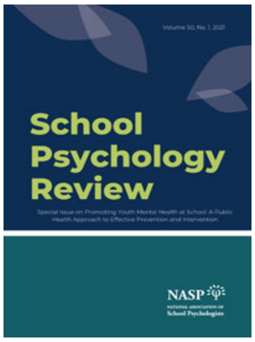


School Psychology Review

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/uspr20

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

Wendy M. Reinke, Keith C. Herman, Aaron Thompson, Christa Copeland, Chynna S. McCall, Shannon Holmes & Sarah A. Owens

SCHOOL PSYCHOLOGY REVIEW
2021, VOL. 50, NO. 1, 17–29
https://doi.org/10.1080/2372966X.2020.1870869

SPECIAL SERIES

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

Wendy M. Reinke[a,b], Keith C. Herman[a,b], Aaron Thompson[a,b], Christa Copeland[a,b], Chynna S. McCall[a,b], Shannon Holmes[a,b], and Sarah A. Owens[a,b]

[a]University of Missouri; [b]Missouri Prevention Science Institute

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 3-year period. Four classes of students were identified using growth mixture modeling: (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 with 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.

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.

ARTICLE HISTORY
Received June 7, 2020
Accepted December 28, 2020

KEYWORDS
prevention, school mental health, fidelity, evidence-based intervention

ASSOCIATE EDITOR
Kristin Hawley

INTRODUCTION

Children of all ages and abilities attend school every day, each with different social, emotional, and behavioral strengths and needs. However, many interrelated variables can impact these experiences, such as a child's externalizing, internalizing, social relationships, and school participation behaviors (Côté et al., 2006). Atypical manifestations of these variables can have detrimental impacts on students' relationships with their peers, school connectedness, and overall academic achievement (Dishion et al., 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 et al., 2014). In fact, 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 [Centers for Disease Control and Prevention (CDC), 2019; Twenge et al., 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 to 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 to 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 youth's mental health needs remain largely unmet at the national level, indicating a need for greater effort in this area. Meeting the needs of youth mental health is important because even the best efforts to support academic growth are undermined when student behavioral or emotional issues exist (Cooper et al., 2020).

CONTACT Wendy M. Reinke reinkew@missouri.edu Department of Educational, School, and Counseling Psychology, University of Missouri, 16 Hill Hall, CA 65211, USA.

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

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 et al., 2019b; Kilgus et al., 2015). 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. 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 et al., 2014). Less than 14% of schools report conducting universal screening of student social and emotional needs (Bruhn et al., 2014). Further, the use of teacher referrals or office discipline referrals to identify students are reactive and occur after precursors to these issues could have already been identified, and have repeatedly been associated with student failure (Bruns et al., 2006; Cash & Nealis, 2004). 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 intervene before a problem becomes ingrained or gets progressively worse (Oakes et al., 2014).

To address this problem, many schools utilize multitiered systems of support (MTSS; McIntosh & Goodman, 2016) to provide the appropriate level of services to each student more effectively. Through 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. Thus, identifying students who require these early intervention services is most effectively achieved by using 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 efficiently determine the social, emotional, and behavioral status of all students (Dowdy et al., 2015; McIntosh et al., 2010) and to determine needed supports for each student (Bruns et al., 2016). Using universal screening data within the context of a comprehensive tiered model of prevention and intervention supports can produce positive outcomes for students. However, this will only be achieved if these models are conducted with fidelity.

