Mental Health Screening within a Tiered Model:

Investigation of a Brief Strength-Based Approach

Natalie Romer

University of Oregon

Research Summary for the Wing Institute

Mental Health Screening within a Tiered Model:

Investigation of a Brief Strength-Based Approach

In the United States, approximately one in five children has a diagnosable mental health problem, and one in ten children experience problems so severe as to impact their daily functioning (Costello, Mustillo, Erkanli, Keeler, & Angold, 2003; Doll, 1996; Department of Health and Human Services, 1999). Even more concerning is that less than half of these children diagnosed with a disorder receive the therapeutic services they require (Ries Merikangas, He, Brody, Fisher, Bourdon, & Kortez, 2010). As a result, schools often function as "de facto" mental health systems for children and adolescents attempting to meet the unmet mental health needs of students and their families (Farmer & Farmer, 1999; Rones & Hoagwood, 2000). Fortunately, schools are well suited to integrate and coordinate mental health services (Doll & Cummings, 2008).

Positive student outcomes achieved via school-based mental health interventions include improved academic skills, attendance, and social behavior (Fleming et al., 2005; Zins, Bloodworth, Weissberg, & Walberg, 2004). In addition, school-based mental health services can be linked to a reduction in special education referrals, improved aspects of school climate, and a decline in discipline referrals, suspensions, and grade retentions (Burns, Walrath, Glass-Siegel, & Weist, 2004; Knoff, 2004). Conversely, failure to prevent mental health problems is linked to increased school drop-out rates, costs associated with intensive mental health care, and an increased number of children entering the juvenile justice system (Fruedenberg & Ruglis, 2007; Hu, 2004; Wasserman, et al., 2004). In order to minimize these costs and improve student outcomes, primary interventions within schools need to begin early through active

mental health promotion so that children and their families are provided the support they

require prior to problems developing (Greenberg, Domitrovitch, & Brumbarger, 2001).

Unfortunately, most students are identified with emotional and behavioral disorders after

opportunities for early intervention have been missed, and the myriad of poor school-related

outcomes have already begun to transpire (Rones & Hoagwood, 2000; Wagner, Kutash,

Duchnowski, & Epstein, 2005).

Three-Tiered Approaches to Providing School-Based Mental Health Services

A focus on the prevention of mental illness and the promotion of factors associated with

resilience within schools provides a promising approach to improving student outcomes and

reducing emotional and behavior problems among children and adolescents. Three-tiered

models of intervention such as response to intervention provide a framework for an integrated

continuum of supports that promote mental wellness and, thereby, prevent the onset of

behavioral, social, and emotional problems (Adleman & Taylor, 2000; Doll & Cummings, 2008;

Merrell & Gueldner, 2008).

If schools are to respond to the call of public and mental health experts to support the

mental wellbeing of children and their families, the infrastructure for systemic, coordinated,

and integrated approaches to mental health programming must be considered (Adelman &

Taylor, 2008; Nastasi, 2004). A public health perspective implies a conceptualization of mental

health service delivery as a continuum of care ranging from mental health promotion to the

treatment of identified mental health disorders across settings within a community, including

schools. Doll and Cummings (2008) identified four goals for population-based mental health

services: "(a) promote the psychological well-being of all students so that they can achieve

developmental competence, (b) promote caretaking environments that nurture students and

allow them to overcome minor risks and challenges, (c) provide protective support to students

at high risk for developmental failures, and (d) remediate social emotional and behavioral

disturbances so that students can develop competence" (p.3). Integrating mental health

supports such as social and emotional learning into three-tiered prevention models has been

recommended as a means for achieving these goals (Doll & Cummings, 2008; Gresham, 2005;

Merrell & Guelder, 2010).

At the universal level of support, prevention efforts involve school-wide approaches

focusing not only on reducing children's aggressive, disruptive, and withdrawn behavior, but

also on developing children's social and emotional competence (Osher et al., 2008). Similar to

three-tiered approaches that have been applied to academics and social behavior, such a model

requires not only prevention and early intervention supports, but also a means of identifying

students who have not responded to universal practices (Greenwood, Horner, & Kratochwill,

2008). Thus, methods for universal mental health screening are needed to identify students

who could benefit from more intensive supports. Although some approaches to universal

screening for mental health and internalizing problems have been recommended (Doll &

Cummings, 2008; Levitt & Merrell, 2009), there is a pressing need for research to investigate

the validity and outcomes associated with these approaches, as well as research and

development of new assessment tools (Levitt, Saka, Romanelli, & Hoagwood, 2007).

Universal Mental Health Screening

Universal mental health screening is necessary if schools plan to make data-based

decisions about the effectiveness of their available supports and to proactively meet the mental

health needs of all students by targeting skills and competencies associated with resilience and

aimed at modifying risk factors. Universal voluntary school-based mental health screening has

been identified as a means for accomplishing this goal (e.g., New Freedom Commission on

Mental Health, 2003); however, research is needed on the use of strength-based measures for

the purpose of mental health screening (Dowdy, Furlong, Eklund, Saeki, & Ritchey, 2010; Levitt,

Saka, Romanelli, & Hoagwood, 2007). Of the existing comprehensive, research-based mental

health screeners for schools, most focus on measuring the presence of social or emotional

concerns (i.e., deficits) (Drummond, 1994; Kamphaus & Reynolds, 2007; Walker & Severson,

1992). This may be problematic, because it cannot be assumed that the absence of symptoms

related to psychopathology alone implies that a student is mentally well or well adjusted (Suldo

& Shaffer, 2008). Many of the most commonly used measures of social-emotional behavior

were also originally developed for diagnostic purposes (i.e., measuring existing symptoms of a

diagnosable disorder) rather than for identifying future risk (Albers, Glover, & Kratochwill,

2007).   Strength-based self-report measures show promise as a viable and socially acceptable

approach to measuring individual characteristics and skills associated with resilience.

Strength-Based Assessment

Behavior rating scales and self-report measures provide an efficient approach to

assessing students' behavioral, social, and emotional functioning (Merrell, 2008). Although

traditional deficit-based assessment measures are useful for identifying disabilities, these

methods do not take into account ecological variables, nor are they likely to inform intervention

planning (Jimerson, Sharkey, Nyborg & