity. The measure evaluates school level fidelity across three key areas: 1) data collection and review of universal screening data, 2) intervention planning and implementation across universal, selective, and indicated level, and 3) progress monitoring and evaluation of the effectiveness of interventions. The measure has a total of 34 items. Items include questions such as, "Did the school use school level data to determine if universal school level or grade level interventions were needed?", "Did the identified Tier 2 supports match the needs identified by the data?", "Did the school gather progress monitoring data for Tier 3 supports?", "Did the school use pre-post data to determine if the intervention was effective?" The mental health consultant and school staff completed the measure together answering whether each item was completed (3=*yes,* 2=*somewhat,* and *1=not at all*) in January of the school year after time for gathering data, implementing interventions, and evaluating interventions were possible. When schools indicated that an item was somewhat completed, they were asked to describe what was meant. The purpose of the measure was to highlight areas for improvement with regard to fully implementing the model. The fidelity tool was administered in the middle of the third year of implementation; given that schools typically take 3-5 years to reach full implementation of school-wide programs, this three year window allowed for sufficient time for schools to reach full implementation of the model as well as variability in implementation between schools. For the purpose of this study, the total score on the fidelity measure was calculated (α = .84). We followed conventional criteria for defining thresholds of effective implementation needed to reach student outcomes; i.e., 80% or higher (Horner et al., 2004). Thus, schools that had 80% of the total possible score were categorized into implementing well = 1. Schools with scores below 80% were categorized as having lower implementation = 0.

Demographic Information. Schools provided demographic information for students, including sex, Free or Reduced Meal (FRM) status, and race.

Analysis

Latent growth curve modeling (LGC) and growth mixture modeling (GMM) using the Mplus version 8 statistical software package (Muthén & Muthén, 2017) were utilized to examine the growth and to identify patterns of growth for total student reported social, emotional, and behavioral problems over a three year period. The analysis occurred in several stages. To test the first hypothesis that there would be decline in total social, emotional, and behavioral problems over time LGC was conducted. First, unconditional LGC models (without covariates) were estimated to determine the shape of the trajectories that would provide guidelines for subsequent analyses. All models were nested within schools to account for inflated interclass or within school variance. The overall fit indices for the LGC models included the comparative fit index (CFI), the Tucker-Lewis index (TLI), and root mean square error of approximation (RMSEA) provided by Mplus. Models are regarded as acceptable if the CFI and TLI are greater than 0.9. A model with an RMSEA of less than 0.05 is regarded as a “good” fit, and an RMSEA of less than 0.08 is “acceptable” (McDonald & Ho, 2002). The next step in the analysis was to fit the conditional LGC model by including the covariate measures at baseline. The conditional models were estimated with sex, race, FRM status, and school level (i.e., elementary, middle, or high school) as covariates. Change in growth over time was evaluated in relationship to whether the overall sample demonstrated a decline or increase or neither over time.

Next, the GMM analyses were conducted and based on the unconditional LGC models (i.e, included growth indices). To determine the relative fit of the models for varying numbers of classes, we used the most accepted and widely cited methods (Bauer & Curran, 2003; Muthén, 2003). First, we compared models with differing numbers of classes using the Akaike information criterion (AIC; Akaike, 1987), the Bayesian information criterion (BIC; Schwartz, 1978), and the sample-size adjusted Bayesian information criterion (aBIC; Sclove, 1987). Typically, the smaller the information criteria, the better the model fit to the data. In addition, we evaluated the classification precision as indicated by estimated posterior class probabilities, summarized by the entropy measure (Ramaswamy, DeSarbo, Reibstein, & Robinson, 1993). Entropy values close to 1.0 indicate higher classification precision. In prior research, entropy values higher than 0.80 have been interpreted to mean good classification (Muthén, 2004). Finally, models with varying numbers of classes were evaluated and compared according to substantive utility, distinctiveness, and interpretability of the resultant class sets. Once the appropriate number of trajectories was determined, these classes were used to determine the association between classes, covariates, and fidelity to the model by means of latent class regression analysis (Guo, Wall, & Amemiya, 2006) and examination of the odds ratio estimates. Further, descriptive data are provided that include the percentage of students in each trajectory.

Missing Data

The Mplus software utilizes full information maximum likelihood estimation under the assumption that the data are missing at random (Arbuckle, 1996), which is a widely accepted way of handling missing data (Muthén & Shedden, 1999; Schafer & Graham, 2002). Overall, 79% of the participants had at least six of the nine assessment time points. The minimum covariance coverage recommended for reliable model convergence is .10 (Muthén & Muthén, 2009). In this study, coverage ranged from .54 to 1.00, well within the recommended range.

Results

Unconditional LGC models were first fit to determine the shape of the trajectories and variances in the growth factors. Using the Sattora-Bentler Scaled Chi-Square Difference test, used for nested data, determined that including a slope parameter significantly improved the fit over that of the intercept model ($\Delta\chi 2_{SB} =$ 760.08(3) $p < 0.001$; CFI = 0.98, TLI = 0.98, RMSEA = 0.05). Although the addition of a quadratic parameter further improved the model fit ($\Delta\chi 2_{SB} =$ 728.05(4) $p < 0.01$; CFI = 0.99, TLI = 0.99, RMSEA = 0.03), the mean of the quadratic term was not significant and visual inspection of the trajectory aligned with a linear slope. Thus, it was determined that the linear model was the most appropriate. The variances in the intercept and slope growth factors were significantly different from zero, suggesting individual differences in pathways of total social, emotional, and behavioral problems. Further, the mean linear slope was negative and significant ($x = -0.47$, $p = .002$), indicating that there was a small, but statistically significant decline in social, emotional, and behavioral problems for students over time.

Based on the fit of the unconditional model, conditional LGC models were estimated by incorporating the baseline covariates into the model with paths from each covariate (sex, race, FRM status, and school level) leading to the growth factors for total social, emotional and behavioral problems. Several variables were significant predictors of the intercept, including sex (B=0.05, p< .05), indicating that males had higher average social, emotional, and behavioral problems in the fall of 2016-2017 than females. Race was also a significant predictor of the intercept, indicating that Black (B=0.08, p< .001), Multiracial (B=0.04, p< .001), and American Indian (B=0.02, p< .05) students reported higher levels of social, emotional, and behavioral problems in the fall of 2016-2017 compared to White students. Asian (B= -.02, p< .05) and Latinx students (B= -0.03, p<.05) reported lower levels of social, emotional, and behavioral problems in the fall of 2016-2017 compared to White students. Students receiving FRM reported higher levels of social, emotional, and behavioral problems in fall of 2016-2017 (B=0.19, p< .001). Lastly, high school students reported higher levels of social, emotional, and behavioral problems in the fall of 2016-2017 (B=0.14, p< .001).

With regard to growth, several variables predicted the linear slope factor. For instance, Asian students had less decline in social, emotional, and behavioral problems over time as compared to White students (B= -0.05, p< .001). Also, female students had less decline in social, emotional, and behavioral problems over time as compared to male students (B= -0.05, p< .001). Further, students in high school had less decline in social, emotional, and behavioral problems than elementary students (B= -0.08, p< .05). Whereas students who received FRM had greater decline over time (B= 0.03, p< .05).

The GMM was an extension of the LGC models, formed by adding a latent categorical variable. As described in the analysis subsection, to determine the best-fitting GMM model, we considered the AIC, BIC, and aBIC indices, with the smaller value indicating a better fit model. In addition, entropy was considered in the determination. Entropy values close to 1.0 indicate higher classification precision or differentiation between classes of students. We then included class prevalence and interpretability (the extent to which an additional class provided unique information) as additional criteria while selecting the best-fitting models. According to the AIC, BIC, and aBIC, significant improvements in model fit were observed in up to six classes as is common for large samples (see Table 1). However, inspection of the five-class and six-class solutions indicated that each produced a new class that mimicked those in the four-class solution, but the new classes represented very few students (less than 1% for each new class), providing little additional information and not necessarily in line with theory. Thus, the four-class solution was deemed the best fitting model for the sample. The four trajectories included a group of students exhibiting high stable levels of social, emotional and behavioral problems (10.64%%), a group of students exhibiting social, emotional and behavioral problems that increased over time (5.69%), a group with social, emotional, and behavioral problems that decreased over time (5.76%), and a group exhibiting consistently low levels of social, emotional and behavioral problems (77.92%). After obtaining the predicted group membership from the GMM analyses, the estimated means were calculated (see Figure 1).

Following identification of the appropriate number of classes, the classes were used to determine the association of each class with school level of fidelity to the comprehensive school-based mental health model. The findings are reported in terms of odds ratios; that is, the odds that youth in the social, emotional, and behavioral problems subclasses were more likely to be from a high fidelity or lower fidelity school.

In relationship to fidelity, one finding was significant and another approached significance. Youth in the increasing class social, emotional, and behavioral problem class were significantly more likely to attend a school with lower fidelity. Further, the association between the high stable class and fidelity approached significance (p=.06) in that students in this class were more likely to attend a low fidelity school. More specifically, students in the high stable class were 1.29 times more likely to be in a lower fidelity school than students in the low stable class, whereas students in the increasing class were 1.28 times more likely to be in a lower fidelity school than students in the low stable class. Students in the decreasing social, emotional, and behavioral problems class, although not statistically significant, were 2.12 times more likely to be in a high fidelity school when compared to those in the low stable class.