Fidelity to MTSS

Several studies suggest that implementing interventions as intended by developers, thereby with fidelity, is associated with better outcomes and increased chances of mirroring efficacy trial results (Durlak & DuPre, 2008; O'Donnell, 2008; Power et al., 2005). Although additional research is needed to confirm this relationship, it is clear that in order to make valid decisions about the effects of an intervention, it is necessary to know the fidelity with which it was implemented (Sanetti & Kratochwill, 2009). Whereas tiered models of prevention and intervention are well-founded empirically, the implementation of large-scale evidence-based programming and prevention frameworks is complex and presents unique challenges to schools. To be successful, schools need to adopt a systematic process for screening, data-based decision-making, and selecting and implementing evidence-based interventions at the universal, selective, and indicated levels (Herman et al., 2019a). 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., Dusenbury et al., 2003). Several barriers impede the implementation of these models, including lack of administrator support, organizational structure, and school personnel certification and training (Bradshaw & Pas, 2011; Domitrovich et al., 2008). As a result, implementation of such large-scale models often takes ~3–5 years to achieve high implementation quality and positive outcomes (Molloy et al., 2013; 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 association between fidelity to scaling up of positive behavior intervention supports (PBIS) and student achievement. Findings indicated that fidelity scores were associated with student math and reading achievement and truancy but not suspensions. 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 identify schools with low and high implementation quality, allow school professionals to better predict outcomes, and help determine implementation gaps. The following provides a description of the model used in this study, referred to as the Coalition model.

Description of the Coalition Model

Since 2016, a Coalition of six school districts and one parochial school, in collaboration with researchers, has 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 Thompson et. al., 2017; Reinke et al, 2018). The model utilizes the early identification system (EIS; Huang et al., 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, including internalizing behavior, externalizing behavior, emotion dysregulation, peer relationship problems, attention and academic issues, school disengagement, and relational aggression (Herman et al., 2020; Thompson et al., 2020). All data are gathered electronically and once surveys are complete, the data are populated in dashboard reports for schools to review in a red, yellow, green format (Reinke et al., 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. 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). Student reports are generated that indicate an area of need (red) for students when they are two standard deviations or higher than their peers within each screener's target area. Yellow indicates students are one standard deviation higher than their peers, and green indicates that they are within the normative range when compared with their peers.

Once universal screening is complete, school teams review the data-based reports, identify areas of concern, determine evidence-based 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 selected interventions. A measure to assess the fidelity to the overall model that includes gathering and using data, selecting appropriate interventions, implementing interventions, and evaluating the effectiveness of interventions implemented across all tiers was developed and utilized by the Coalition. Fidelity is multidimensional as defined in the literature, including adherence, dosage, quality, responsiveness of participants, and program differentiation (Dane & Schneider, 1998; Reinke et al., 2013). The Coalition measure targets adherence, the bottom-line measure of fidelity (Carroll et al., 2007) and was developed by identifying the essential components of the model, thus attending to program differentiation (Dusenbury et al., 2003). 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.

Demographic Characteristics Associated With Mental Health

Several demographic characteristics are associated with differential mental health outcomes, including socioeconomic status, race, gender, and age. For instance, youth who live in poverty are also more likely to have stressors related to poorer quality education, unsafe housing, and greater experiences with racial discrimination (Alegria et al., 2015). These stressors can lead to mental health issues over time. Furthermore, the proportion of Black (39%) and Latinx (33%) youth living in poverty is more than double that of non-Latino White (14%) and Asian (14%) youth (Kids Count Data Center, Children in Poverty 2014). Importantly, experiences with racism have been linked to poorer mental health. Youth from minority backgrounds are faced with racism that is systemic, meaning that racial bias and discrimination are weaved into all facets of society, including the schooling system. A comprehensive review of research on discrimination among children and adolescents found that exposure to discrimination predicted worse mental health (e.g., anxiety and depression symptoms) in 76% of the 127 associations examined (Priest et al., 2013). In addition, age and gender may be predictive of mental health problems. Diagnoses of depression and anxiety are more common with increased age and have been shown to affect girls to a greater extent than boys (Ghandour et al., 2019; Hamblin, 2016). For these reasons, free or reduced meal status (FRM; a proxy for SES), race, gender, and grade level (a proxy for age) will be examined as covariates in this study because these indicators are surrogates for societal contexts which place youth differentially at risk for mental health problems.

 DOI: 10.1080/2372966X.2020.1870869

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, a person-centered approach was applied to identify patterns of student risk over time. 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 were determined. First, the overall growth of social, emotional, and behavioral problems over a 3-year period of time was examined. Because all schools were partially to fully implementing the comprehensive mental health model with universal screening and tiered supports, a small and steady decline in youth-reported total problems was expected over the study period. Next, using growth mixture modeling (GMM), different subgroups of student total problems were expected to emerge with unique patterns of growth over time. Specifically, it was 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. It was 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. Several demographic variables were included in the final models because they are often associated with differential student risk for mental health problems.

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 countywide 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% Pacific Islander. Participants were from elementary (52.5%), middle (30.3%), and high school (17.2%) settings. Thirty-six percent of the sample received 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 et al., 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 Huang et al. (2019), Reinke et al. (2018), and Thompson et al. (2017). Students in 3rd through 12th grade completed the measure three times per year across 3 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 min 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 3 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: (a) data collection and review of universal screening data; (b) intervention planning and implementation across universal, selective, and indicated levels; and (c) 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 prepost 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 3-year window allowed for sufficient time for schools to reach full implementation of the model as well as for assessment of variability in implementation between schools. For the purpose of this study, the total score on the fidelity measure was calculated (α = .84). Conventional criteria for defining thresholds of effective implementation needed to reach student outcomes were followed (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, FRM status, and race.

Analysis

Latent growth curve (LGC) modeling and 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 3-year period. The analysis occurred in several stages. To test the first hypothesis that there would be a 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 < 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 et al., 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 et al., 2006) and examination of the odds ratio estimates. We also converted odds ratios to Cohen’s *d* estimates to provide an additional indication of effect sizes (Cohen, 1988). Further, descriptive data are provided that include the percentage of students in each trajectory.

Missing Data

Baseline data on the student report of total problems was used to determine if it predicted missingness. Although higher levels of total problems were a statistically significant predictor of missing data, this association was of low practical significance ($B = 0.026$ (.001) $p < .001$ [OR = 1.02]). The Mplus software utilizes full information maximum likelihood (FIML) estimation which is the state of the art approach to handle missing values under missingness at random (Arbuckle, 1996). The observed deviations from the missing at random assumption are incredibly small in the present dataset and thus FIML was retained as the method for handling missing data. Overall, 52% of the participants had at least five of the nine assessment time points. Further, there was no association between missing 4 or more time points with the fidelity measure ($\chi^2 = 0.69(1)$, $p = .41$). The minimum covariance coverage recommended for reliable model convergence is .10 (Muthén & Muthén, 2000). 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 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 when compared with White students. Asian ($B = -0.02$, $p < .05$) and Latinx students ($B = -0.03$, $p < .05$) reported lower levels of social, emotional, and behavioral problems in the fall of 2016 when compared with White students. Students receiving FRM reported higher levels of social, emotional, and behavioral problems in fall of 2016 ($B = 0.19$, $p < .001$). Lastly, high school students reported higher levels of social, emotional, and behavioral problems in the fall of 2016 ($B = 0.14$, $p < .001$).

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 when compared with White students ($B = -0.05$, $p < .001$). Also, female students had less decline in social, emotional, and behavioral problems over time when compared with 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 ($B = 0.03$, $p < .05$). Table 1 provides the results of the LGC model for total problems and covariates.

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 2). However, inspection of the

Table 1. Results for LGC Model for Total Problems With Covariates

	Intercept		Slope	
	ß	*B* (SE)	*ß*	*B* (SE)
Black	0.08***	2.52 (0.54)	−0.02	−0.27 (0.17)
Asian	−0.02*	−1.22 (0.57)	−0.05***	−1.17 (0.23)
Latinx	−0.03**	−1.47 (0.45)	0.002	0.05 (0.26)
Multiracial	0.04***	2.00 (0.46)	0.01	0.15 (0.19)
American Indian	0.02*	3.85 (1.66)	−0.002	−0.18 (0.55)
Middle School	−0.001	−0.04 (0.67)	0.04	0.43 (0.23)
High School	0.14***	3.82 (0.64)	−0.08*	−0.78 (0.36)
FRM	0.19***	4.79 (0.29)	0.03*	0.27 (0.14)
Sex	0.05*	1.10 (0.48)	−0.05**	−0.44 (0.16)

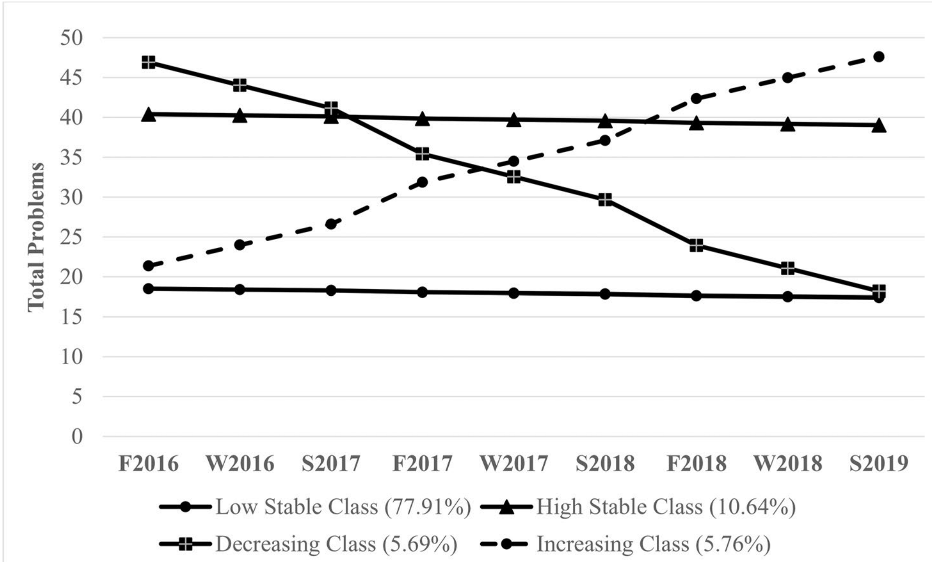
*$p < .05$, **$p < .01$, ***$p < .001$.

Table 2. Fit Indices of Growth Mixture Modeling for Student Reported Total Social, Emotional, and Behavioral Problems

LC	AIC	BIC	aBIC	Entropy
2	997,833.588	997,970.152	997,916.127	0.73
3	996,725.806	996,886.470	996,822.910	0.73
4	996,155.608	996,340.372	996,267.278	0.74
5	995,579.174	995,788.037	995,705.410	0.75
6	995,333.918	995,566.881	995,474.720	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 3 Years



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 (<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 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 1.30 times more likely to be in a lower fidelity school than students in the low stable class ($d = .14$), 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 ($d = .13$). Students in the decreasing social, emotional, and behavioral problems class, although not statistically significant, were equally as likely to be in a high fidelity school when compared with 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 ($d = .39$) times more likely to be Black, 1.71 ($d = .30$) times more likely to be Multiracial, 3.54 ($d = .70$) times more likely to receive FRM, and 3.02 ($d = .67$) times more likely to be in high school than students in the low stable class. Also, Asian students were 2.34 ($d = .47$) 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 ($d = .56$) times more likely to be in the increasing social, emotional, and behavioral problems class when compared with the low stable class, and Asian students and female students were less likely to be in this class (OR: 4.30 [$d = .80$] and 1.48 [$d = .22$], respectively). Lastly, students in the decreasing social, emotional, and behavioral problems class were more likely to be Black (OR: 1.77 [$d = .32$]), male (OR: 1.29 [$d = .14$]), and receiving FRM (2.85 [$d = .58$]) and less likely be in middle school (OR: 2.69 [$d = .55$]) than students in the low stable class. See Table 3 for GMM results.

 DOI: 10.1080/2372966X.2020.1870869

Table 3. Odds Ratios Showing Association Between Total Social, Emotional, and Behavioral Problems Subclasses, Fidelity, and Covariates

	High Stable	Decreasing	Increasing
Fidelity	1.30	1.08	1.28*
Black	1.78***	1.77***	1.20
Asian	0.43*	1.28	0.23**
Latinx	0.66	0.68	0.97
Multiracial	1.80*	1.26	1.16
American Indian	2.14	1.48	1.16
Middle School	1.27	0.37***	0.98
High School	2.30***	0.81	0.97
FRM	3.91**	2.85***	2.75***
Sex	1.09	1.29*	0.77*

Note. low fidelity= 1; FRM = 1; male = 1.
$^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

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 3–12 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 3-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 student social–emotion problems.

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 United States. Several national surveys have found consistent evidence of increasing youth mental health problems during the past decade (CDC, 2019; Twenge 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 social, emotional, and behavioral health risk factors rather than waiting until those issues result in office referrals, absenteeism, suspensions, and poor academic performance. Although causal inferences cannot be determined by the 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.

The subsequent analyses in this study examined the relations between fidelity of implementation and growth patterns of student risk over time. Using GMM, a person-centered approach, to identify patterns of student risk over time, identified four types of trajectories. The majority of students (78%) were characterized by a normative pattern of low stable levels of social, emotional, and behavioral problems across the 3 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 class patterns. These patterns are closely aligned with public health models, which indicate that in optimal preventive approaches where universal supports are adequate, between 15% and 20% of the student population would benefit from more selective or indicated interventions (Herman et al., 2019b). Most importantly, students with increasing social, emotional, and behavioral problems were more likely to be in schools with lower fidelity to the model at year three of implementation. Thus, implementation of the Coalition model may be associated with lower student risk for maladaptive trajectories of social, emotional, and behavior risk over time.

The effect sizes of fidelity on student trajectories ranged from small to moderate using traditional interpretation guidelines (Cohen, 1988). However, even small effects on population-level interventions such as the Coalition can dramatically benefit population health (National Research Council and Institute of Medicine, 2009). Moreover, many scholars have advocated for the refinement of traditional guidelines to include attention to context, typically observed effect sizes, and potential for the cumulation of effects over time (Funder & Ozer, 2019; Kraft, 2020). As one example, based on comparative empirical benchmarks, Dynarski (2017) concluded that "in real-world settings, a fifth of a standard deviation [0.20 SD] is a large effect." Moreover, a review of preregistered, randomized trials of promising educational interventions found a median effect size of .03 (see Yeager et al., 2019). Using these collective guidelines—attentive to the scale, duration, and context of the Coalition intervention—would place these effects in the medium to large range (Dynarski, 2017; Funder & Ozer, 2019; Kraft, 2020).

Sociodemographic characteristics also predicted student trajectories. Youth of color are more likely to live in poverty and experience discrimination, both of which are linked to poorer mental health problems (Kids Count Data Center, Children in Poverty, 2014; Priest et al. 2013). 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 to note because consistent evidence suggests that youth of color are significantly less likely to seek or be referred for mental services (Marrast et al., 2016). Thus, the universal screening and supports offered by a model like the Coalition holds the 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 to fall into a class characterized by decreasing symptoms when compared with 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 countywide 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 3 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 multitiered 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. In this project, all 54 schools implemented universal screening triannually over a 3-year period and most schools used these data in appropriate ways to guide decisions about multitiered supports. Regardless, the results of these analyses strongly suggest that schools that more fully adhered to the model experienced better student outcomes when compared with schools with lower rates of adherence.

Limitations

The study describes the countywide 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 GMM analyses to describe patterns of student growth over time. This rich data stream allowed for the examination of 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

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 et al., 2008; Horner et al., 2012; Molloy et al., 2013). Thus, it is reasonable to categorize schools as high and low implementers at this 3-year time-point. Furthermore, the fidelity measure focused almost entirely on adherence to the essential components of the model. Thus, the quality of implementation and responsiveness of participants to the program were not measured. Also, we were unable to use independent observers to assess fidelity in this study. Both expanding the measurement of fidelity to the model and using independent observers are areas for future research. 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 the decision to change their culture and context 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. However, 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 3-year period. Given that each school that reached the 80% threshold likely did so at different points in the prior 3 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

Many youth come to schools in need of additional supports for their mental health. Students struggling with mental health issues will not meet their full academic and behavioral potential without needed supports. 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. Implementing these practices with high fidelity will likely lead to better and important outcomes for youth.

DISCLOSURE

The authors have no conflicts of interest to report.

FUNDING

This work was supported by Children Services Fund and Institute of Education Sciences (R305H170023 and R305C190014).

ORCID

Shannon Holmes https://orcid.org/0000-0001-8719-9224

REFERENCES

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

Alegria, M., Green, J. G., McLaughlin, K. A., & Loder, S. (2015). Disparities in child and adolescent mental health and mental health services in the U.S. http://www.mamh.org/Portals/0/Uploads/Documents/Public/Disparities%20in%20Child%20and%20Adolescent%20Mental%20Health.pdf

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*. 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. https://doi.org/10.1037/1082-989X.8.3.384

Bradshaw, C. P., & Pas, E. T. (2011). A statewide scale up of positive behavioral interventions and supports: A description of the development of systems of support and analysis of adoption and implementation. *School Psychology Review*, *40*(4), 530–548. https://doi.org/10.1080/02796015.2011.12087528

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. https://doi.org/10.1353/etc.2014.0039

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*(2), 156–170. https://doi.org/10.1037/ort0000083

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science*, *2*(1) 40. https://doi.org/10.1186/1748-5908-2-40

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

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

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Erlbaum.

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*(2), 86–95. https://doi.org/10.1037/pla0000114

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*(1), 68–85. https://doi.org/10.1007/s10802-005-9001-z

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

Dane, A., & Schneider, B. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review*, *18*, 23–45.

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*(3), 603–619. https://doi.org/10.1017/S0954579410000313

Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., Romanelli, L. H., Leaf, P. J., Greenberg, M. T., & Ialongo, N. S. (2008). Maximizing the

implementation quality of evidence-based preventive interventions in schools: A conceptual framework. *Advances in School Mental Health Promotion*, *1*(3), 6–28. https://doi.org/10.1080/1754730x.2008.9715730
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*(2-3), 178–197. https://doi.org/10.1080/10474412.2014.929951
Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, *41*(3-4), 327–350. https://doi.org/10.1007/s10464-008-9165-0
Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research*, *18*(2), 237–256. https://doi.org/10.1093/her/18.2.237
Dynarski, S. M. (2017). *For better learning in college lectures, lay down the laptop and pick up a pen*. The Brookings Institution.
Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168. https://doi.org/10.1177/2515245919847202
Ghandour, R., Sherman, L., Vladutiu, C., Ali, M., Lynch, S., Bitsko, R., & Blumberg, S. (2019). Prevalence and treatment of depression, anxiety, and conduct problems in US children. *The Journal of Pediatrics*, *206*, 256–267. https://doi.org/10.1016/j.jpeds.2018.09.021
Guo, J., Wall, M., & Amemiya, Y. (2006). Latent class regression on latent factors. *Biostatistics (Oxford, England))*, *7*(1), 145–163. https://doi.org/10.1093/biostatistics/kxi046
Hamblin, E. (2016). Gender and children and young people's emotional and mental health: Manifestations and responses. A rapid review of the evidence. National Children's Bureau. https://www.ncb.org.uk/resources/all-resources/filter/well-being-mental-health/gender-sensitive-approaches-addressing
Herman, K. C., Reinke, W. M., Huang, F. L., Thompson, A. M., & Doyle-Baker, L. (2020). Investigating the psychometric properties of the Early Identification System – Student Report in a middle school sample. *School Psychology*. Advance online. https://doi.org/10.1037/spq0000414
Herman, K. C., Reinke, W. M., & Thompson, A. (2019a). 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.
Herman, K. C., Reinke, W. M., Thompson, A., & Hawley, K. (2019b). The Missouri Prevention Center: A multidisciplinary approach to reducing the societal prevalence and burden of youth mental health problems. *The American Psychologist*, *74*(3), 315–328. https://doi.org/10.1037/amp0000433
Horner, R., Sugai, G. M., Kincaid, D., George, H., Lewis, T., Eber, L., Barret, S. B., & 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
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*(1), 3–12. https://doi.org/10.1177/10983007040060010201
Huang, F., Reinke, W. M., Thompson, A., Herman, K. C. & County Schools Mental Health Coalition. (2019). An Investigation of the psychometric properties of the Early Identification System (EIS)-Student Report. *Journal of Psychoeducational Assessment*, *37*, 473–485.
Kids Count Data Center. Children in Poverty. (2014). *The Annie E. Casey Foundation, KIDS COUNT Data Center*. https://datacenter.kidscount.org.
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*(2), 159–165. https://doi.org/10.1037/spq0000118
Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, *49*(4), 241–253. https://doi.org/10.3102/0013189X20912798
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.
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: Planning, Administration, Evaluation*, *46*(4), 810–824. https://doi.org/10.1177/0020731416662736
McDonald, R. P., & Ho, M.-H. R. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods*, *7*(1), 64–82. https://doi.org/10.1037/1082-989x.7.1.64
McIntosh, K., & Goodman, S. (2016). *Integrated multi-tiered systems of support blending RTI and PBIS*. Guilford Press.
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). 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*(4), 209–218. https://doi.org/10.1177/1098300713491980
Merikangas, K. R., He, J. P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., Benjet, C., Georgiades, K., & 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.
Molloy, L., Moore, J., Trail, J., Van Epps, J., & Hopfer, S. (2013). Understaning the real-world implementation quality and "active ingredients" of PBIS. *Prevention Science*, *14*(6), 593–605. https://doi.org/10.1007/s11121-012-0343-9
Muthén, B. (2003). Statistical and substantive checking in growth mixture modeling: Comment on Bauer and Curran (2003). *Psychological Methods*, *8*(3), 369–377. https://doi.org/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). 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. (2017). *Mplus Users Guide*. Muthén and Muthén.

National Research Council and Institute of Medicine. (2009). *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities*. National Academies Press.

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*(3), 159–170. https://doi.org/10.1080/1045988X.2014.895572

O'Donnell, C. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research*, *78*(1), 33–84. https://doi.org/10.3102/0034654307313793

Pas, E. T., & Bradshaw, C. P. (2012). Examining the association between implementation and outcomes: State-wide scale-up of school-wide positive behavior intervention and supports. *The Journal of Behavioral Health Services & Research*, *39*(4), 417–433. https://doi.org/10.1007/s11414-012-9290-2

Patterson, G. R., Reid, J. B., & Dishion, T. (1992). *Antisocial boys: A social interactional approach*. (Vol. 4). Castalia Publishing Company.

Power, T. J., Blom-Hoffman, J., Clarke, A. T., Riley-Tillman, T. C., Kelleher, C., & Manz, P. H. (2005). Reconceptualizing intervention integrity: A partnership-based framework for linking research with practice. *Psychology in the Schools*, *42*(5), 495–507. https://doi.org/10.1002/pits.20087

Priest, N., Paradies, Y., Trenerry, B., Truong, M., Karlsen, S., & Kelly, Y. (2013). A systematic review of studies examining the relationship between reported racism and health and wellbeing for children and young people. *Social Science & Medicine*, *95*, 115–127.

Ramaswamy, V., DeSarbo, W., Reibstein, D., & Robinson, W. (1993). An empirical pooling approach for estimating marketing mix elasticities with PIMS data. *Marketing Science*, *12*(1), 103–124. https://doi.org/10.1287/mksc.12.1.103

Reinke, W. M., Herman, K. C., Stormont, M., Newcomer, L., & David, K. (2013). Illustrating the multiple facets and levels of fidelity of implementation to a teacher classroom management intervention. *Administration and Policy in Mental Health*, *40*(6), 494–506. https://doi.org/10.1007/s10488-013-0496-2

Reinke, W. M., Thompson, A., Herman, K. C., Holmes, S., Owens, S., Cohen, D., Tanner-Jones, L., Henry, L., Green, A., Copeland, C., & County Schools Mental Health Coalition (2018). The County Schools Mental Health Coalition: A model for community level impact. *School Mental Health*, *10*(2), 173–180. https://doi.org/10.1007/s12310-017-9227-2

Rimm-Kaufman, S. E., Fan, X., Chiu, Y. J., & You, W. (2007). The contribution of the responsive classroom approach on children's academic achievement: Results from a three year longitudinal study. *Journal of School Psychology*, *45*(4), 401–421. https://doi.org/10.1016/j.jsp.2006.10.003

Sanetti, L. M. H., & Kratochwill, T. R. (2009). Toward developing a science of treatment integrity: Introduction to the special series. *School Psychology Review*, *38*(4), 445–459.

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*(3), 333–343. https://doi.org/10.1007/BF02294360

Sugai, G., & Horner, R. (2006). A promising approach for expanding and sustaining school-wide positive behavior support. *School Psychology Review*, *35*(2), 245–259. https://doi.org/10.1080/02796015.2006.12087989

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*(2), 84–98. https://doi.org/10.1007/s12310-013-9116-2

Thompson, A. M., Reinke, W. M., Holmes, S., Danforth, L., Herman, K. C., & the County School Mental Health Coalition. (2017). The county school mental health coalition: A model for a systematic approach to supporting youth. *Children & Schools*, 39, 209–218.

Thompson, A., Huang, F., Smith, T., Reinke, W. M., & Herman, K. C. (2020). Confirmatory factor structure an predictive validity of the Early Identification System-Student Report in a community sample of high school students. *School Mental Health*. https://doi.org/10.1007/s12310-020-09396-5

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*(3), 185–199. https://doi.org/10.1037/abn0000410

Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., Tipton, E., Schneider, B., Hulleman, C. S., Hinojosa, C. P., Haunesku, D., Romero, C., Flint, K., Roberts, A., Trott, J., Iachan, R., Buontempo, J., Yang, S. M., Carvalho, C. M., ... Dweck, C. S. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature, 573*, 364–369.

AUTHOR BIOGRAPHICAL STATEMENTS

Dr. Wendy M. Reinke is a Professor in the Department of Education, School, and Counseling Psychology at the University of Missouri. She is the cofounder and codirector of the Missouri Prevention Science Institute. She has an extensive grant and publication record including over 100 peer-reviewed publications and over $40 million in grant funding in the areas of prevention and early intervention of child emotional and behavior disturbances. She is also the Director of the National Center for Rural School Mental Health and the codeveloper and leadership team member for the Family Access Center of Excellence and the Boone County Schools Mental Health Coalition.

Dr. Keith C. Herman is a Curator's Distinguished Professor in Department of Education, School, and Counseling Psychology at the University of Missouri. He is the cofounder and codirector of the Missouri Prevention Science Institute. He has an extensive grant and publication record including over 120 peer-reviewed publications in the areas of prevention and early intervention of child emotional and behavior disturbances and culturally sensitive education interventions.

Dr. Aaron M. Thompson is an Associate Professor in School of Social Work at the University of Missouri. He is the Associate Director of the Missouri Prevention Science Institute. Dr. Thompson developed Self-management Training And Regulation Strategy (STARS) intervention, an evidence-based Tier 2 support for youth with disruptive behaviors. He is also the codeveloper and leadership team member for the Family Access Center of Excellence and the Boone County Schools Mental Health Coalition.

Dr. Christa Copeland is a postdoctoral fellow at the Missouri Prevention Science Institute. She earned her doctoral degree in school psychology from the University of Missouri. Her research interests focus on implementation science and evidence-based practices in schools.

Dr. Chynna McCall is a postdoctoral fellow at the Missouri Prevention Science Institute. She earned her doctoral degree from the University of Northern Colorado. Her research interests focus on social emotional learning and equity.

Dr. Shannon Holmes is an Assistant Professor in the University of Missouri school psychology program and a faculty affiliate of the Missouri Prevention Science Institute. Her research interests focus on implementation science, particularly the measurement of fidelity to implementation to improve school intervention practices.

Dr. Sarah Owens is a Clinical Assistant Professor in school psychology at the University of Missouri and the Director of the Boone County Schools Mental Health Coalition. She is a graduate of the school psychology program at the University of Missouri. Her research interests focus on implementation science and fidelity to intervention.