The Report Committee for Benjamin Judd Paly certifies that this is the approved version of the following Report:

Comparing Two School-Based Methods for Identifying Behavioral and Emotional Risk in Youth: Traditional Identification Practices and Self-Report Universal Screening

Committee:

Erin Rodriguez, Supervisor

David Klingbeil

Comparing Two School-Based Methods for Identifying Behavioral and Emotional Risk in

Youth: Traditional Identification Practices and Self-Report Universal Screening

by

Benjamin Judd Paly

Report

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

May 2019

Abstract

Comparing Two School-Based Methods for Identifying Behavioral and Emotional

Risk in Youth: Traditional Identification Practices and Self-Report Universal

Screening

Benjamin Judd Paly, M.A.

The University of Texas at Austin, 2019

Supervisor: Erin Rodriguez

Traditional identification methods in schools for determining students at-risk for emotional and behavioral disorders tend to rely on teacher referral. There is evidence that systematic approaches to screening for emotional and behavioral risk more effectively capture the full range of internalizing and externalizing symptoms and are less vulnerable to biases. The proposed study seeks to compare traditional identification methods for identifying youth with elevated behavioral and emotional risk (BER) with a self-report universal screening procedure in a school setting. The study will explore discrepancies between the two identification methods, including the degree to which they agree/disagree on "at risk" students, the racial/ethnic, socioeconomic status (SES), and language characteristics of identified students, and the presenting symptoms of students identified via the two methods. It is hypothesized that the two identification methods will frequently disagree on the risk status of individual students and that patterns based on racial/ethnic background, SES, language status, and symptom presentation will emerge. Data will analyzed using chi-squared goodness of fit, one-way repeated measures ANOVA, and Cohen's kappa coefficient.

Table of Contents

Introduction

Approximately one in every 4 - 5 youth meets the criteria for a mental health disorder that causes impairment over the course of their lifetime (Merikangas et al., 2010). Youth with mental health disorders are more likely to experience numerous negative long-term outcomes that affect quality of life, including academic failure, economic hardship, and poorer physical and mental health in adulthood (Bradshaw, Schaeffer, Petras, & Lalongo, 2010). Studies have consistently found that early prevention for youth at elevated risk for emotional and behavioral disorders is more likely to lead to successful outcomes and is more cost-effective than allowing early risk to develop over time into psychopathology (Saxena et al., 2004). Schools are considered by many to be an ideal place to identify youth mental health needs because of their universal coverage (Dowdy, Ritchey, & Kamphaus, 2010). Schools typically identify students through a teacher-initiated referral process that relies heavily on teacher discretion (Kalberg, Lane, & Menzies, 2010). In recent years, there has been a trend towards a school-based public health model that includes universal screening for emotional and behavioral risk (Albers, Glover, & Kratochwill, 2007). However, recent surveys suggest this model is still only utilized in a small minority of schools across the country (Bruhn, Woods-Groves, & Huddle, 2014).

The following proposed study addresses a key aspect of the provision of mental health services: the early and accurate identification of behavioral and emotional risk in youth. The proposed study will investigate how the results of a universal screening

1

procedure for the identification of behavioral and emotional risk compares with traditional identification practices that are widespread in schools today. The proposed study will first investigate disparities in identification rates through traditional identification methods based on student characteristics such as racial/ethnic background, socioeconomic status, and language status. The research questions also include direct comparisons between traditional identification practices and universal screening to determine the extent to which identification rates differ based on racial/ethnic background, socioeconomic status, language status, and symptom presentation. Finally, the researchers will conduct analyses on the level of agreement between the two identification methods. Results of the study will contribute the literature on the school-based identification of mental health needs in youth and the potential benefits of a universal screening procedure.

Integrative Analysis

EMOTIONAL AND BEHAVIORAL DISORDERS

Emotional and behavioral disorders (EBDs) encompass a broad, multidimensional range of difficulties related to youth mental health, and include but are not limited to externalizing disorders, mood disorders, anxiety disorders, and psychotic disorders (Rutherford, Quinn, Mathur, Rutherford Jr., & Rutherford Jr., 2014). Students with EBDs often struggle to build and maintain relationships with adults and peers, develop self-awareness and emotion regulation, adhere to rules of conduct, and complete age-appropriate academic work (Halfon & Newacheck, 1999; Wheeler & Mayton, 2014).

There is an urgent need to meet the mental health needs of youth (Mitchell, Tynes, Umaña-Taylor, & Williams, 2015). Early social-emotional and behavioral challenges are associated with a multitude of negative long-term outcomes, including poor academic performance, school dropout, substance abuse, early pregnancy, unemployment, and socioeconomic disadvantage in adulthood (McLeod & Kaiser, 2004; Mitchell et al., 2015; Sheridan et al., 2017). Youth mental health issues such as adolescent depression have been linked to poorer long-term health outcomes, including higher health-care utilization and increased work impairment due to physical health (Halfon & Newacheck, 1999; Keenan-Miller, Hammen, & Brennan, 2007). From a financial perspective, the burden of youth mental health challenges is substantial; when considering treatment costs, productivity losses, poor health, and criminal activity, the burden of youth mental health difficulties in the United States is estimated to cost $247 billion annually (Griffith, 2010).

Half of all lifetime mental health disorders start by the age of 14 (Kessler et al.,

2005). There is robust evidence to support the importance of early intervention for youth

with or at risk for mental health difficulties (Saxena et al., 2004; Weissberg & Bell,

1997). Preventative intervention when youth are at risk for a mental health disorder

represents the most effective and cost-effective strategy for reducing long-term burden

(McGorry & Purcell, 2009). Screening for EBDs represents one of the first steps in

preventative efforts (Albers et al., 2007).

Prevalence of EBDs

The U.S. Department of Health and Human Services (DHHS) estimates that at

least one in five youth has a mental health disorder at some point during childhood and

adolescence, and that at least 10% have a serious emotional disturbance during this time

period that dramatically affects their ability to function socially, academically, and

emotionally (Brauner & Stephens, 2006).

Anxiety disorders are the most common mental health disorder in children; 12-

month estimates range from 8.6% to 20.9% for any anxiety disorder (Costello, Egger, &

Angold, 2005). According to data from the National Comorbidity Survey-Adolescent

Supplement (NCS-A), a nationally representative sample of over 10,000 youths aged 13-

18, the lifetime prevalence of Major Depressive Disorder is 11% and the 12-month

prevalence is 7.5% (Avenevoli, Swendsen, He, Burstein, & Merikangas, 2015). The

NCS-A found that 9.6% of youth meet the criteria for a behavior disorder (Merikangas et

al., 2010). Regarding diagnoses from a mental health professional, it is estimated that

nearly 8% of children in the United States have been diagnosed with anxiety or

depression, and 5.4% have been diagnosed with behavior or conduct disorders

(Ghandour, Kogan, Blumberg, Jones, & Perrin, 2012).

Sociodemographic Variations in Prevalence of EBDs

There is an ongoing discussion about the associations between race, ethnicity,

socio-economic status, and other sociodemographic variables and the prevalence of

mental health morbidity (Williams & Earl, 2007). Large epidemiological studies have

shown relatively similar prevalence rates for major classes of mental health disorders

across racial and ethnic subgroups (Merikangas et al., 2010). Other studies have found

small variations in prevalence rates based on racial and ethnic background (Alegría et al.,

2012; Breslau, Kendler, Su, Gaxiola-Aguilar, & Kessler, 2005). A review of psychiatric

disorders using data from the NCS-A determined that Hispanic youth had a lower

lifetime risk of substance use than Non-Hispanic white individuals, and that Non-

Hispanic black individuals had a lower lifetime risk for mood, anxiety, and substance use

disorders (Breslau et al., 2005). Research from the Center for Disease Control (CDC)

suggests higher rates of suicidal ideation and attempted suicide in Latino adolescents

when compared to Non-Hispanic white and black youth (Center for Disease Control,

1999). Within the Latino subgroup in the United States, researchers found an increased

rate of psychiatric disorders among US-born, English language proficient, and third

generation individuals (Alegría et al., 2012). A study of Mexican-Americans similarly

5

determined that US-born Mexican Americans had significantly higher rates of mental health disorders than Mexican immigrants (Vega et al., 1998).

There is agreement in the field that low socioeconomic status (SES) is a risk factor for mental health problems. SES can be conceptualized to include an index of indicators such as household income, parental education, and parental occupation (Reiss, 2013). A meta-analysis of 55 studies addressing the relationship between SES and mental health found that disadvantaged youth were two to three times more likely to develop psychopathology (Reiss, 2013). A systematic review of youth depression and anxiety specifically found that the prevalence of these internalizing disorders was 2.49 times higher in youth from low-SES backgrounds (Lemstra et al., 2008). More research is needed to understand the extent of variations in prevalence and the ways in which referral disparities, access to mental health services, and additional cultural and sociodemographic characteristics may influence the prevalence and treatment of mental health disorders (Flores et al., 2002).

Access to Care

A 2014 review of service use for mental, emotional, and behavioral disorders in children and adolescents found that only 45% of youth with a mental health diagnosis receive treatment of any kind, and 24% of those individuals receive care within the school system (Costello, He, Sampson, Kessler, & Merikangas, 2014). Given that access to mental health treatment is far from universal (Kataoka, Zhang, & Wells, 2002), the low use of services amongst even those who had access to a diagnosis represents a dramatic underutilization of mental health services in community and school settings across the general population. An analysis of a nationally

6

representative sample of children ages six to 11 determined that among children with EBDs, 17.8% received both medication and psychosocial services, 28.8% received psychosocial services only, 6.8% received medication only, and 46.6% received neither service (Simon, Pastor, Reuben, Huang, & Goldstrom, 2015). Since the early 1990s, schools have been the primary care delivery setting for child mental health needs (Atkins, Cappella, Shernoff, Mehta, & Gustafson, 2017). The analysis by Simon et al. (2015) found that 18.6% of children with EBDs received services in school only, 11.4% received services in a community care setting, and 17.3% received psychosocial services in both settings.

Academic investigations of the barriers to care for the treatment of mental health disorders have identified numerous obstacles that prevent the delivery of services, including but not limited to health insurance constraints, other financial concerns, a lack of transportation, the stigma of mental health services, and a lack of well-qualified providers (Committee on School Health, 2004). Even if families are able to initiate services, there is a significant association between barriers to care and dropping out of treatment (Kazdin, Holland, & Crowley, 1997). Given these barriers to care in community settings and the universal nature of the school setting, there is an emerging consensus that schools are uniquely positioned to be a primary provider of mental health assessments and interventions for youth (Atkins, Hoagwood, Kutash, & Seidman, 2010).

Disparities in Access to Care

The National Institutes of Medicine defines disparity as "differences in treatment or access not justified by the differences in health status or preferences of the groups" (Smedley, Stith, & Nelson, 2003). Differences in health status between groups do not

necessarily represent a disparity; only differences that negatively and systematically affect an underprivileged group are classified as disparities (Dehlendorf, Bryant, Huddleston, Jacoby, & Fujimoto, 2010). Racial/ethnic minority children are more likely to have unmet mental health needs, including less overall access to care, delays in treatment, lower quality of care, and premature termination (Kataoka et al., 2002; Snowden & Yamada, 2005). As discussed previously, differences in service utilization rates are not a reflection of differences in prevalence rates amongst minority youth (Merikangas et al., 2010). Factors influencing this underutilization of services include practical barriers (e.g. insurance and language status) and cultural beliefs (e.g. attitudes towards care) (Guo, Kataoka, Bear, & Lau, 2014). An analysis of 2002-2007 nationally representative Medical Expenditure Panel Surveys (MEPS) data found persistent racial/ethnic disparities across multiple measures of mental health care usage, including any mental health care, any outpatient mental health care, and any psychotropic drug usage (Cook, Barry, & Busch, 2013). Another review of 2006-2012 MEPS data determined that black and Hispanic children averaged significantly fewer mental health visits than their White peers after adjusting models for other demographic characteristics (age, gender, geographic location), mental health impairment, and insurance status (Marrast, Himmelstein, & Woolhandler, 2016).

Additional factors, such as the location of service delivery, problem type, and critical points in the care process, provide further information on the nature of disparities in access to care. An analysis of Medicaid claims for 23,601 children found differences in service usage between racial/ethnic groups for both in-school and out-of-school service

8

usage; for all disorders, Hispanic children had significantly lower usage of in-school mental health services (Locke et al., 2017). However, racial/ethnic disparities in usage appear to be smaller in school settings than in clinical settings, suggesting that some of the barriers to care that minority parents face in seeking clinical services may not inhibit school-based services, and that schools represent a critical opportunity for addressing overall disparities in the provision of mental health services (Cummings, Ponce, & Mays, 2010). Disparities also exist based on symptom presentation, and can be partially attributed to the perception that there is a greater need for intervention for disruptive disorders compared to internalizing disorders (Wu et al., 1999). Students are significantly more like to receive services for externalizing disorders, such as behavior disorders and attention-deficit hyperactivity disorder (ADHD), than for internalizing disorders, including anxiety and depression (Merikangas et al., 2011). Racial/ethnic disparities in service utilization also appear to be influenced by problem type (Gudiño, Lau, Yeh, McCabe, & Hough, 2009). A two-year longitudinal study of youth mental health services usage found that Non-Hispanic White youth were the only group for whom exclusively internalizing problems at baseline led to higher rates of mental health service usage, while minority youth with externalizing and/or comorbid problems at baseline were more likely to receive mental health services than their Non-Hispanic White peers (Gudiño et al., 2009). When considering critical points in the process of providing care, it appears that disparities in the initiation of services are a primary driver of disparities in overall mental health care usage, suggesting that policies to improve identification and reduce barriers to initial access to care are vital for reducing disparities (Cook et al., 2013).

9

TRADITIONAL IDENTIFICATION PRACTICES

Throughout this study proposal, *traditional identification practices* will be used as an umbrella term to cover the most common ways in which students are currently identified in schools as needing mental health services. While there is variation in "typical" identification practices across school districts, they are often teacher initiated, and tend to rely heavily on patterns of office discipline referrals (ODRs) and teacher anecdotes about student behavior (Kalberg et al., 2010). Teachers spend substantial time with their students each day and are an invaluable resource for identifying students in need of behavioral, emotional, and academic supports (Anderson, Lubig, & Smith, 2012). However, a relatively unstructured, teacher-initiated identification process may result in challenges to equitable identification, due to biases against certain groups of students and biases regarding the necessity for treatment of different symptom presentations. These challenges will be discussed in more detail below.

The concepts of disparity and bias both warrant examination in the context of typical identification practices in schools. Disparity in identification practices refers to the over- or under-identification of the mental health needs of certain groups of students in a way that systematically disadvantages underprivileged groups (Dehlendorf et al., 2010). Biases are errors based on beliefs and emotions, either conscious or subconscious, that are wrong or irrelevant and may adversely affect specific groups of people, such as those of a specific racial/ethnic group (Couchenour & Chrisman, 2016). In the context of identifying mental health risk, it is important to consider the ways in which biases towards certain racial/ethnic groups may be one source of disparities in the identification of mental health problems and the provision of mental health services. Other potential causes of disparities, including cross-cultural factors and socioeconomic variables, will also be discussed.

10

Numerous academic studies have documented the ways in which disparities exist in the school-based referral processes (Guo et al., 2014; Tenenbaum & Ruck, 2007). Disparities in discipline practices must also be explored because discipline patterns are often factored into traditional identification practices; patterns of ODRs are one way that students are typically flagged as potentially needing additional behavioral and emotional supports (Kalberg et al., 2010).

Disparities in Traditional Identification and Discipline Practices

School-based disparities in identification and discipline practices based on student race/ethnicity have been extensively catalogued in academic literature (Martinez, McMahon, & Treger, 2016; Carter, Skiba, Arredondo, & Pollock, 2017; Krezmien, Leone, & Achilles, 2006; Silva, Langhout, Kohfeldt, & Gurrola, 2015). A study of Latino and Asian American youth found a significant direct effect of race and ethnicity on school-based mental health referrals after controlling for externalizing problems, school bonding, impairment, and academic performance (Guo et al., 2014). Latino students were four times more likely to be referred for mental health services than Asian American students (Guo et al., 2014). In a review of four meta-analyses on teacher referral practices, three of the four meta-analyses found a small but significant effect of race/ethnicity, concluding that Latino and African American students were more likely to receive referrals about negative behaviors and less likely to receive referrals for positive behaviors than their White peers (Tenenbaum & Ruck, 2007). A 2012 review of the National Comorbidity Survey Adolescent Supplement found almost no racial or ethnic differences in the identification of mental health needs by school personnel (Alegría et

11

al., 2012), which suggests that further research is needed on the extent to which race and ethnicity influences referral practices.

Racial and ethnic disparities in discipline practices are well documented (Krezmien et al., 2006). Regression analyses of 2005-2006 discipline data for 364 elementary and middle schools indicated that African American students were 2.19 (elementary school) to 3.78 (middle school) times more likely to receive an office discipline referral for problem behavior than their White peers (Skiba et al., 2011). In a study of a school-wide behavioral intervention, Silva et al. (2015) found that African American and Latino boys were significantly more likely to receive a bad conduct report for safety or self-responsibility. The literature on this topic suggests that relying heavily on teacher discretion and disciplinary mechanisms for identification of mental health needs may disproportionately affect specific racial and ethnic groups.

Implicit Bias

Teacher referral is the most common pathway for the identification of mental health needs, and is also considered one of the most vulnerable parts of the system to bias, as behavioral and performance expectations vary amongst teachers and it relies heavily on individual discretion (Severson, Walker, Hope-Doolittle, Kratochwill, & Gresham, 2007). There is evidence that teacher biases towards certain racial/ethnic groups of youth contribute to disparities in the identification of mental health needs and discipline practices (Skiba, Michael, Nardo, & Peterson, 2002; Tenenbaum & Ruck, 2007). As explicit racial attitudes have become less biased in the latter part of the 20th

12

century and into the 21st century, racial inequality and discriminatory outcomes are now frequently linked to implicit biases (Devine, Forscher, Austin, & Cox, 2012). Implicit biases are unconscious and involuntarily activated beliefs that include negative evaluations of individuals based on their membership in a social group, such as a gender or race (Kelly, 2013). The effects of implicit bias can be situationally exacerbated, such as situations that involve ambiguity, time constraints, or cognitive overload, all of which frequently occur for teachers in the classroom setting (Staats, 2016). A review of the academic literature on implicit bias in the school setting reveals that implicit bias can influence teacher expectations for students (Tenenbaum & Ruck, 2007) and interpretations of behavior (Skiba et al., 2002). The review of meta-analyses by Tenenbaum and Ruck (2007) found that teachers had more positive expectations for White youth than their Hispanic and African American peers. During a task in which preschool teachers were given a vignette about a child, teachers kept their gaze on African American boys longer when challenging behaviors were described (Gilliam, Maupin, Reyes, Accavitti, & Shic). Additionally, there is evidence of differential patterns of treatment at the classroom level in which African American students are more likely to receive ODRs for infractions that are more subjective and open to interpretation (Skiba et al., 2011). An analysis of ODR data for over one million students determined that disparities in subjective ODRs (e.g. defiance), as compared to objective ODRs (e.g. truancy), disproportionately affected minority students and explained the vast majority of the variance in total ODR disparities (Girvan, Gion, McIntosh, & Smolkowski, 2017). Ferguson's (2001) ethnographic study of elementary school discipline found that teachers

perceived misbehaviors by African American male students as threatening and dangerous, while teachers perceived the same misbehaviors by White male students as developmentally appropriate. Amongst elementary school teachers, measures of explicit and implicit racial bias had a near zero correlation, which suggests that stated beliefs about equality may not be representative of subconscious influences on behavior, and that merely stating the intention of removing bias from the identification process is likely not enough to meaningfully reduce disparities (Van den Bergh, Denessen, Hornstra, Voeten, & Holland, 2010).

Symptom Presentation

The existence of disparities in BER identification is complex and multifaceted. Several additional factors, including perceptions about the severity of presenting symptoms, cross-cultural and linguistic factors, and other sociodemographic variables warrant exploration. There is evidence that teachers minimize the risk associated with certain types of mental health needs and symptom presentations. In general, teachers tend to perceive internalizing/overcontrolled symptoms of depression, anxiety, and social withdrawal as less serious, less concerning, and less likely to be referred for assessment and treatment than externalizing problems (Chang & Sue, 2003). In a study in which teachers were given vignettes of students with internalizing and externalizing disorders and asked whether a referral was warranted, teachers did not explicitly state a bias towards one symptom presentation (Pearcy, Clopton, & Pope, 1993). However, the same teachers did report that they more frequently referred for externalizing problems in

14

practice; the authors suggest this is due to the realities of behavior management in the classroom, which prioritizes disruptive externalizing problems over withdrawn internalizing problems (Pearcy et al., 1993). These patterns may partially explain why adults are more likely to recognize the need for treatment in children with externalizing disorders than internalizing disorders (Chavira, Stein, Bailey, & Stein, 2004).

Cross-cultural Factors and Language Status

The cultural context of behaviors and interpretations of behaviors also warrant discussion, particularly as it relates to students whose families recently immigrated to the United States and are classified as English Language Learners (ELLs). Students who are non-native English speakers may present different behavioral and social skills than their native English speaking peers (Blatchley & Lau, 2010). Additionally, students experiencing the stresses of acculturation could present symptoms in the classroom that mimic signs of EBDs (Blatchley & Lau, 2010). There is a documented lack of professional development for teachers regarding working specifically with ELLs, and teachers cite systemic challenges such as communication with students and their families as a major obstacle to fully meeting their needs (Hansen-Thomas, Richins, Kakkar, & Okeyo, 2016). Other systemic factors, such as lower school-initiated parental engagement and limited support services to meet language needs, are documented and plausibly related to referrals for mental health services (Niehaus & Adelson, 2014). The researcher is unaware of any academic articles directly exploring the relationship between ELL status and teacher referral for mental health issues, but given the potential for cross-

cultural misinterpretation and miscommunication, there is reason to suspect that this factor may influence traditional identification practices for BER.

Socioeconomic Status

Finally, the role of SES in traditional identification practices requires exploration. As stated previously, there is agreement in the field that low SES increases BER (Lemstra et al., 2008; Reiss, 2013). In the context of school-based identification of BER, SES as a risk factor for BER is complicated by teacher perceptions of students from low-SES backgrounds. There is evidence that teachers hold lower expectations for students from low-SES backgrounds, particularly for males (Auwarter & Aruguete, 2008). Students from low-SES backgrounds are also more likely to have relationships with teachers that are higher in conflict and lower in closeness than their more affluent peers (McGrath & Van Bergen, 2015). Systemic factors related to SES also warrant discussion. Multiple studies suggest that parents of low SES students tend to be less engaged with their child's school, which may be related to a lack of engagement efforts by the school as well as lower parent self-efficacy regarding advocating on behalf of their child in the school setting (Hill & Taylor, 2004). In addition, teachers of low SES students tend to have less experience and training than teachers working in higher income schools, which suggests they have less exposure to the spectrum of normal and abnormal behaviors (Barbarin & Aikens, 2015). Further investigation is needed to understand whether these teacher-level factors and systemic factors influence identification rates above and beyond the increased prevalence rates expected for low-SES youth.

16

Additional Concerns with Traditional Identification Practices

A review of the literature on traditional identification practices revealed several other potential obstacles to the effective identification of student mental health needs in the school setting. First, there is a concern that some teachers do not consider behavior problems and social-emotional development to be a part of their responsibilities, which could lead to under-reporting and lower referral rates in these areas (Severson, Walker, Hope-Doolittle, Kratochwill, & Gresham, 2007). Additionally, studies suggest there is usually a substantial time lapse, often five years or greater, between the recognition of initial behavioral or emotional symptoms by someone outside of the family and school personnel formally recognizing the need for specialized services (Duncan, Forness, & Hartsough, 1995). Furthermore, a survey of 152 high school teachers found that while they perceived that they were expected to identify internalizing symptoms in their students, they felt less capable of recognizing and accurately referring for internalizing disorders when compared to externalizing disorders (Papandrea & Winefield, 2011). Finally, attempts to avoid the stigma of labeling EBDs, the lack of services for identified students, and the emphasis on academics over social-emotional development all have the potential to hinder traditional identification practices in schools (Severson et al., 2007).

UNIVERSAL SCREENING

Over the past two decades, the academic world and schools across the country have invested resources in the development of multi-tiered systems of support (MTSS) models (Eagle, Dowd-Eagle, Snyder, & Holtzman, 2015). The MTSS model provides a comprehensive framework to address the diverse academic, social, emotional, and behavioral needs of students

through systematic supports that vary in intensity based on need (Utley & Obiakor, 2015). The authorization of IDEA in 2004 codified into federal law the national movement towards an emphasis on models of primary prevention in the school setting (C. R. Cook, Volpe, & Livanis, 2010). Frameworks such as Response to Intervention (RTI) and School-wide Positive Behavioral Interventions and Supports (PBIS) have gained traction as data-driven models for serving the needs of an entire school population through primary, secondary, and tertiary supports (Gersten & Dimino, 2006). These models aim to use data to more effectively deliver services to a diverse group of students (Lane et al., 2011). The potential for systematic support structures is especially promising at the elementary level, as children are less likely to be significantly behind and growth trajectories are still being determined (Campbell & Ramey, 1995).

A critical component of the MTSS model is universal screening (Albers et al., 2007). *Universal screening* is the systematic and standardized process of assessing an entire school population for predetermined criteria in specific domains (e.g. social-emotional development, academics), with the goal of identifying "at risk" students in need of intervention (Donohue, Goodman-Scott, & Betters-Bubon, 2015). Universal screening data can be used to identify broad areas of need across a school population, which can lead to adjustments in schoolwide systems such as academic curricula or a mental health intervention model (Dowdy et al., 2015). Universal screening data can also be used to identify individual students "at risk" in one or more domains, who subsequently become the targets of secondary and tertiary interventions (Albers et al., 2007). There is evidence that the majority of students identified via a universal screening procedure may not have previously been identified, and that screening can lead to increases in service utilization (Husky, Sheridan, McGuire, & Olfson, 2011).

18

This shift towards MTSS and universal screening has largely occurred in the domain of academic screening (Cook et al., 2010). In 2014, a nationwide survey of 454 school or district-level administrators representing a range of school levels, locales, and SES levels found that 81% indicated they used some type of academic screening tool (Bruhn et al., 2014). Typical academic screeners are curriculum-based measures that compare a student's performance to established academic benchmarks three to four times a year (Kalberg et al., 2010).

Universal Screening for Emotional and Behavioral Disorders

The same principles that guide school-based academic screening, including its systematic and universal design, its implementation at regularly scheduled intervals, and its function of flagging "at risk" students, apply to universal screening for EBDs (Dowdy et al., 2010). School-based universal screening for EBDs attempts to capture broad indicators associated with mental health functioning, serving as an initial indicator of risk rather than as a diagnostic tool (Levitt, Saka, Hunter Romanelli, & Hoagwood, 2007). If the universal screening procedure is working effectively, it will identify students with elevated behavioral and emotional risk (BER), a term that captures a range of early symptoms of disorders that may later require special education placement or warrant a diagnosis (Kamphaus, 2012). Screening and early identification is a first step towards offering necessary additional services rather than an end goal (Dowdy, Kamphaus, Twyford, & Dever, 2014). Given the critical importance of early intervention for ensuring positive outcomes for children, universal screening has the potential to be an integral part a more effective 21st century mental health system (Weist, Rubin, Moore, Adelsheim, & Wrobel, 2007). Mental Health America, the American Academy of

Pediatrics, and the United States Preventative Services Task Force are just a few of the numerous organizations that have endorsed mental health universal screening for youth in either school or clinical settings as a best practice (Mental Health America, 2013).

Universal screening and early identification of mental health needs are essential components of a MTSS designed to serve all students with the necessary supports to be successful (Albers et al., 2007). Despite the growing popularity of MTSS to address academic needs, relatively few schools have expanded this framework to include monitoring of emotional and behavioral risks (Bruhn et al., 2014). In 2005, only 2% of schools reported using universal screening tools that went beyond academic domains (Evans, 2005). In 2014, the nationwide survey of school administrators found that 12.6% of respondents indicated that their school or district conducted school-wide screening for BER (Bruhn et al., 2014). These numbers suggest that universal screening for emotional and behavioral disorders is slowly gaining traction but is still only utilized in a small minority of schools.

Status of Research on Universal Screening Tools

In order to better understand the current status of academic research on tools available for universal screening for BER in the school setting, a literature review was conducted in May, 2017. The researcher searched the Academic Search Complete, Education Source, ERIC and PsycINFO databases using the following search terms: "universal screen*" OR "systematic screen*" OR "risk screen*" AND school* OR students AND social OR emotion* OR behavior* OR psychosocial OR "mental health" OR "mental disorders" OR "mental illness." The search was limited to peer reviewed articles published since 1997. Out of a total of 1,079 articles, 87

empirical articles and 24 review articles were selected for further consideration because they included key words in the title or abstract related to universal screening, emotional and behavioral disorders, and the school setting.

Several trends emerged from the review. Empirical articles tended to focus on the validation of tools rather than school outcomes associated with universal screening. 51 of the 87 empirical articles focused on establishing psychometric properties for a screening tool. The empirical articles included the use of 31 different screening tools. The most common screening tool was the Behavioral and Emotional Screening System (BESS), which was used by researchers in 24 of the 87 empirical articles. Academic articles on the use of universal screening for EBDs also appear to focus primarily on the elementary school setting. 60 of 87 empirical articles included elementary schools, compared to 24 studies that included middle schools and 21 studies that included high schools. The screening tools most often used in the reviewed articles relied on teacher report. 73 of the 87 articles included a teacher report, and the teacher was the only reporter in all but 8 of those studies. Comparatively, 7 studies incorporated a parent report and 15 studies incorporated a student report.

Validated Tools

A multitude of validated tools for assessing BER through a universal screening process are currently available, and new tools are continuing to be developed (Dever, Raines, & Barclay, 2012). These tools vary in terms of the robustness of the evidence supporting their usage, their format, the required reporters, and their current usage in schools (Feeney-Kettler, Kratochwill, Kaiser, Hemmeter, & Kettler, 2010; Jenkins et al., 2014). Given these variations, school personnel should consider the appropriateness for the intended usage, the psychometric

properties, and the usability of the instrument when selecting a universal screening tool (Glover & Albers, 2007).

Commonly cited tools in the literature with evidence suggesting adequate psychometric properties include, but are not limited to, the Behavioral and Emotional Screening System (BESS) the Social Skills Improvement System Performance Screening Guide (SSIS-PSG), the Strengths and Difficulties Questionnaire (SDQ), the Systematic Screening for Behavior Disorders (SSBD), and the Social, Academic, and Emotional Behavior Risk Screener (SAEBRS) (Jenkins et al., 2014; Miller et al., 2015; von der Embse, Pendergast, Kilgus, & Eklund, 2016). Each of the tools listed above was included in at least three of the academic articles included in the review of tools available for universal screening.

Typical Reporters and Informant Discrepancies

One of the decisions school personnel need to be make when considering options for a universal screening tool is who will be reporting (Dowdy et al., 2010). Some tools, such as the BESS and the SDQ, offer options for multiple reporters, including teacher, parent, and student self-report forms, while others, such as the SSBD, only include a teacher report (Jenkins et al., 2014). Additional decisions may need to be made regarding the logistics of soliciting responses from reporters. For example, within teacher reporters, decisions need to be made about whether ratings will be provided by an instructional teacher (e.g. academic subject teacher), non-instructional teacher (e.g. advisory supervisor), or both (Lane et al., 2011). Within instructional teachers, there is evidence that general education teachers and special education teachers may rate students differently, with special education teachers reporting lower levels of BER (Tanner,

Eklund, Kilgus, & Johnson, 2018). Some tools, such as the SSBD, use a multiple gating procedure, which includes multiple rounds of screening (Dowdy, Dever, Raines, & Moffa, 2016).

Discrepant reporting between different informants is a well-established phenomenon, and should be considered the norm rather than an aberration from the norm (Achenbach, McConaughy, & Howell, 1987). A metal-analyses of 119 studies on informant discrepancies in clinical settings determined that similar reporters (e.g. teacher-teacher) tend to correlate around .6, two outside reporters (e.g. teacher-parent) correlate at roughly .28, and subject-informant reporters (teacher-child) correlate at approximately .22 (Achenbach et al., 1987). A recent analysis of universal screening data for BER using the SAEBRS and SDQ found interrater correlations between teachers to be approximately .7 for both instruments (Tanner et al., 2018).

De Los Reyes and Kazdin (2005) proposed the Attribution Bias Context (ABC) Model as a framework for understanding informant discrepancies in a clinical setting. According to this model, discrepancies exist due to varying attributions regarding the cause of the behavior, perspectives on whether the behaviors warrant treatment, and the informant's goals for the assessment process (De Los Reyes & Kazdin, 2005). If multiple reporters are included in a screening process, these factors related to reporting discrepancies should be considered.

It is important to consider the ways in which cultural differences may partially explain discrepancies between reporters (Lau et al., 2004). A student's cultural background and language abilities should be considered to avoid the misinterpretation of

23

behaviors as maladaptive when they may be culturally appropriate (Dowdy et al., 2014).

An analysis of parent, teacher, and student self-report data found that teacher-student

discrepancies for internalizing behavior problems were higher for African American and

Asian and Pacific Islander students than for White students (Lau et al., 2004). When

conducting a universal screening procedure with a culturally diverse population, schools

should consider factors such as language status, level of acculturation, and the

psychometric properties of the instrument in multicultural settings (Dowdy et al., 2014).

Mischievous Responding

An additional consideration when collecting student self-report data centers

around the honesty of youth responses and mischievous responders. Mischievous

responding describes an individual's pattern of responses that includes extreme,

untruthful answers to multiple responses, often indicating multiple high-risk behaviors in

exceedingly unlikely combinations (Furlong, Fullchange, & Dowdy, 2017). There is

evidence that mischievous responding, in addition to factors such as social desirability,

unengaged responding, and response inconsistencies, can compromise the reliability and

validity of youth survey responses (Cornell, Klein, Konold, & Huang, 2012). An analysis

of over 1,800 high school responses to a universal screener for complete mental health

found that roughly 2% of students could be classified as mischievous responders, which

suggests that schools should be aware of the possibility of mischievous responding, and

that untruthful answers are not universal enough to compromise the functioning of the

system (Furlong et al., 2017). Schools could also consider the setting in which the self-

report measures are completed; a review of student response patterns determined that more responses were flagged for invalidity when administered in a classroom setting with peers versus after class (Spirrison, Gordy, & Henley, 1996). Given that no high stakes decisions are being made based on universal screening data, the possibility of untruthful responding should be a factor for schools to be aware of rather than a critical obstacle for the universal screening process (Furlong et al., 2017).

STRENGTHS AND DIFFICULTIES QUESTIONNAIRE

There are several well-validated tools available for universal screening for BER (Jenkins et al., 2014). The guidelines for instrument selection outlined by Glover and Albers (2007) suggest that the primary considerations should be appropriateness for intended usage, technical adequacy, and usability. After reviewing all well-established instruments available for universal screening for BER in the school setting, the researcher determined that the self-report version of the Strengths and Difficulties Questionnaire (SDQ) is the most appropriate tool for the proposed study. The evidence supporting the SDQ with regards to its appropriateness for intended usage, technical adequacy, and usability is discussed below, as well as a brief comparison with other available tools.

Appropriateness of usage

The SDQ is frequently cited as an appropriate tool for universal screening for BER (Jenkins et al., 2014). Appropriateness includes alignment with the constructs of interest and population fit (Glover & Albers, 2007). The constructs of interest in the case of universal screening for BER are overall mental health functioning as well as broad

measures of internalizing and externalizing problems (Kamphaus, 2012). The SDQ

includes an overall Total Difficulties score, as well as Emotional Symptoms and Conduct

Problems subscale scores that capture these broad concepts of risk. With regards to

population fit, the SDQ youth self-report was designed to be developmentally appropriate

for 11-16 year old students (R. Goodman, Meltzer, & Bailey, 2003). In the context of

universal screener for BER, the SDQ self-report is more appropriate than the teacher or

parent report for several reasons. Self-report measures appear to be the best way to gather

information about internalizing disorders, which are the most likely to be overlooked by

traditional identification practices (Raines, Dever, Kamphaus, & Roach, 2012). In an

analysis of a clinical sample, the SDQ self-report more accurately identified

psychopathology than the parent report, largely because of the more accurate

identification of internalizing symptoms (Kovacs & Sharp, 2014). The youth report also

appears to be more usable, which will be discussed in further detail below.

Technical adequacy

The technical adequacy of the SDQ as a mental health screener in adolescent

populations is well-supported in the literature. Numerous studies have supported the self-

report version of the SDQ as a reliable (Muris, Meesters, Eijkelenboom, & Vincken,

2004) and valid (R. Goodman et al., 2003) measure of mental health in adolescents. The

five factor model outlined by Goodman (1997) is generally well-supported by

confirmatory factor analysis (Hoofs, Jansen, Mohren, Jansen, & Kant, 2015; Richter,

Sagatun, Heyerdahl, Oppedal, & Røysamb, 2011) and there is evidence of adequate

26

discriminant and concurrent validity (R. Goodman et al., 2003; Muris et al., 2004). An

analysis of criterion validity in a clinical sample determined that the SDQ classified

psychopathology similarly to the Youth Self-Report (YSR) and Child Behavior Checklist

(CBCL), two of the gold-standard broadband mental health screeners (Kovacs & Sharp,

2014). The cutoff scores proposed by Goodman (2001) for the Total Problems score led

to high specificity at 94% and low sensitivity at 23%. Other cutoff scores have been

proposed for American samples that increase the sensitivity and decrease the specificity

of the measure (Kovacs & Sharp, 2014). In the context of using the SDQ as a screening

tool, high specificity would allow schools to confidently rule out students who do not

need further screening or intervention. Finally, there is evidence supporting the SDQ's

usage in multicultural populations, as cross-cultural comparisons show more variance

within populations than between populations, and norms from multiple populations can

be used as culturally appropriate references (Achenbach et al., 2008).