There were also several significant differences across classes by student demographics. For instance, students in the high stable social, emotional, and behavioral problems class were 2.01 times more likely to be Black, 1.71 times more likely to be Multi-Racial, 3.54 times more likely to receive FRM, and 3.02 times more likely to be in high school than students in the low stable class. Also, Asian students were 2.34 times less likely to be in the high stable social, emotional, and behavioral problems class than in the low stable class. Whereas, students receiving FRM were 2.75 times more likely to be in the increasing social, emotional, and behavioral problems class when compared to the low stable class, and Asian students and female students were less likely to be in this class (OR: 4.30 and 1.48 respectively). Lastly students in the decreasing social, emotional, and behavioral problems class were more likely to be Black (OR: 1.77), male (OR: 1.29), receiving FRM (2.85) and less likely be in middle school (OR: 2.69) than students in the low stable class.

Discussion

The purpose of this study was to investigate the association between fidelity and outcomes in a comprehensive school mental health model that conducts universal social, emotional and behavioral screening with students in grades 3 to 12 three times each year for the explicit purpose of reducing social, emotional and behavioral health risks. As such, we were interested in whether there were any changes in total social, emotional, and behavioral problems as reported by students in this sample over a three year implementation period. Hypotheses were supported in that there was a significant, though small, decline in student reported social, emotional, and behavioral problems over time. Additionally, lower levels of implementation of the model predicted worsening or persisting student social emotion problem trajectories whereas higher implementation was associated with improving and lower risk trajectories.

The overall decline in student-reported social, emotional and behavior problems over time contrasts sharply with national data regarding the growing prevalence of these problems among youth in the U.S. Several national surveys have found consistent evidence of increasing youth mental health problems during the past decade (CDC, 2018; Tweege et al., 2019). The declining risk in these schools also supports the need for schools to engage in systematic universal screening so they may act upon early indicators of oncoming social, emotional, and behavioral health risk factors rather than waiting until those issues result in office referrals, absenteeism, suspensions, and poor academic performance. Although we cannot draw causal inferences from our study design or from these comparative data, the findings are promising in that they suggest that the implementation of the Coalition model—a model of prevention—is one potential explanation for the declining rates of mental health concerns.

This explanation is bolstered by the subsequent analyses in this study where we examined the relations between fidelity of implementation and growth patterns of student risk over time. We used GMM, a person-centered approach, to identify patterns of student risk over time and found four types of trajectories—which largely confirmed our hypotheses derived from prior studies. That is, the majority of students (78%) were characterized by a normative pattern of low stable levels of social, emotional, and behavioral problems across the three years. The second group of students (6%) fell into a pattern of decreasing total problems over time. This finding is promising given that it represents a sizeable subgroup of youth who improved over time, and thus aligns with the goal of the Coalition to reduce the prevalence of youth mental health concerns. The final two subgroups of students fell into high stable (11%) and increasing (6%) social, emotional, and behavioral problems classes patterns. These patterns are closely aligned with public health models which indicate that in optimal preventative approaches where universal supports are adequate between 15 to 20% of the student population would benefit from more selective or indicated interventions (Herman et al., 2019). Most importantly, students with stable high and increasing social, emotional, and behavioral problems were more likely to be in schools with lower fidelity to the model at year 3 of implementation. Thus, implementation of the Coalition model was associated with lower student risk for maladaptive trajectories of social, emotional and behavior risk over time. In turn, high implementation of the model predicted student membership in the more adaptive trajectories: stable low and decreasing patterns of risk.

Sociodemographic characteristics also predicted student trajectories. Black and multiracial students and those who qualified for FRM were more likely to have elevated social emotional risks at baseline and stable high problems over time. This is important because consistent evidence suggests that youth of color are significantly less likely to seek or be referred for mental services (Marrast, Himmelstein, & Woolhandler, 2016). Thus, the universal screening and supports offered by a model like the Coalition holds potential for identifying youth in need and connecting them with services. Older students were also more likely to have higher social, emotional, and behavioral problems which is consistent with developmental literature showing the onset of youth mental health disorders peaks in middle school and remain high through early adulthood (Merikangas et al., 2010). One important note is that Black youth and those who qualified for FRM were more likely fall into a class characterized by decreasing symptoms compared to those in the stable low class. This is likely an artifact of the higher base rate of total problems for these youth at baseline but points to the promise of comprehensive school-wide initiatives such as the Coalition to lower these rates given that students in high implementation schools were also more likely to be in the decreasing class.