Usability

The usability of the self-report SDQ was a primary consideration given the

limited number of schools currently conducting universal screening for BER and the

perceived obstacles (Bruhn et al., 2014). School personnel have raised a variety of

concerns about conducting a universal screening process for BER, including but not

limited to the costs associated with a screening tool, personnel time needed to complete a

screening process, ease of administration and scoring, and how to meet the needs of all

identified students (Humphrey & Wigelsworth, 2016). The SDQ offers several

advantages when compared with other common universal screening tools. First, the SDQ

is a free and publicly available instrument, which contrasts with the costs of other

validated tools (Harrison, Vannest, & Reynolds, 2013). School personnel have the option

to score the assessment by hand for free or with online software for $0.25 per screener.

Additionally, the SDQ is a brief questionnaire that can be completed in roughly five

minutes (Robert Goodman & Scott, 1999), eliminating concerns about lengthy

assessments that take away class time. School-based mental health personnel report that it

can be difficult to get teachers to follow through on completing measures (Connors,

Arora, Curtis, & Stephan, 2015). A student self-report eliminates this barrier as teachers

are not tasked with completing screeners for their entire class. The researcher was unable

to find data on user perceptions of the SDQ youth self-report. However, research on the

parent report suggests that it may be more a more acceptable tool than more intensive

broadband measures such as the CBCL (Robert Goodman & Scott, 1999). A review of

the social acceptability of the SDQ revealed multiple strengths of the SDQ, including its

accessible reading levels and that it does not require a professional with an advanced skill

set to interpret scores (Harrison et al., 2013).

Comparing the SDQ with Other Available Tools

The SDQ self-report compares favorably with several other potential screeners in

the context of this study. First, it is free to administer and there are both free and low-cost

options for scoring. Cost is frequently cited as a barrier to universal screening (Bruhn et

al., 2014), so the choice of a free measure increases the potential usability of the study

findings. Other available measures, such as the BESS, can cost hundreds or thousands of

dollars depending on the number of reporters and school size. Second, the SDQ self-

report is a time-efficient measure that takes approximately five minutes to complete,

which compares favorably with screeners such as the SSBD, which can take up to an

hour per classroom (Harrison et al., 2013). Third, the psychometric properties of the SDQ

are adequate for usage, and its high specificity (.94) when compared to the BESS (.64-

.82) and SRSS (.74-.95) is ideal for a screener that is designed to rule out students

without BER (Harrison et al., 2013). Finally, the study aims to use a student self-report

for the universal screening process. Several of the tools commonly cited in the literature,

such as the SRSS, SSBD, and SAEBRS, only have teacher report forms.

PREVIOUS RESEARCH COMPARING TRADITIONAL IDENTIFICATION WITH UNIVERSAL SCREENING

While there appears to be sufficient evidence to validate multiple emotional and

behavioral universal screening tools in the school setting, relatively little academic work

has been done to compare the results of a universal screening procedure with typical

identification practices. There is significant value in understanding any possible

discrepancies, both as a means of adequately serving the needs of a school population and

as a check on potential biases in the identification process. Several published articles

begin to address this issue.

Eklund et al. (2009) compared students referred through traditional teacher

referral with a universal screening procedure. A student was classified as "at risk"

29

through traditional teacher referral if they received a referral to the school's child study team, a referral for testing for special education eligibility, were currently enrolled in special education, or received non-special education services such as general-education counseling or in-class accommodations. Teachers completed the BESS for each student in the participating grades. 13 of the 24 students identified by the BESS had not previously been identified through traditional teacher referral. Students that were only identified as "at risk" by the BESS scored significantly lower on a measure of school engagement than students only identified by traditional teacher referral. Despite the robust design, the study has limitations. The demographic characteristics of the elementary school, which was 73% Hispanic or Latino, 68% socioeconomically disadvantaged, and 40% English language learners, provided potentially rich data for understanding the relationship between these characteristics and the referral process. However, no between group differences were analyzed for these characteristics. Furthermore, the relatively small sample size of 48 students suggests that power wasn't sufficient to examine these comparisons, and that replication is necessary to further support the findings.

Dowdy, Doane, Eklund, and Dever (2013) compared teacher nomination and screening practices to examine discrepancies in the identification process. They found that a structured rating scale identified more students as "at risk" than a nomination procedure. However, the teacher nomination process was based on a survey given to each teacher in which they were asked to list any students in their class that they believed were

30

at risk behaviorally or emotionally. The survey was essentially another form of universal screening and was not representative of typical identification practices.

Eklund and Dowdy (2014) evaluated the ability of a universal screener to identify students at risk for emotional and behavioral disorders who might not have otherwise been identified through traditional teacher referral. Students identified through the teacher report BESS were compared to students currently receiving services at the school through traditional identification practices. The study spanned 20 elementary schools, which included 867 students and 216 teachers. Of the 160 students identified as "at risk" using the BESS, only 61 had been previously identified through a traditional referral process. The BESS also failed to identify a significant number of students who were previously identified by the school, which suggests that screening may need to be used in conjunction with a traditional referral process. Students identified by the BESS had significantly higher externalizing and internalizing symptoms than students identified by traditional referral.

The Proposed Study

Schools typically rely on traditional identification methods, which tend to be based on teacher discretion, to identify students with elevated BER. Universal screening for BER, which systematically evaluates BER through a standardized process, is infrequently utilized by schools despite the existence of a multitude of validated tools. Disparities in youth mental health needs identification and service provision exist within the school setting that cannot be attributed differences in prevalence rates. A review of the academic literature on BER identification suggests that student characteristics such as race/ethnicity, presenting symptoms, SES, and language status could influence school-based identification and contribute to existing disparities. The connection between school characteristics and mental health identification is less well studied; preliminary investigation suggests that school population SES and school location do not change the likelihood of using a systematic screening process, but less is known on how these school characteristics might influence traditional identification practices (Bruhn et al., 2014).

The researcher is unaware of any academic articles directly addressing racial/ethnic, SES, or language-based disparities in identification rates for behavioral and emotional risk when comparing systematic universal screening with traditional identification practices. The author is also unaware of any academic research that directly compares a student self-report universal screening procedure to traditional identification practices. Finally, the researcher knows of only one article, Eklund & Dowdy (2014), that begins to analyze discrepancies in presenting symptoms of students identified through the

32

two practices, and that study relies on a teacher report. Further investigation into potential

discrepancies between the identification methods could bring attention to shortcomings in

traditional identification practices and provide further support for an evidence-based

screening procedure, particularly in schools with higher racial/ethnic minority and ELL

populations.

The proposed study will compare traditional identification methods for BER with

a student self-report universal screening procedure for BER in a school setting. The study

will include a universal screening procedure and the examination of school records of

referrals and services received, sociodemographic data, and language status. The

racial/ethnic characteristics, presenting symptoms, SES, and language status of students

identified through the two methods will be analyzed for potential discrepancies. The

proposed research questions and hypotheses based on existing literature are outlined

below.

RESEARCH QUESTIONS AND HYPOTHESES

1. To what extent do traditional identification practices and a student-report

   universal screening process agree on students "at risk" and not "at risk" for

   emotional and behavioral disorders?

   Students will be categorized as "Yes, Yes", "Yes, No", "No, Yes", or "No, No"

based on whether they were identified with elevated BER via each of the two

identification methods. We hypothesize that agreement between the two identification

methods will be below .5. We also predict that the self-report universal screener will

identify significantly more students than traditional identification methods. Previous

studies of agreement levels between universal screening tools and traditional

identification methods have reported agreement levels below 50% and that significantly

more students were identified with a universal screening process (Eklund & Dowdy,

2014; Eklund et al., 2009).

2. How do levels of externalizing and internalizing problems, as measured by the

   SDQ self-report, compare for students identified as "at risk" through traditional

   identification, students identified through universal screening, and students not

   identified by either method?

   The average internalizing and externalizing score, as measured by the SDQ, will

be calculated and compared for students "at risk" through traditional identification

methods, for students "at risk" through universal screening, and students not identified as

"at risk" by either method. We predict that the average level of internalizing symptoms

will be significantly higher for students identified as "at risk" through the SDQ universal

screener than the other two groups and that the average level of externalizing symptoms

will be significantly higher for students identified as "at risk" through traditional

identification methods than the other two groups. Teachers are less likely to recognize

and refer students with symptoms of an internalizing disorder through traditional

identification methods (Cunningham & Suldo, 2014; Pearcy et al., 1993), and self-report

measures are more likely to accurately capture internalizing symptoms (Kovacs & Sharp,

2014). Externalizing symptoms, which often disrupt the classroom learning environment,

are more likely to trigger a teacher-initiated referral through traditional identification

methods (Bradshaw, Buckley, & Ialongo, 2008).

3. To what extent do the group of students identified through universal screening

    and the group of students identified through traditional identification methods

    differ based on racial/ethnic characteristics?

    A breakdown by the race/ethnicity of students identified by traditional

identification methods will be compared to a breakdown by race/ethnicity of students

identified by universal screening. We predict that of all students identified through

universal screening, the proportion of students who are Asian or Hispanic will be

significantly higher when compared to the racial/ethnic breakdown of students identified

through traditional identification methods. Similarly, we predict that of all students

identified through universal screening, the proportion of students who are white or

African American will be significantly lower when compared to the racial/ethnic

breakdown of students identified through traditional identification methods. Large

epidemiological studies suggest similar rates of EBDs amongst youth in different

racial/ethnic groups (Merikangas et al., 2010), and investigations of youth self-report

measures have found little variation in reported problems by racial/ethnic group (Lau et

al., 2004). However, there does appear to be a consensus that White students are more

likely to be identified and receive services for internalizing disorders than racial/ethnic

minority students (Gudiño et al., 2009; Merikangas et al., 2011). There is also evidence

that African American students may be over-identified in schools for emotional

disturbance while Asian and Hispanic youth are under-identified (Bear, Finer, Guo, & Lau, 2014; Coutinho & Oswald, 2000).Given these trends, we predict that there will be an overrepresentation of White and African American students through traditional identification practices when compared to universal screening.

4. To what extent do the group of students identified through universal screening and the group of students identified through traditional identification methods differ based on SES?

   A breakdown by the SES of students identified by traditional identification methods will be compared to a breakdown by SES of students identified by universal screening. We hypothesize that students from low SES backgrounds, as measured by the free and reduced lunch program, will be overrepresented through traditional identification methods when compared to universal screening. This prediction is based on evidence that teachers may hold more negative beliefs about low SES students (Auwarter & Aruguete, 2008; McGrath & Van Bergen, 2015), as well as systemic factors related to parent engagement and teacher experience (Barbarin & Aikens, 2015; Hill & Taylor, 2004), which will cause low SES students to be overidentified above and beyond the increase in prevalence expected due to low SES as a risk factor for psychopathology.

5. To what extent do the group of students identified through universal screening and the group of students identified through traditional identification methods differ based on ELL status?

A breakdown by the ELL status of students identified by traditional identification methods will be compared to a breakdown by the ELL status of students identified by universal screening. We hypothesize that students with lower levels of English proficiency will be underrepresented by traditional identification methods when compared to universal screening. This prediction is based on the belief that, in general, teachers' lack of multicultural competency will lead to a cautious approach with ELL students and subsequent under-referral for services, and that universal screening will be a more sensitive method for identifying psychopathology in this population.

METHODS

Participants

A public middle school in the Austin area will be recruited to participate in the study. The school will need to have at least 132 students in 6th through 8th grade based on power analyses discussed below. The ideal school will be racially/ethnically diverse. However, the researcher acknowledges that broad racial/ethnic diversity in schools is the exception rather than the norm. An example of a middle school that reflects the racial/ethnic demographics of Austin Independent School District (AISD) is Fullmore Middle School, which has 997 students and is 66% Hispanic, 22.4% White, 6.2% African American, 2.6% Asian, and 2.4% two or more races. Fullmore Middle School is 28.6% English Language Learners and 64.2% socioeconomically disadvantaged. A school of this size and diversity would allow adequate power to examine all of the outlined research questions. If the researchers are not able to work with a school with this level of

diversity, the number of racial/ethnic groups that are included in analyses may need to be limited based the demographics of the participating school.

All students in 6th, 7th, and 8th grade will participate in the study as a part of the universal screening process. Exclusion criteria include a moderate or more severe intellectual disability, which could compromise the student's ability to self-reflect, as well as limited English proficiency for students whose first language does not have an available version of the SDQ. The SDQ is offered in 78 different languages, so the researchers do not anticipate any participants will be excluded based on this criterion.

Procedure

The researchers will work with the participating school to conduct a universal screening procedure roughly six weeks into the school year. Once the school and researchers have agreed on an appropriate date and time of the school day for screening, the researchers will go to the school to facilitate the screening sessions. Each student will complete the SDQ self-report, which should take approximately 5-10 minutes. At the start of the session, teachers will be asked if any students need the Spanish version of the SDQ. Students who are not able to complete the SDQ independently due to reading difficulties or a disability will have the screener privately read aloud to them by a research assistant. Make-up dates will be scheduled for students who are absent during the screening session.

Due to the universal nature of the screening process, it is not required that families consent to this initial student participation. However, a protocol will be

established to ensure the confidentiality of sensitive student data. Data collected during the screening process will not leave school grounds until all identifying information has been replaced with participant ID numbers. At that point, each student's SDQ responses will be entered into a statistical software package and classified as either "at risk" or "not at risk" based on their overall Total Difficulties score. Each student's scores for the Emotional Symptoms and Conduct Problems subscales will also be compiled.

In the spring of the same academic year, the researchers will work with school staff to collect school records on student characteristics and identification status. The study will adopt similar criteria to those used by Eklund et al. (2009) to determine if a student has been identified with elevated BER by traditional identification methods. In the Eklund et al. (2009) study, children were considered identified if there had been a referral to the child study team, a referral for special education testing, current enrollment in special education, or the receipt of non-special education services. The only modification to the criteria for the proposed is the exclusion of purely academic referrals or services. A student will be considered "identified" by traditional identification methods if there is a behavioral and/or emotional component to any of the following: a referral to the school's child study team, a referral for testing for special education eligibility, current enrollment in special education, or the receipt of non-special education services such as general-education counseling or in-class accommodations. At this time, student racial/ethnic data will also be collected from the school's student data system. Schools typically keep student demographic information organized in a central computerized system, including all of the data relevant for this study. Data on SES will

be collected using free and reduced lunch as a proxy variable. Currently in Texas, a family of four is eligible for reduced meals if the combined income is under $44,955 and they are eligible for free meals if their income is under $31,590. The final student characteristic that will be collected is language status. Students in Texas are classified as ELLs according to the Texas English Language Proficiency Assessment System (TELPAS). The test assesses English language proficiency of K-12 ELLs in four domains: listening, speaking, reading, and writing. A single composite proficiency rating is created from these four domains. Every student will be classified by their proficiency rating: beginning, intermediate, advanced, or advanced high. All students who have either placed out of the ELL designation or are native English speakers will be classified as non-ELL.

By the end of the academic school year, the researchers will have the BER status based on the universal screening process, the identification status based on traditional identification methods, self-reported internalizing and externalizing symptoms, and the racial/ethnic background, SES, and language status of each student in the school. Students who permanently leave the school after the universal screening process will be removed from the data set. Data from D.C. and Chicago public schools suggest that up to 7-8% of students may transfer into a school after September (Whitesell, Stiefel, & Schwartz, 2016). To limit the potential impact of late entries and transfers, the first round of screening will occur roughly 6 weeks after the school year starts. At the end of the year, the sociodemographic characteristics of students who left mid-year will be

examined for representativeness of the school population. It will not be feasible to follow students who transfer out mid-year, which is a limitation that will be discussed below.

Measures

The Strengths and Difficulties Questionnaire (SDQ) will be the primary measure in this study. The SDQ was originally developed as an expansion of the Rutter parent questionnaire by Goodman (1997). The SDQ is a brief screener of behavioral and emotional functioning that is appropriate for youth aged 3-16. It was developed based on the *Diagnostic and Statistical Manual of Mental Disorders (4th edition)*, practical considerations, and factor analysis (Goodman, 2001). The 25 items cover different attributes, some positive and some negative, that are classified into one of five subscales: emotional symptoms, conduct problems, hyperactivity-inattention, peer problems, and prosocial behavior (Goodman, 2001). Respondents use a three-point Likert scale to report whether items are "not true", "somewhat true", or "certainly true". Examples of items on the self-report version include "I get very angry and often lose my temper" and "I fight a lot. I can make other people do what I want". A total difficulties score is calculated based on the sum of the first four subscales and ranges from 0-40; scores are classified as "normal", "borderline", or "abnormal" based on cutoff scores (Dever et al., 2012). The SDQ includes parent, teacher, and youth self-report forms. This study will use the SDQ youth self-report, which is appropriate for youth aged 11-17.

The psychometric properties of the SDQ have been extensively studied since its creation in 1997, and substantial evidence exists that suggests the SDQ is a reliable and valid measure.

The internal consistency of the self-report measure is r = .80, with subscale coefficients ranging from .41 to .81 (Goodman, 2001). Test-retest reliability at four to six months is .62 (Goodman, 2001). Confirmatory Factor Analysis in multiple studies have confirmed that the five factor model originally outline by Goodman (1997) is a good fit for the SDQ self-report (Hoofs et al., 2015; Richter et al., 2011; van de Looij-Jansen, Goedhart, de Wilde, & Treffers, 2011). There is evidence of strong concurrent validity with corresponding scales of the Youth Self-Report (YSR) version of the ASEBA (Muris et al., 2004). In a clinical sample that concurrently collected data with the Child Behavior Checklist (CBCL) and YSR forms of the ASEBA; the SDQ Total Difficulties score had a correlation of .83 with the YSR Total Problems and .71 with the CBCL Total Problems (Kovacs & Sharp, 2014). The self-report version of the SDQ has been shown to satisfactorily discriminate between clinical and non-clinical samples (Goodman et al., 2003). A study of 7,912 students found significant predictive validity of the SDQ youth self-report form on child psychopathology over a three-year time period (Goodman & Goodman, 2009). The sensitivity of the self-report form is .23 and the specificity is .94 based on the original cutoff scores proposed by Goodman (Jenkins et al., 2014). The Spanish version of the SDQ self-report also appears to have adequate internal consistency (.75) and supports the same five-factor structure as the English version (Ortuño-Sierra, Fonseca-Pedrero, Paino, Sastre i Riba, & Muñiz, 2015).

There appear to be several limitations to the SDQ youth self-report that should be noted. The reliability of the Peer Problems scale appears to be questionable (Robert Goodman, 2001; Kovacs & Sharp, 2014). The sensitivity of .23 for the youth self-report version is lower than would be desirable. However, given the context of its usage as a screening tool, the low sensitivity does not mean that it is not a useful tool. Finally, despite being frequently cited as a universal screening tool, the researcher found a limited number of studies using the SDQ as a school-based screening tool and is not aware of any published studies in the United States that examine the self-report version at the middle school level.

Statistical Analysis

Hypotheses

1. We hypothesize that agreement between the two identification methods will be below .5. We also predict that the self-report universal screener will identify significantly more students than traditional identification methods.

2. We predict that the average level of internalizing symptoms will be significantly higher for students identified as "at risk" through the SDQ universal screener than the other two groups and that the average level of externalizing symptoms will be significantly higher for students identified as "at risk" through traditional identification methods than the other two groups. However, given that youth with elevated symptoms in one category are more likely to be higher in the other

43

category, it is possible that the effects of any potential discrepancies in identification method will be canceled out.

3. We predict that of all students identified through universal screening, the proportion of students who are Asian or Hispanic will be significantly higher when compared to the racial/ethnic breakdown of students identified through traditional identification methods. Similarly, we predict that of all students identified through universal screening, the proportion of students who are white or African American will be significantly lower when compared to the racial/ethnic breakdown of students identified through traditional identification methods.

4. We hypothesize that students from low-SES backgrounds, as measured by the free- and reduced lunch program, will be overrepresented through traditional identification methods when compared to universal screening.

5. We hypothesize that students with lower levels of English proficiency will be underrepresented by traditional identification methods when compared to universal screening.

Preliminary Analysis

Question 1, which addresses the agreement between the two identification methods, will be analyzed using a Cohen's kappa coefficient. This coefficient is a measure of inter-rater agreement for categorical data, which in this context is "at risk" or "not at risk". The kappa coefficient provides the level of agreement greater than that expected by chance. Question 2, which compares internalizing and externalizing

symptoms amongst three groups, will be analyzed with a one-way ANOVA. A one-way ANOVA compares the means of two or more independent samples. Assumptions of normality, independence of observations, and homogeneity of variance will be analyzed. All comparisons between traditional identification methods and universal screening (questions 3, 4, and 5), will be conducted using a chi-squared goodness of fit test. The chi-squared goodness of fit test compares observed frequency distributions with a theoretical distribution. For questions 3, 4, and 5, the frequency expected will be based on the results of the SDQ universal screener, which will be considered the "true" indicator of psychopathology in the sample. The frequency observed will be the students identified via traditional identification methods. A significant result, which would occur if the calculated chi-squared test statistic exceeds the critical value, would suggest that the distribution of observed frequencies significantly differs from the theoretical distribution expected.

Post Hoc

Significant results in a chi-squared goodness of fit test will be followed up with analysis of standardized residuals as outlined in Sharpe (2015) to determine which differences between expected and observed frequencies contributed the most to the significant result. Confidence intervals will be built around standardized residuals to account for differing group sizes in the comparison. For a significant result in the one-way ANOVA, post hoc analysis using a Bonferroni adjustment will be used. A Bonferroni correction keeps the family-wise alpha at .05 and reduces the risk of a type I

45

error. The kappa coefficient will be interpreted according to guidelines proposed by Landis and Koch (1977): less than or equal to 0 = poor, .01-.20 = slight, .21-.40 = fair, .41-.60 = moderate, .61-80 = substantial, and .81-1 = almost perfect.

Power Analysis

A power analysis was conducted using G*Power software to determine the number of participants needed to find significant results. A power analysis for finding a significant chi-squared goodness of fit test requires 122 participants to obtain a moderate effective size (w = .3) at a .80 power level with an alpha of .05 and four groups (df = 3). A power analysis for finding a significant one-way ANOVA requires 159 participants to obtain a moderate effect size (f = .25) at a .80 power level with an alpha of .05 and three groups. Cohen's kappa does not have a minimum sample size because an inferential test is not being conducted. Given the estimate that up to 8% of students could transfer mid-year based on previous data, the school population will need to have at least 172 students.

Discussion

SUMMARY

      The proposed study seeks to compare traditional identification methods for identifying youth with elevated behavioral and emotional risk (BER) with a self-report universal screening procedure in a school setting. The study will explore discrepancies between the two identification methods, including the degree to which they agree/disagree on "at risk" students, the racial/ethnic characteristics of identified students, and the presenting symptoms of students identified via the two methods. It is expected that the two identification methods will frequently disagree on the risk status of individual students and that patterns based on racial/ethnic background and symptom presentation will emerge. If these hypotheses are supported, the present study could provide additional evidence supporting the use of a systematic universal screening procedure to identify BER in the school setting in order to reduce identification disparities.

LIMITATIONS

      The proposed study has several limitations that should be noted. First, the universal screening procedure only relies on a student self-report. There is evidence that the sensitivity and specificity of the SDQ are higher when multiple reporters are included (Goodman, Ford, Simmons, Gatward, & Meltzer, 2000). However, considerations of usability suggest that including only a self-report may be more feasible for schools, which makes the proposed study more realistic for real-world application. Second, the

47

administration of a second round of universal screening in the spring of the academic

year would allow for investigation of the stability of screening data over time and the

added value of a second screening timepoint. Third, there are potential areas of analysis

that are beyond the scope of the proposed study. Analysis of agreement between the two

identification methods for individual students (i.e. for what types of students were the

methods most likely to disagree) could provide additional valuable information. Other

student variables, such as age and academic performance, could provide additional levels

of understanding of the differences between the two identification methods. Additionally,

the use of only four major racial/ethnic categories, and the lack of a multiracial category,

may not fully capture the nuances of race and ethnicity. Given the lack of racial/ethnic

diversity in many public schools, it will be challenging to find a school with adequate

representation for each group, and one or more of the racial/ethnic groups may have to be

dropped from analysis. Another potential limitation is the inability to follow students who

leave during the academic year. While not anticipated, it is possible that a number of

students beyond those incorporated into the power analysis could leave the school, and

that students who leave are not representative of the entire school. Analyses will be

conducted on transfer students to examine their representativeness. Finally, it should be

noted that the study is only addressing one part of the process of providing appropriate

mental health services based on need. There is evidence that screening-triggered referrals

are less likely to result in caregiver consent than teacher-initiated referrals (Guo, Kim,

Bear, & Lau, 2017). Schools will have to go beyond problem recognition and engage

families in the referral process for the results of a universal screening procedure to lead to better outcomes for students.

STRENGTHS

The proposed study has multiple strengths that make it a potentially valuable contribution to the field. First, it uses a tool, the SDQ self-report version, that is both well-validated and understudied in the context of universal screening. Despite numerous academic articles validating its usage, the researchers are unaware of any academic studies using the self-report version of the SDQ as a universal screening tool in schools. Second, the proposed study will add to the literature on student self-report universal screening, which is relatively understudied compared to teacher-report universal screening. There are reasons to believe that a self-report measure may be a more valid and time-efficient approach to universal screening, and this study will meaningful contribute to that discussion. Third, the proposed study asks a basic but mostly unanswered question: is the new direction in the field of identifying student BER (universal screening) more effective than the status quo (traditional identification methods)? The researchers believe that taking time to ask these fundamental questions will help inform future directions of research within the field. Finally, a strength of the proposed study is its usability in the real world. The measure used in the study is both cost- and time-efficient. If the study does show that a student self-report universal screener adds value to traditional identification practices, schools could implement a similar process with relatively few resources and training.

IMPLICATIONS AND FUTURE RESEARCH

The results of this study could have implications for real-world practice and future research. Every school is tasked to at least some extent with identifying BER. Therefore, the results of the study have implications for all schools. The results of this study will meaningfully comment on the degree to which the emerging best practice of universal screening is effective at reducing disparities based on student racial/ethnic background and symptom presentation. If evidence emerges that universal screening does significantly reduces disparities, there is an additional compelling reason for school administrators to move towards systematic universal screening. For school administrators that cite concerns about inadequate resources to provide mental health services, the results of this study could help the school more effectively advocate for necessary resources. Any changes in school policies on BER identification can and will have an impact on the student body. Although identification does not guarantee appropriate treatment, it is a necessary first step.

If the results of the proposed study are compelling, future research should seek to replicate and expand on this study's findings. Future studies could explore similar research questions at the high school level, in racially/ethnically homogenous schools, and with the use of other validated self-report measures. Promising results from this study could encourage an increased focus on self-report measures to complement the current focus on teacher-report measures in the academic literature. Finally, future research could explore the integration of self-report universal screening with traditional identification methods to maximize the likelihood that student mental health needs are identified.

Ideally, the strengths of each system could be utilized to create a comprehensive system that is more effective than either approach on its own.

References

Achenbach, T. M., Becker, A., Döpfner, M., Heiervang, E., Roessner, V., Steinhausen, H.-C., & Rothenberger, A. (2008). Multicultural assessment of child and adolescent psychopathology with ASEBA and SDQ instruments: research findings, applications, and future directions. *Journal of Child Psychology and Psychiatry*, *49*(3), 251–275. https://doi.org/10.1111/j.1469-7610.2007.01867.x

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*(2), 213–232. https://doi.org/10.1037//0033-2909.101.2.213

Albers, C. A., Glover, T. A., & Kratochwill, T. R. (2007). Where are we, and where do we go now? *Journal of School Psychology*, *45*(2), 257–263. https://doi.org/10.1016/j.jsp.2006.12.003

Alegría, M., Lin, J. Y., Green, J. G., Sampson, N. A., Gruber, M. J., & Kessler, R. C. (2012). Role of referrals in mental health service disparities for racial and ethnic minority youth. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(7), 703–711.

Anderson, D. L., Lubig, J., & Smith, M. (2012). Meeting the Needs of All Students: How Student Teachers Identify Individualization. *Education and Research Perspectives*, *39*, 1–23.

Atkins, M. S., Cappella, E., Shernoff, E. S., Mehta, T. G., & Gustafson, E. L. (2017). Schooling and Children's Mental Health: Realigning Resources to Reduce Disparities and Advance Public Health. *Annual Review of Clinical Psychology*, *13*(1), 123–147. https://doi.org/10.1146/annurev-clinpsy-032816-045234

Atkins, M. S., Hoagwood, K. E., Kutash, K., & Seidman, E. (2010). Toward the Integration of Education and Mental Health in Schools. *Administration and Policy in Mental Health and Mental Health Services Research*, *37*(1–2), 40–47. https://doi.org/10.1007/s10488-010-0299-7

Auwarter, A. E., & Aruguete, M. S. (2008). Effects of Student Gender and Socioeconomic Status on Teacher Perceptions. *The Journal of Educational Research*, *101*(4), 242–246. https://doi.org/10.3200/JOER.101.4.243-246

Avenevoli, S., Swendsen, J., He, J.-P., Burstein, M., & Merikangas, K. R. (2015). Major Depression in the National Comorbidity Survey–Adolescent Supplement: Prevalence, Correlates, and Treatment. *Journal of the American Academy of Child & Adolescent Psychiatry*, *54*(1), 37-44.e2. https://doi.org/10.1016/j.jaac.2014.10.010

Barbarin, O. A., & Aikens, N. (2015). Overcoming the educational disadvantages of poor children: How much do teacher preparation, workload, and expectations matter. *American Journal of Orthopsychiatry*, *85*(2), 101–105. https://doi.org/10.1037/ort0000060

Bear, L., Finer, R., Guo, S., & Lau, A. S. (2014). Building the gateway to success: An appraisal of progress in reaching underserved families and reducing racial disparities in school-based mental health. *Psychological Services*, *11*(4), 388–397. https://doi.org/10.1037/a0037969

Blatchley, L. A., & Lau, M. Y. (2010). Culturally Competent Assessment of English Language Learners for Special Education Services, *38*(7), 8.

Bradshaw, C. P., Buckley, J. A., & Ialongo, N. S. (2008). School-based service utilization among urban children with early onset educational and mental health problems: The squeaky wheel phenomenon. *School Psychology Quarterly*, *23*(2), 169–186. https://doi.org/10.1037/1045-3830.23.2.169

Bradshaw, C. P., Schaeffer, C. M., Petras, H., & Ialongo, N. (2010). Predicting Negative Life Outcomes from Early Aggressive–Disruptive Behavior Trajectories: Gender Differences in Maladaptation Across Life Domains. *Journal of Youth and Adolescence*, *39*(8), 953–966. https://doi.org/10.1007/s10964-009-9442-8

Brauner, C. B., & Stephens, C. B. (2006). Estimating the Prevalence of Early Childhood Serious Emotional/Behavioral Disorders: Challenges and Recommendations. *Public Health Reports*, *121*(3), 303–310.

Breslau, J., Kendler, K. S., Su, M., Gaxiola-Aguilar, S., & Kessler, R. C. (2005). Lifetime risk and persistence of psychiatric disorders across ethnic groups in the United States. *Psychological Medicine*, *35*(3), 317–327. https://doi.org/10.1017/S0033291704003514

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.

Campbell, F. A., & Ramey, C. T. (1995). Cognitive and School Outcomes for High-Risk African-American Students at Middle Adolescence: Positive Effects of Early Intervention. *American Educational Research Journal*, *32*(4), 743–772. https://doi.org/10.2307/1163334

Carter, P. L., Skiba, R., Arredondo, M. I., & Pollock, M. (2017). You can't fix what you don't look at: Acknowledging race in addressing racial discipline disparities. *Urban Education*, *52*(2), 207–235.

Center for Disease Control. (1999). Youth Risk Behavior Surveillance --- United States, 1999. Retrieved April 21, 2018, from https://www.cdc.gov/mmwr/preview/mmwrhtml/ss4905a1.htm

Chang, D. F., & Sue, S. (2003). The effects of race and problem type on teachers' assessments of student behavior. *Journal of Consulting and Clinical Psychology*, *71*(2), 235–242. https://doi.org/10.1037/0022-006X.71.2.235

Chavira, D. A., Stein, M. B., Bailey, K., & Stein, M. T. (2004). Child anxiety in primary care: Prevalent but untreated. *Depression and Anxiety*, *20*(4), 155–164. https://doi.org/10.1002/da.20039

Committee on School Health. (2004). School-Based Mental Health Services. *Pediatrics*, *113*(6), 1839–1845. https://doi.org/10.1542/peds.113.6.1839

Connors, E. H., Arora, P., Curtis, L., & Stephan, S. H. (2015). Evidence-Based Assessment in School Mental Health. *Cognitive and Behavioral Practice*, *22*(1), 60–73. https://doi.org/10.1016/j.cbpra.2014.03.008

Cook, B. L., Barry, C. L., & Busch, S. H. (2013). Racial/Ethnic Disparity Trends in Children's Mental Health Care Access and Expenditures from 2002 to 2007. *Health Services Research*, *48*(1), 129–149. https://doi.org/10.1111/j.1475-6773.2012.01439.x

Cook, C. R., Volpe, R. J., & Livanis, A. (2010). Constructing a Roadmap for Future Universal Screening Research Beyond Academics. *Assessment for Effective Intervention*, *35*(4), 197–205. https://doi.org/10.1177/1534508410379842

Cornell, D., Klein, J., Konold, T., & Huang, F. (2012). Effects of validity screening items on adolescent survey data. *Psychological Assessment*, *24*(1), 21–35. https://doi.org/10.1037/a0024824

Costello, E. J., Egger, H. L., & Angold, A. (2005). The Developmental Epidemiology of Anxiety Disorders: Phenomenology, Prevalence, and Comorbidity. *Child and Adolescent Psychiatric Clinics of North America*, *14*(4), 631–648. https://doi.org/10.1016/j.chc.2005.06.003

Costello, E. J., He, J., Sampson, N. A., Kessler, R. C., & Merikangas, K. R. (2014). Services for adolescents with psychiatric disorders: 12-month data from the National Comorbidity Survey–Adolescent. *Psychiatric Services*, *65*(3), 359–366.

Couchenour, D., & Chrisman, J. K. (2016). *The SAGE Encyclopedia of Contemporary Early Childhood Education* (Vols. 1–3). Thousand Oaks, California. https://doi.org/10.4135/9781483340333

Coutinho, M. J., & Oswald, D. P. (2000). Disproportionate Representation in Special Education: A Synthesis and Recommendations. *Journal of Child and Family Studies*, *9*(2), 135–156. https://doi.org/10.1023/A:1009462820157

Cummings, J. R., Ponce, N. A., & Mays, V. M. (2010). Comparing Racial/Ethnic Differences in Mental Health Service Use Among High-Need Subpopulations Across Clinical and School-Based Settings. *Journal of Adolescent Health*, *46*(6), 603–606. https://doi.org/10.1016/j.jadohealth.2009.11.221

Cunningham, J. M., & Suldo, S. M. (2014). Accuracy of Teachers in Identifying Elementary School Students Who Report At-Risk Levels of Anxiety and Depression. *School Mental Health*, *6*(4), 237–250. https://doi.org/10.1007/s12310-014-9125-9

De Los Reyes, A., & Kazdin, A. E. (2005). Informant Discrepancies in the Assessment of Childhood Psychopathology: A Critical Review, Theoretical Framework, and Recommendations for Further Study. *Psychological Bulletin*, *131*(4), 483–509. https://doi.org/10.1037/0033-2909.131.4.483

Dehlendorf, C., Bryant, A. S., Huddleston, H. G., Jacoby, V. L., & Fujimoto, V. Y. (2010). Health disparities: definitions and measurements. *American Journal of Obstetrics and Gynecology*, *202*(3), 212–213. https://doi.org/10.1016/j.ajog.2009.12.003

Dever, B. V. ., bdever@gsu. ed., Raines, T. C. ., & Barclay, C. M. . (2012). Chasing the Unicorn: Practical Implementation of Universal Screening for Behavioral and Emotional Risk. *School Psychology Forum*, *6*(4), 108–118.

Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. L. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology*, *48*(6), 1267–1278. https://doi.org/10.1016/j.jesp.2012.06.003

Donohue, P., Goodman-Scott, E., & Betters-Bubon, J. (2015). Using Universal Screening for Early Identification of Students at Risk: A Case Example from the Field. *Professional School Counseling*, *19*(1), 133–143. https://doi.org/10.5330/1096-2409-19.1.133

Dowdy, E., Dever, B. V., Raines, T. C., & Moffa, K. (2016). A Preliminary Investigation Into the Added Value of Multiple Gates and Informants in Universal Screening for Behavioral and Emotional Risk. *Journal of Applied School Psychology*, *32*(2), 178–198. https://doi.org/10.1080/15377903.2016.1165327

Dowdy, E., Doane, K., Eklund, K., & Dever, B. V. (2013). A comparison of teacher nomination and screening to identify behavioral and emotional risk within a sample of underrepresented students. *Journal of Emotional and Behavioral Disorders*, *21*(2), 127–137.

Dowdy, E., Furlong, M., Raines, T. C., Bovery, B., Kauffman, B., Kamphaus, R. W., … Murdock, J. (2015). Enhancing School-Based Mental Health Services with a Preventive and Promotive Approach to Universal Screening for Complete Mental Health. *Journal of Educational & Psychological Consultation*, *25*(2–3), 178–197.

Dowdy, E., Kamphaus, R. W., Twyford, J. M., & Dever, B. V. (2014). Culturally Competent Behavioral and Emotional Screening. In M. D. Weist, N. A. Lever, C. P. Bradshaw, & J. S. Owens (Eds.), *Handbook of School Mental Health* (pp. 311–321). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4614-7624-5_23

Dowdy, E., Ritchey, K., & Kamphaus, R. W. (2010). School-Based Screening: A Population-Based Approach to Inform and Monitor Children's Mental Health Needs. *School Mental Health*, *2*(4), 166–176. https://doi.org/10.1007/s12310-010-9036-3

Eagle, J. W., Dowd-Eagle, S. E., Snyder, A., & Holtzman, E. G. (2015). Implementing a Multi-Tiered System of Support (MTSS): Collaboration Between School Psychologists and Administrators to Promote Systems-Level Change. *Journal of Educational and Psychological Consultation*, *25*(2–3), 160–177. https://doi.org/10.1080/10474412.2014.929960

Eklund, K., & Dowdy, E. (2014). Screening for Behavioral and Emotional Risk Versus Traditional School Identification Methods. *School Mental Health*, *6*(1), 40–49. https://doi.org/10.1007/s12310-013-9109-1

Eklund, K., Renshaw, T. L., Dowdy, E., Jimerson, S. R., Hart, S. R., Jones, C. N., & Earhart, J. (2009). Early identification of behavioral and emotional problems in youth: Universal screening versus teacher-referral identification. *California School Psychologist*, *14*, 89–95. https://doi.org/10.1007/BF03340954

Evans, D. L., Annenberg Public Policy Center, & Annenberg Foundation Trust at Sunnylands (Eds.). (2005). *Treating and preventing adolescent mental health disorders: what we know and what we don't know: a research agendafor improving the mental health of our youth*. Oxford ; New York: Oxford University Press.

Feeney-Kettler, K. A., Kratochwill, T. R., Kaiser, A. P., Hemmeter, M. L., & Kettler, R. J. (2010). Screening Young Children's Risk for Mental Health Problems: A Review of Four Measures. *Assessment for Effective Intervention*, *35*(4), 218–230. https://doi.org/10.1177/1534508410380557

Ferguson, A. (2001). Bad Boys: Public Schools in the Making of Black Masculinity by Ann Arnett Ferguson. *American Journal of Sociology*, *106*(5), 1440–1441. https://doi.org/10.1086/320824

Flores, G., Fuentes-Afflick, E., Barbot, O., Carter-Pokras, O., Claudio, L., Lara, M., … Weitzman, M. (2002). The Health of Latino Children: Urgent Priorities, Unanswered Questions, and a Research Agenda. *JAMA*, *288*(1), 82–90. https://doi.org/10.1001/jama.288.1.82

Furlong, M. J., Fullchange, A., & Dowdy, E. (2017). Effects of mischievous responding on universal mental health screening: I love rum raisin ice cream, really I do! *School Psychology Quarterly*, *32*(3), 320–335. https://doi.org/10.1037/spq0000168

Gersten, R., & Dimino, J. A. (2006). RTI (Response to Intervention): Rethinking Special Education for Students with Reading Difficulties (Yet Again). *Reading Research Quarterly*, *41*(1), 99–108.

Ghandour, R. M., Kogan, M. D., Blumberg, S. J., Jones, J. R., & Perrin, J. M. (2012). Mental Health Conditions Among School-Aged Children: Geographic and Sociodemographic Patterns in Prevalence and Treatment. *Journal of Developmental & Behavioral Pediatrics*, *33*(1), 42–54. https://doi.org/10.1097/DBP.0b013e31823e18fd

Gilliam, W. S., Maupin, A. N., Reyes, C. R., Accavitti, M., & Shic, F. (n.d.). Do Early Educators' Implicit Biases Regarding Sex and Race Relate to Behavior Expectations and Recommendations of Preschool Expulsions and Suspensions?, 18.

Girvan, E. J., Gion, C., McIntosh, K., & Smolkowski, K. (2017). The relative contribution of subjective office referrals to racial disproportionality in school discipline. *School Psychology Quarterly: The Official Journal of the Division of School Psychology, American Psychological Association*, *32*(3), 392–404. https://doi.org/10.1037/spq0000178

Glover, T. A., & Albers, C. A. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology*, *45*(2), 117–135. https://doi.org/10.1016/j.jsp.2006.05.005

Goodman, A., & Goodman, R. (2009). Strengths and Difficulties Questionnaire as a Dimensional Measure of Child Mental Health. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(4), 400–403. https://doi.org/10.1097/CHI.0b013e3181985068

Goodman, R., Meltzer, H., & Bailey, V. (2003). The Strengths and Difficulties Questionnaire: a pilot study on the validity of the self-report version. *International Review of Psychiatry*, *15*(1–2), 173–177. https://doi.org/10.1080/0954026021000046137

Goodman, Robert. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

Goodman, Robert. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(11), 1337–1345.

Goodman, Robert, Ford, T., Simmons, H., Gatward, R., & Meltzer, H. (2000). Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *The British Journal of Psychiatry*, *177*(6), 534–539. https://doi.org/10.1192/bjp.177.6.534

Goodman, Robert, & Scott, S. (1999). Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: Is Small Beautiful? *Journal of Abnormal Child Psychology*, *27*(1), 17–24. https://doi.org/10.1023/A:1022658222914

Griffith, A. K. (2010). Committee on the Prevention of Mental Disorders and Substance Abuse Among Children, Youth, and Young Adults, National Research Council and the Institute of Medicine: <Emphasis Type="Italic">Preventing Mental, Emotional, and Behavioral Disorders Among Young People: Progress and Possibilities</Emphasis>. *Journal of Child and Family Studies*, *19*(5), 675–676. https://doi.org/10.1007/s10826-009-9341-3

Gudiño, O. G., Lau, A. S., Yeh, M., McCabe, K. M., & Hough, R. L. (2009). Understanding Racial/Ethnic Disparities in Youth Mental Health Services: Do Disparities Vary by Problem Type? *Journal of Emotional and Behavioral Disorders*, *17*(1), 3–16. https://doi.org/10.1177/1063426608317710

Guo, S., Kataoka, S. H., Bear, L., & Lau, A. S. (2014). Differences in School-Based Referrals for Mental Health Care: Understanding Racial/Ethnic Disparities Between Asian American and Latino Youth. *School Mental Health*, *6*(1), 27–39. https://doi.org/10.1007/s12310-013-9108-2

Guo, S., Kim, J. J., Bear, L., & Lau, A. S. (2017). Does Depression Screening in Schools Reduce Adolescent Racial/Ethnic Disparities in Accessing Treatment? *Journal of Clinical Child & Adolescent Psychology*, *46*(4), 523–536. https://doi.org/10.1080/15374416.2016.1270826

Halfon, N., & Newacheck, P. W. (1999). Prevalence and Impact of Parent-Reported Disabling Mental Health Conditions Among U.S. Children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *38*(5), 600–609. https://doi.org/10.1097/00004583-199905000-00023

Hansen-Thomas, H., Richins, L. G., Kakkar, K., & Okeyo, C. (2016). I do not feel I am properly trained to help them! Rural teachers' perceptions of challenges and needs with English-language learners. *Professional Development in Education*, *42*(2), 308–324. https://doi.org/10.1080/19415257.2014.973528

Harrison, J. R., Vannest, K. J., & Reynolds, C. R. (2013). Social Acceptability of Five Screening Instruments for Social, Emotional, and Behavioral Challenges. *Behavioral Disorders*, *38*(3), 171–189.

Hill, N. E., & Taylor, L. C. (2004). Parental School Involvement and Children's Academic Achievement: Pragmatics and Issues. *Current Directions in Psychological Science*, *13*(4), 161–164. https://doi.org/10.1111/j.0963-7214.2004.00298.x

Hoofs, H., Jansen, N. W. H., Mohren, D. C. L., Jansen, M. W. J., & Kant, I. J. (2015). The Context Dependency of the Self-Report Version of the Strength and Difficulties Questionnaire (SDQ): A Cross-Sectional Study between Two Administration Settings. *PLoS One; San Francisco*, *10*(4), e0120930. http://dx.doi.org.ezproxy.lib.utexas.edu/10.1371/journal.pone.0120930

Humphrey, N., & Wigelsworth, M. (2016). Making the case for universal school-based mental health screening. *Emotional and Behavioural Difficulties*, *21*(1), 22–42. https://doi.org/10.1080/13632752.2015.1120051

Husky, M. M., Sheridan, M., McGuire, L., & Olfson, M. (2011). Mental Health Screening and Follow-up Care in Public High Schools. *Journal of the American Academy of Child & Adolescent Psychiatry*, *50*(9), 881–891. https://doi.org/10.1016/j.jaac.2011.05.013

Jenkins, L. N., Demaray, M. K., Wren, N. S., Secord, S. M., Lyell, K. M., Magers, A. M., … Tennant, J. (2014). A Critical Review of Five Commonly Used Social-Emotional and Behavioral Screeners for Elementary or Secondary Schools. *Contemporary School Psychology*, *18*(4), 241–254. https://doi.org/10.1007/s40688-014-0026-6

Kalberg, J. R., Lane, K. L., & Menzies, H. M. (2010). Using systematic screening procedures to identify students who are nonresponsive to primary prevention efforts: Integrating academic and behavioral measures. *Education and Treatment of Children*, *33*(4), 561–584.

Kamphaus, R. W. ., rkamphaus@gsu. ed. (2012). Screening for Behavioral and Emotional Risk: Constructs and Practicalities. *School Psychology Forum*, *6*(4), 89–97.

Kataoka, S. H., Zhang, L., & Wells, K. B. (2002). Unmet need for mental health care among US children: Variation by ethnicity and insurance status. *American Journal of Psychiatry*, *159*(9), 1548–1555.

Kazdin, A. E., Holland, L., & Crowley, M. (1997). Family experience of barriers to treatment and premature termination from child therapy. *Journal of Consulting and Clinical Psychology*, *65*(3), 453–463. https://doi.org/10.1037/0022-006X.65.3.453

Keenan-Miller, D., Hammen, C. L., & Brennan, P. A. (2007). Health Outcomes Related to Early Adolescent Depression. *Journal of Adolescent Health*, *41*(3), 256–262. https://doi.org/10.1016/j.jadohealth.2007.03.015

Kelly, D. (2013). Implicit Bias and Social Cognition. In B. Kaldis, *Encyclopedia of Philosophy and the Social Sciences*. 2455 Teller Road, Thousand Oaks California 91320 United States: SAGE Publications, Inc. https://doi.org/10.4135/9781452276052.n172

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime Prevalence and Age-of-Onset Distributions of DSM-IV Disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *62*(6), 593–602. https://doi.org/10.1001/archpsyc.62.6.593

Kovacs, S., & Sharp, C. (2014). Criterion validity of the Strengths and Difficulties Questionnaire (SDQ) with inpatient adolescents. *Psychiatry Research*, *219*(3), 651–657. https://doi.org/10.1016/j.psychres.2014.06.019

Krezmien, M. P., Leone, P. E., & Achilles, G. M. (2006). Suspension, race, and disability: Analysis of statewide practices and reporting. *Journal of Emotional and Behavioral Disorders*, *14*(4), 217–226.

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Lane, K. L., Kalberg, J. R., Menzies, H., Bruhn, A., Eisner, S., & Crnobori, M. (2011). Using Systematic Screening Data to Assess Risk and Identify Students for Targeted Supports: Illustrations Across the K-12 Continuum. *Remedial and Special Education*, *32*(1), 39–54. https://doi.org/10.1177/0741932510361263

Lau, A. S., Garland, A. F., Yeh, M., Mccabe, K. M., Wood, P. A., & Hough, R. L. (2004). Race/Ethnicity and Inter-Informant Agreement in Assessing Adolescent Psychopathology. *Journal of Emotional and Behavioral Disorders*, *12*(3), 145–156. https://doi.org/10.1177/10634266040120030201

Lemstra, M., Neudorf, C., D'Arcy, C., Kunst, A., Warren, L. M., & Bennett, N. R. (2008). A Systematic Review of Depressed Mood and Anxiety by SES in Youth Aged 10-15 Years. *Canadian Journal of Public Health / Revue Canadienne de Sante'e Publique*, *99*(2), 125–129.

Levitt, J. M., Saka, N., Hunter Romanelli, L., & Hoagwood, K. (2007). Early identification of mental health problems in schools: The status of instrumentation. *Journal of School Psychology*, *45*(2), 163–191. https://doi.org/10.1016/j.jsp.2006.11.005

Locke, J., Kang-Yi, C. D., Pellecchia, M., Marcus, S., Hadley, T., & Mandell, D. S. (2017). Ethnic Disparities in School-Based Behavioral Health Service Use for Children With Psychiatric Disorders. *Journal of School Health*, *87*(1), 47–54. https://doi.org/10.1111/josh.12469

Marrast, L., Himmelstein, D. U., & 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*(4), 810–824. https://doi.org/10.1177/0020731416662736

Martinez, A., McMahon, S. D., & Treger, S. (2016). Individual- and School-Level Predictors of Student Office Disciplinary Referrals. *Journal of Emotional and Behavioral Disorders*, *24*(1), 30–41. https://doi.org/10.1177/1063426615588289

McGorry, P., & Purcell, R. (2009). Youth mental health reform and early intervention: encouraging early signs. *Early Intervention in Psychiatry*, *3*(3), 161–162. https://doi.org/10.1111/j.1751-7893.2009.00128.x

McGrath, K. F., & Van Bergen, P. (2015). Who, when, why and to what end? Students at risk of negative student–teacher relationships and their outcomes. *Educational Research Review*, *14*, 1–17. https://doi.org/10.1016/j.edurev.2014.12.001

McLeod, J. D., & Kaiser, K. (2004). Childhood Emotional and Behavioral Problems and Educational Attainment. *American Sociological Review*, *69*(5), 636–658. https://doi.org/10.1177/000312240406900502

Merikangas, K. R., He, J., 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 Study-Adolescent Supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*(10), 980–989. https://doi.org/10.1016/j.jaac.2010.05.017

Merikangas, K. R., He, J., Burstein, M., Swendsen, J., Avenevoli, S., Case, B., … Olfson, M. (2011). Service Utilization for Lifetime Mental Disorders in U.S. Adolescents: Results of the National Comorbidity Survey–Adolescent Supplement (NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry*, *50*(1), 32–45. https://doi.org/10.1016/j.jaac.2010.10.006

Miller, F. G., Cohen, D., Chafouleas, S. M., Riley-Tillman, T. C., Welsh, M. E., & Fabiano, G. A. (2015). A comparison of measures to screen for social, emotional, and behavioral risk. *School Psychology Quarterly*, *30*(2), 184–196. https://doi.org/10.1037/spq0000085

Mitchell, K. J., Tynes, B., Umaña-Taylor, A. J., & Williams, D. (2015). Cumulative experiences with life adversity: Identifying critical levels for targeting prevention efforts. *Journal of Adolescence*, *43*, 63–71. https://doi.org/10.1016/j.adolescence.2015.05.008

Muris, P., Meesters, C., Eijkelenboom, A., & Vincken, M. (2004). The self-report version of the Strengths and Difficulties Questionnaire: Its psychometric properties in 8- to 13-year-old non-clinical children. *British Journal of Clinical Psychology*, *43*(4), 437–448. https://doi.org/10.1348/0144665042388982

Niehaus, K., & Adelson, J. L. (2014). School Support, Parental Involvement, and Academic and Social-Emotional Outcomes for English Language Learners. *American Educational Research Journal*, *51*(4), 810–844. https://doi.org/10.3102/0002831214531323

Ortuño-Sierra, J., Fonseca-Pedrero, E., Paino, M., Sastre i Riba, S., & Muñiz, J. (2015). Screening mental health problems during adolescence: Psychometric properties of the Spanish version of the Strengths and Difficulties Questionnaire. *Journal of Adolescence*, *38*, 49–56. https://doi.org/10.1016/j.adolescence.2014.11.001

Papandrea, K., & Winefield, H. (2011). It's Not Just the Squeaky Wheels That Need the Oil: Examining Teachers' Views on the Disparity Between Referral Rates for Students with Internalizing Versus Externalizing Problems. *School Mental Health*, *3*(4), 222–235. https://doi.org/10.1007/s12310-011-9063-8

Pearcy, M. T., Clopton, J. R., & Pope, A. W. (1993). Influences on Teacher Referral of Children to Mental Health Services: Gender, Severity, and Internalizing Versus Externalizing Problems. *Journal of Emotional and Behavioral Disorders*, *1*(3), 165–169. https://doi.org/10.1177/106342669300100304

Position Statement 41: Early Identification of Mental Health Issues in Young People. (2013, October 17). Retrieved January 29, 2018, from http://www.mentalhealthamerica.net/positions/early-identification

Raines, T. C. ., Tara. Raines@unlv. ed., Dever, B. V. ., Kamphaus, R. W. ., & Roach, A. T. . (2012). Universal Screening for Behavioral and Emotional Risk: A Promising Method for Reducing Disproportionate Placement in Special Education. *Journal of Negro Education*, *81*(3), 283–296.

Reiss, F. (2013). Socioeconomic inequalities and mental health problems in children and adolescents: A systematic review. *Social Science & Medicine*, *90*, 24–31. https://doi.org/10.1016/j.socscimed.2013.04.026

Richter, J., Sagatun, Å., Heyerdahl, S., Oppedal, B., & Røysamb, E. (2011). The Strengths and Difficulties Questionnaire (SDQ) – Self-Report. An analysis of its structure in a multiethnic urban adolescent sample. *Journal of Child Psychology and Psychiatry*, *52*(9), 1002–1011. https://doi.org/10.1111/j.1469-7610.2011.02372.x

Rutherford, R. B., Quinn, M. M., Mathur, S. R., Rutherford Jr., R. B., & Rutherford Jr., R. B. (2014). *Handbook of Research in Emotional and Behavioral Disorders*. New York, UNITED STATES: Guilford Publications. Retrieved from http://ebookcentral.proquest.com/lib/utxa/detail.action?docID=362559

Saxena, S., Hosman, C., Jané Llopis, E., Organisation mondiale de la santé, Dept. of Mental Health and Substance Abuse, & Prevention Research Centre (Pays-Bas). (2004). *Prevention of mental disorders: effective interventions and policy options : summary report*. Geneva: World Health Organization.

Severson, H. H., Walker, H. M., Hope-Doolittle, J., Kratochwill, T. R., & Gresham, F. M. (2007). Proactive, early screening to detect behaviorally at-risk students: Issues, approaches, emerging innovations, and professional practices. *Journal of School Psychology*, *45*(2), 193–223. https://doi.org/10.1016/j.jsp.2006.11.003

Sharpe, D. (2015). Your Chi-Square Test is Statistically Significant: Now What?, *20*(8), 10.

Sheridan, S. M., Witte, A. L., Holmes, S. R., Coutts, M. J., Dent, A. L., Kunz, G. M., & Wu, C. (2017). A randomized trial examining the effects of Conjoint Behavioral Consultation in rural schools: Student outcomes and the mediating role of the teacher–parent relationship. *Journal of School Psychology*, *61*, 33–53. https://doi.org/10.1016/j.jsp.2016.12.002

Silva, J. M., Langhout, R. D., Kohfeldt, D., & Gurrola, E. (2015). "Good" and "Bad" Kids? A Race and Gender Analysis of Effective Behavioral Support in an Elementary School. *Urban Education*, *50*(7), 787–811. https://doi.org/10.1177/0042085914534859

Simon, A. E., Pastor, P. N., Reuben, C. A., Huang, L. N., & Goldstrom, I. D. (2015). Use of Mental Health Services by Children Ages Six to 11 With Emotional or Behavioral Difficulties. *Psychiatric Services*, *66*(9), 930–937. https://doi.org/10.1176/appi.ps.201400342

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

Skiba, R. J., Michael, R. S., Nardo, A. C., & Peterson, R. L. (2002). The Color of Discipline: Sources of Racial and Gender Disproportionality in School Punishment. *The Urban Review*, *34*(4), 317–342. https://doi.org/10.1023/A:1021320817372

Smedley, B., Stith, A., & Nelson, A. (2003). *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care"*. https://doi.org/10.17226/12875

Snowden, L. R., & Yamada, and A.-M. (2005). Cultural Differences in Access to Care. *Annual Review of Clinical Psychology*, *1*(1), 143–166. https://doi.org/10.1146/annurev.clinpsy.1.102803.143846

Spirrison, C. L., Gordy, C. C., & Henley, T. B. (1996). After-Class Versus In-Class Data Collection: Validity Issues. *Journal of Psychology; Provincetown, Mass., Etc.*, *130*(6), 635–644.

Staats, C. (2016). Understanding Implicit Bias, 11.

Tanner, N., Eklund, K., Kilgus, S. P., & Johnson, A. H. (2018). Generalizability of Universal Screening Measures for Behavioral and Emotional Risk. *School Psychology Review*, *47*(1), 3–17. https://doi.org/10.17105/SPR-2017-0044.V47-1

Tenenbaum, H. R., & Ruck, M. D. (2007). Are teachers' expectations different for racial minority than for European American students? A meta-analysis. *Journal of Educational Psychology*, *99*(2), 253–273. https://doi.org/10.1037/0022-0663.99.2.253

Utley, C. A., & Obiakor, F. E. (2015). Special issue: Research perspectives on multi-tiered system of support. *Learning Disabilities: A Contemporary Journal*, *13*(1), 1–2.

van de Looij-Jansen, P. M., Goedhart, A. W., de Wilde, E. J., & Treffers, P. D. A. (2011). Confirmatory factor analysis and factorial invariance analysis of the adolescent self-report Strengths and Difficulties Questionnaire: How important are method effects and minor factors?:

Factorial validity of self-report SDQ. *British Journal of Clinical Psychology*, *50*(2), 127–144. https://doi.org/10.1348/014466510X498174

van den Bergh, L., Denessen, E., Hornstra, L., Voeten, M., & Holland, R. W. (2010). The Implicit Prejudiced Attitudes of Teachers: Relations to Teacher Expectations and the Ethnic Achievement Gap. *American Educational Research Journal*, *47*(2), 497–527. https://doi.org/10.3102/0002831209353594

Vega, W. A., Kolody, B., Aguilar-Gaxiola, S., Alderete, E., Catalano, R., & Caraveo-Anduaga, J. (1998). Lifetime Prevalence of DSM-III-R Psychiatric Disorders Among Urban and Rural Mexican Americans in California. *Archives of General Psychiatry*, *55*(9), 771–778. https://doi.org/10.1001/archpsyc.55.9.771

von der Embse, N. P., Pendergast, L. L., Kilgus, S. P., & Eklund, K. R. (2016). Evaluating the applied use of a mental health screener: Structural validity of the Social, Academic, and Emotional Behavior Risk Screener. *Psychological Assessment*, *28*(10), 1265–1275. https://doi.org/10.1037/pas0000253

Weissberg, R. P., & Bell, D. N. (1997). A Meta-Analytic Review of Primary Prevention Programs for Children and Adolescents: Contributions and Caveats. *American Journal of Community Psychology*, *25*(2), 207–214. https://doi.org/10.1023/A:1024666329372

Weist, M. D., Rubin, M., Moore, E., Adelsheim, S., & Wrobel, G. (2007). Mental Health Screening in Schools. *The Journal of School Health; Kent*, *77*(2), 53–58.

Wheeler, J. J., & Mayton, M. R. (2014). The Integrity of Interventions in Social Emotional Skill Development for Students with Emotional and Behavioral Disorders. In P. Garner, K. James, & E. Julian, *The SAGE Handbook of Emotional and Behavioral Difficulties* (pp. 385–398). 1 Oliver's Yard, 55 City Road London EC1Y 1SP: SAGE Publications Ltd. Retrieved from http://sk.sagepub.com/reference/hdbk_emotionbehaviordifficulties/n28.i2449.xml

Whitesell, E. R., Stiefel, L., & Schwartz, A. E. (2016). Unexpected Arrivals: The Spillover Effects of Mid-Year Entry on Stable Student Achievement in New York City. *Educational Evaluation and Policy Analysis*, *38*(4), 692–713. https://doi.org/10.3102/0162373716661242

Williams, D. R., & Earl, T. R. (2007). Commentary: Race and mental health more questions than answers. *International Journal of Epidemiology*, *36*(4), 758–760. https://doi.org/10.1093/ije/dym114

Wu, P., Hoven, C. W., Bird, H. R., Moore, R. E., Cohen, P., Alegria, M., … Roper, M. T. (1999). Depressive and Disruptive Disorders and Mental Health Service Utilization in Children and Adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, *38*(9), 1081–1090. https://doi.org/10.1097/00004583-199909000-00010