Implications

The study represents an advanced translational study to bring effective practices to scale across multiple school districts to impact the population health of students. This may be the first study to document an association between effective implementation of a comprehensive county-wide school mental health model and improvements in youth-reported mental health symptoms. A prior evaluation of school-wide PBIS implemented at scale found that a single implementation measure predicted some student academic outcomes based on school records (Pas & Bradshaw, 2012). However, the study only had outcome data from one time point and did not include student social emotional outcome measures. Here we collected student self-reported social, emotional, and behavioral symptoms at nine time-points across three years. A separate study found that statewide implementation of a comprehensive school guidance system predicted student reports of their grades and perceptions of climate (Lapan et al., 1997). However, this study was cross-sectional and thus was unable to rule out other potential factors, including baseline functioning of the school.

The present study has potential implications for school practitioners and researchers attempting to implement and disseminate effective practices in schools on a large scale. It is noteworthy that schools were able to implement the Coalition school mental health model over a several year period, with many implementing with high fidelity. This contrasts with the low rates of multi-tiered social emotional screening and intervention implementation typically observed in schools. Schools often struggle to implement even singular elements of these models such as universal screening (Bruhn et al., 2014) let alone connecting screening scores to evidence-based practices, procuring training and/or materials to implement a solution, and then adhering to the solution as it was designed. In this project, all 54 schools implemented universal screening tri-annually over a three year period and most schools used these data in appropriate ways to guide decisions about multi-tiered supports. Regardless, the results of these analyses strongly suggest that schools who more fully adhered to the model experienced better student outcomes compared to schools with lower rates of adherence.

What are the critical elements that distinguish this project from other efforts to widely disseminate comprehensive school mental health approaches that may inform future efforts in this regard? First, it would be disingenuous to not acknowledge the importance of ongoing funding to support this initiative. It would have been difficult to develop the online and reporting features of the data tool, the EIS, without this funding. Additionally, the high quality of implementation achieved by most schools in the project would not have been possible without technical assistance offered by consultants funded by the project. At the same time, it would be simplistic to think that funding alone achieved these outcomes. Abundant research indicates that simply increasing youth access to mental health supports does not improve their outcomes (Bickman et al., 2000; Weisz et al., 2006). Moreover, the tax initiative that supported this project has been passed by nearly a dozen other communities; it is unclear whether these initiatives have in any way impacted the delivery of services in schools or youth mental health. Rather it is the synergy between the strategic investment of funding to support youth mental health in combination with the systemic development of a model rooted in science, both evidence-based practices and implementation science. In addition, the success and/or failure of these large scale, system wide efforts to alter context and culture in schools with respect to using data to drive decisions requires more than just funding, consultation, and researcher involvement. Real change in schools and communities requires the ongoing input of all involved—from superintendents, building administrators, student support personnel, teachers, and students. Any attempt to make such large scale and broad changes without input from all of these groups will likely experience push back and potential rejection resulting in possible failure. Thus, findings from the present study may be utilized to advocate for the strategic investment in youth mental health initiatives and public policies that are guided by science and ongoing evaluation.

A second critical element was the development and use of an efficient and functional data system (the EIS) to inform school practice. Although many universal screening tools exist, the EIS has the following advantages over other similar tools: (1) it is efficient (elementary teachers take about 10 minutes to complete an entire class and students can complete in less than 15 minutes); (2) it is locally normed so that schools identify the students most in need based on their local context; (3) it is parallel across development such that the same items are administered from elementary through high school, and (4) the EIS domains are directly mapped to a menu of evidence-based practices to assist school personnel to select effective options to address areas of concern. Use of the EIS and the evidence-based practices mapped to the EIS domains hold the potential to increase the widespread adoption of evidence-based universal screening and relying on screening scores to drive subsequent implementation of scientifically supported practices at the universal and targeted levels.

Third, ongoing technical assistance from mental health consultants supported school personnel to collect and interpret the data and select and plan subsequent supports. That the EIS domain and total scores are locally normed or scored relative to the risk within each school building is important compared to the manner in which some commercially available tools are scored. Most commercially available tools rely on isolated samples of youth—albeit large samples—to derive cutoff scores that identify a youth as being at risk. All that said, all data is sample specific—and so are cutoff scores. The benefit to schools who rely on data normed on their own building is that those scores accurately identify youth in that building, situated in that community, who are at risk relative to their own peers rather than to similar aged youth in a completely different part of the country exposed to completely different risks and stressors. Relying on such locally normed data permitted the consultants to also tailor interventions for each building based on concerns in their own data as well as existing strengths and limited and highly variable resources. For instance, schools without behavior support teams are first helped to create these teams and plan effective meetings. Schools with more behavior capacity are taught how to use the data to plan intervention supports at the universal, selective, and indicated levels.

Fourth, the collection of ongoing fidelity and outcome data helped ensure the program was meeting school and county objectives. The fidelity data in particular helps guide building level improvements and informs district level leadership of what areas of improvement need to be planned for from year to year. Additionally, school level leadership and student support personnel are supported with the use of efficient tools to measure meaningful outcomes in relation to more intensive interventions delivered at each tier of support (e.g., measures of response to intervention of a social skills curriculum). Such technical support requires the input and guidance of researchers who can not only identify and map such relevant measures onto a system of available interventions but can also set up feasible ways to use those measures to appraise intervention impact. Such feedback is extremely formative and motivating when a school psychologist or educator who works towards change has their efforts confirmed using data reflecting improvements in student functioning—which may have the effect of these school personnel doing more of what works and relying on data to make those decisions rather than guessing or doing the same thing over and over.

Finally, district leadership commitment and support was essential to the success of the model. The Coalition is a partnership between all county superintendents and university personnel. We meet as a Board on a monthly basis to discuss progress and challenges and find solutions. In addition, as researchers, we seek to include the coalition as partners in all that we do to scientifically examine the model either through including them as co-authors on publications and research reports or counting them as co-investigators and budgeting for their needs in grants that we submit to extend aspects of this work. Without a doubt the Coalition would not exist or persist absent the local funding and collaboration—and equally important is the need to have honest, unvarnished and genuine exchanges between district leadership and university partners. To be sure, administrative turnover in school leadership is challenging to maintain such momentum—but within a coalition, even losing a superintendent results in a new leader having a cohort of peers that continue the activities of the model.

Regarding the fidelity tool, development of tools to evaluate large scale projects may be best informed by the hybrid model used in this study. Although direct observations of fidelity practices are often viewed as the gold standard they are often impractical in standard practices and may be insensitive to critical features of effective implementation (Pas & Bradshaw, 2012). Here we employed a hybrid model that incorporated the ratings of an external evaluator and of school personnel in each building. Additionally, the tool focused on core components of full implementation based on a detailed theory of change as well as advanced implementation skills.

Limitations

The study describes the county-wide implementation and evaluation of a comprehensive school mental health model brought to scale. All schools in a medium-sized county were part of the Coalition and had access to intervention supports, thus, random assignment to condition was not an option. Without experimental manipulation, causal inferences are not warranted. Of course, large scale translational research in schools is complicated and other designs are often needed to build a consensus of information to determine promising effects. Here we employed sophisticated growth and growth mixture modeling analyses to describe patterns of student growth over time. This rich data stream allowed us to examine the relation between fidelity to the overall model and student risk trajectories. The finding suggested fidelity was associated with student symptom trajectories in ways that support the impact of that model. The finding is particularly promising given the lower power associated with the "small" sample of schools (n-54) that served as the degrees of freedom for these fidelity analyses.

As noted, the fidelity tool was administered during the third full year of implementation, thus, early implementation data was not available. Although less than desirable, it is important to note that full implementation of school-wide initiative takes time with full implementation occurring for most schools 3-5 years after initiation (Bradshaw, Reinke, Brown, Bevins & Leaf, 2008; Horner et al., 2012; Molloy, Moore, Trail, Van Epps, & Hopfer, 2013). Thus, it is reasonable to categorize schools as high and low implementers at this third year time-point. It is also noteworthy to mention the difficulty of modeling or capturing all that high fidelity schools and educators do to support students who are at risk. That is, schools that implement the Coalition model with a high degree of fidelity are also likely doing other things to the benefit of students' social, emotional, and behavioral growth. The personnel who make decision to change their culture and context by using data and thinking about the social emotional risk experienced by their students may also be engaged in other training and supports not fully modeled here. Likewise, schools that dismiss the data or do not rely on the data and the model to shape, confirm and evaluate practices may be unwittingly engaging in iatrogenic behaviors or practices that contribute to unintended harm, though more research needs to vet this possibility. It is also likely that this single time-point measurement reflected the culmination of implementation activities during the prior three year period. Given that each school that reached the 80% threshold likely did so at different points in the prior three years (e.g., some during the first and second year, some during year 3) it is important to note that this study artifact likely underestimates the true effect size of full model implementation. That is, students in some high implementation schools may have only been exposed to full implementation for a single year. Future studies will need to examine the effect of sustained model implementation on student outcomes to determine if even larger impacts are observed.

Conclusion

To truly transform our schools and communities into supportive contexts that fully foster positive youth development—we must first and foremost reimagine our schools as places where we socialize our children and not simply as academic institutions where we teach reading, writing, and mathematics. As such, educators must come to understand their role in shaping healthy social, emotional, and behavioral contexts for youth. To address the complexity of student needs we must embrace public health approaches to identify, monitor, and use data to select, drive and evaluate practices targeting those unique risks. Furthermore, once these efforts are in place we must ensure that we stick with the plan if we are to cultivate youth into healthy adults who will contribute to the civic fabric of our communities and pass along these values to future generations.

References

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika, 52*(3), 317–332. doi:10.1007/BF02294359

American Psychiatric Association. 4th ed. Washington, DC: Author; 1994. Diagnostic and statistical manual of mental disorders.

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques*. Mahwah, NJ: Erlbaum.

Bauer, D. J., & Curran, P. J. (2003). Overextraction of latent trajectory classes: Much ado about nothing? Reply to Rindskopf (2003), Muthen (2003), and Cudeck and Henly (2003). *Psychological Methods, 8*(3), 384–393. doi:10.1037/1082-989X.8.3.384

Bickman, L., Lambert, E. W., Andrade, A. R., & Penaloza, R. V. (2000). The Fort Bragg continuum of care for children and adolescents: mental health outcomes over 5 years. Journal of consulting and clinical psychology, 68(4), 710.

Bradshaw, C. P., Reinke, W. M., Brown, L. D., Bevans, K. B., & Leaf, P. J. (2008). Implementation of school-wide positive behavioral interventions and supports (PBIS) in elementary schools: Observations from a randomized trial. *Education and Treatment of Children*, *31*(1), 1-26.

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

Bruns, E. J., Duong, M. T., Lyon, A. R., Pullmann, M. D., Cook, C. R., Cheney, D., & McCauley, E. (2016). Fostering SMART partnerships to develop an effective continuum of behavioral health services and supports in schools. *American Journal of Orthopsychiatry, 86*, 156–170.

Burns, M. K., & Coolong-Chaffin, M. (2006). Response to intervention: The role of and effect on school psychology. *School Psychology Forum, 1*, 3–15.

Burns, B. J., Phillips, S. D., Wagner, H. R., Barth, R. P., Kolko, D. J., Campbell, Y., & Landsverk, J. (2004). Mental health need and access to mental health services by youths involved with child welfare: A national survey. *Journal of the American Academy of Child & Adolescent Psychiatry, 43*, 960–970. doi:10.1097/01.chi.0000127590.95585.65

Calear, A. L., & Christensen, H. (2010). Systematic review of school-based prevention and early intervention programs for depression. *Journal of Adolescence, 33*, 429–438. 10.1016/j.adolescence.2009.07.004

Cash, R. E., & Nealis, L. K. (2004, July). Mental health in the schools: It's a matter of public policy. Paper presented at the National Association of School Psychologists Public Policy Institute, Washington, DC.

Centers for Disease Control and Prevention. (2019). *Youth Risk Behavior Survey: Data Summary and Trends 2007-2017*. https://www.cdc.gov/healthyyouth/data/yrbs/pdf/trendsreport.pdf

Cooper, J., Brown, T., & Yu, M. (2020). A retrospective narrative review of filial therapy as a school-based mental health intervention. International *Journal of Play Therapy, 29*, 86–95. https://doiorg.proxy.mul.missouri.edu/10.1037/pla0000114.supp (Supplemental)

Costello E. J., Mustillo S., Erkanli A., Keeler G., Angold A. (2003). Prevalence and development of psychiatric disorders in childhood and adolescence. *Archives of General Psychiatry, 60*, 837–844. doi:10.1001/archpsyc.60.8.837

Côté, S. M., Vaillancourt, T., LeBlanc, J. C., Nagin, D. S., & Tremblay, R. E. (2006). The development of physical aggression from toddlerhood to pre-adolescence: A Nationwide longitudinal study of canadian children. *Journal of Abnormal Child Psychology, 34*, 71-85. doi: 10.1007/s10802-0059001-z

Cowan, K. C., Vaillancourt, K., Rossen, E., & Pollitt, K. (2013). *A framework for safe and successful schools* [Brief]. Bethesda, MD: National Association of School Psychologists.

deBoo, G.M., & Prins, P.J. (2007). Social incompetence in children with ADHD: Possible moderators and mediators in social-skills training. *Clinical Psychology Review*, *27*,78–97.

Dishion, T. J., Véronneau, M. H., & Myers, M. W. (2010). Cascading peer dynamics underlying the progression from problem behavior to violence in early to late adolescence. *Development and Psychopathology, 22*, 603-619. doi: 10.1017/S0954579410000313

Dowdy, E., Furlong, M., Raines, T.C., Bovery, B., Kauffman, K., Kamphaus, R., Dever, B., Price, M., & Murdock, J. (2015). Enhancing school-based mental health services with a preventative and promotive approach to universal screening for complete mental health. Journal of *Educational and Psychological Consultation, 25*, 178-197.

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*, 405–432. 10.1111/j.1467-8624.2010.01564.x

Green, J. G., McLaughlin, K. A., Alegría, M., Costello, E. J., Gruber, M. J., Hoagwood, K., . . .Kessler, R. C. (2013). School mental health resources and adolescent mental health service use. *Journal of the American Academy of Child & Adolescent Psychiatry, 52*, 501–510. 10.1016/j.jaac.2013.03.002

Guo, J., Wall, M., & Amemiya, Y. (2006). Latent class regression on latent factors. *Biostatistics, 7*, 145–163. doi:10.1093/biostatistics/kxi046

Hammen, C., Rudolph, K.D, (2003). Childhood mood disorders. In: Mash EJ, Barkley RA, editors. *Child Psychopathology*, *2nd Edition*. New York: Guilford Press; 2003. pp. 233–278.

Herman, K.C., Reinke, W.M., Thompson, A., & Hawley, K. (2019). The Missouri Prevention Center: A multidisciplinary approach to reducing the societal prevalence and burden of youth mental health problems. *American Psychologist, 74*, 315-328.

Herman, K.C., Reinke, W.M., & Thompson, A. (2019). Prevention Science: An Approach to Inform the Future of School Psychology. In M. Burns (Ed.), *Introduction to School Psychology: Controversies and Current Practice* (pp. 195-213). Oxford University Press.

Hinshaw, S.P., & Lee, S.S, (2003). Conduct and oppositional defiant disorders. In: Mash EJ, Barkley RA, editors. *Child Psychopathology*, *2nd Edition*. New York: Guilford Press, pp. 144–198.

Horner, R., Todd, A., Lewis-Palmer, T., Irvin, L. K., Sugai, G., & Boland, J., (2004). The School-wide Evaluation Tool (SET): A research instrument for assessing school-wide positive behavior supports. *Journal of Positive Behavior Interventions, 6,* 3-12.

Horner, R., Sugai, G., Kincaid, D., George, H., Lewis, T., Eber, L., ... & Algozzine, B. (2012). What Does it Cost to Implement School--wide PBIS? Research Gate. https://www.researchgate.net/profile/Heather_George/publication/267623671_What_Does_it_Cost_to_Implement_School--wide_PBIS/links/5453ab1c0cf2cf51647c1fad.pdf

Huang-Pollock, C.L., Mikami, A.Y., Pfiffner, L., & McBurnett, K, (2009). Can executive functions explain relation between ADHD and social adjustment? *Journal of Abnormal Child Psychology*, *37*, 679–691.

Kataoka, S. H., Zhang, L., & Wells, K. B. (2002). Unmet need for mental health care among U.S. children: Variation by ethnicity and insurance status. *The American Journal of Psychiatry, 159*, 1548-1555. doi: 10.1176/appi.ajp.159.9.1548

Kilgus, S. P., Reinke, W.M., Jimerson, S.R. (2015). Understanding mental health intervention and assessment within a multi-tiered framework: Contemporary science, practice, and policy. *School Psychology Quarterly, 30*, 159-165.

Kincaid D, Childs K, & George H (2010). School-wide Benchmarks of Quality (Revised). Tampa, Florida: University of South Florida.

Klein, J., Cornell, D., & Konold, T. (2012). Relationships between bullying, school climate, and student risk behaviors. *School Psychology Quarterly, 27*, 154–169. 10.1037/a0029350

Lapan, R. T., Gysbers, N. C., & Sun, Y. (1997). The impact of more fully implemented guidance programs on the school experiences of high school students: A statewide evaluation study. *Journal of Counseling & Development, 75(4),* 292-302.

Lochman, J. E., & Wells, K. C. (2002). The Coping Power program at the middle-school transition: Universal and indicated prevention effects. *Psychology of Addictive Behaviors, 16*, S40–S54. 10.1037/0893-164X.16.4S.S40

Marrast, L., Himmelstein, D., Woolhandler, S. (2016). Racial and ethnic disparities in mental health care for children and young adults: A national study. *International Journal of Health Services, 46,*810-824.

McDonald, R. P., & Ho, M.-H. R. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods, 7*, 64–82. doi:10.1037/1082-989X.7.1.64

McInstosh, K., & Goodeman, S. (2016). *Integrated Multi-Tiered Systems of Support Blending RTI and PBIS*. New York: Guilford Press.

McIntosh, K., Ty, S. V., & Miller, L. D. (2014). Effects of school-wide positive behavioral interventions and supports on internalizing problems: Current evidence and future directions. *Journal of Positive Behavior Interventions, 16*, 209–218. doi:10.1177/1098300713491980

McIntosh, K., Reinke, W.M., & Herman, K.C. (2010). School-wide analysis of data for social behavior problems: Assessing outcomes, selecting targets for intervention, and identifying need for support. In G. Peacock, R. Ervin, E. Daly, & K. Merrell (Eds.), *Practical handbook of school psychology: Effective practices for the 21st Century.* (pp. 135-156). New York: Guilford Press.

Merikangas, K. R., He, J. P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., ... & Swendsen, J. (2010). Lifetime prevalence of mental disorders in US adolescents: results from the National

Comorbidity Survey Replication–Adolescent Supplement (NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry, 49(*10), 980-989.

Muthén, B. (2003). Statistical and substantive checking in growth mixture modeling: Comment on Bauer and Curran (2003). *Psychological Methods, 8*(3), 369–377. doi:10.1037/1082-989X.8.3.369

Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 345–368). Newbury Park, CA: Sage Publications.

Muthén, B., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research, 24*(6), 882–891.

Muthén, B., & Muthén, L. K. (2009). *Mplus Users Guide*. Los Angeles, CA: Muthén and Muthén.

Muthen, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics, 6*, 463–469. doi:10.1111/j.0006-341X.1999.00463.x

Nagin, D. S., & Tremblay, R. E. (2001). Analyzing developmental trajectories of distinct but related behaviors: A group-based method. *Psychological Methods, 6*(1), 18–34. doi:10.1037/1082-989X.6.1.18

Neil, A. L., & Christensen, H. (2009). Efficacy and effectiveness of school-based prevention and early intervention programs for anxiety. *Clinical Psychology Review, 29*, 208–215. 10.1016/j.cpr.2009.01.002

Oakes, W. P., Lane, K. L., Cox, M. L., & Messenger, M. (2014). Logistics of behavior screenings: How and why do we conduct behavior screenings at our school? *Preventing School Failure: Alternative Education for Children and Youth, 58*, 159–170. doi:10.1080/1045988X.2014.895572

Pullmann, M. D., Ague, S., Johnson, T., Lane, S., Beaver, K., Jetton, E., & Rund, E. (2013). Defining engagement in adolescent substance abuse treatment. *American Journal of Community Psychology, 52*, 347–358. 10.1007/s10464-013-9600-8

Ramaswamy, V., DeSarbo, W., Reibstein, D., & Robinson, W. (1993). An empirical pooling approach for estimating marketing mix elasticities with PIMS data. *Marketing Science, 12*, 103–124.

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*, 1-13.

Reinke, W. M., Herman, K. C., & Tucker, C. (2006). Building and sustaining communities that prevent mental disorders: Lesson from the field of education. *Psychology in the Schools, 43*, 313-329.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*(2), 147–177. doi:10.1037/1082-989X.7.2.147

Schwartz, G. (1978). Estimating the dimensions of a model. *The Annals of Statistics, 6*, 461–464.

Sclove, L. S. (1987). Application of a model-selection criteria to some problems in multivariate analysis. *Psychometrika, 52*, 333–343. doi:10.1007/BF02294360

Shinn, M. R., & Walker, H. M. (Eds.). (2010). *Interventions academic and behavior problems in a three-tier model, including Response-to-Intervention*. Bethesda, MD: National Association of School Psychologists.

Sugai G, Lewis-Palmer T, Todd A, et al. (2001). School-wide Evaluation Tool (SET). Eugene, OR: Center for Positive Behavioral Supports. University of Oregon.

Suldo, S. M., Gormley, M. J., DuPaul, G. J., & Anderson-Butcher, D. (2014). The impact of school mental health on student and school-level academic outcomes: Current status of the research and future directions. *School Mental Health, 6*, 84–98. doi:10.1007/s12310-013-9116-2

Twenge, J. M., Cooper, A. B., Joiner, T. E., Duffy, M. E., & Binau, S. G. (2019). Age, period, and cohort trends in mood disorder indicators and suicide-related outcomes in a nationally representative dataset, 2005–2017. *Journal of Abnormal Psychology, 128*, 185–199.

Weisz, J. R., Jensen-Doss, A., & Hawley, K. M. (2006). Evidence-based youth psychotherapies versus usual clinical care: a meta-analysis of direct comparisons. *American Psychologist, 61(7),* 671.

Table 1. Fit Indices of Growth Mixture Modeling for Student Reported Total Social, Emotional and Behavioral Problems

LC	AIC	BIC	aBIC	Entropy
2	997833.588	997970.152	997916.127	0.73
3	996725.806	996886.470	996822.910	0.73
4	**996155.608**	**996340.372**	**996267.278**	**0.74**
5	995579.174	995788.037	995705.410	0.75
6	995333.918	995566.881	995474.720	0.70

Note. LC = Latent class; AIC = Akaike information criterion; BIC = Baysian information criterion; aBIC = adjusted Baysian information criterion. Smaller values indicate better fit of the model. Entropy values close to 1.0 indicate higher classification precision.

Figure 1. Developmental trajectories of total social, emotional, and behavioral problems over three years.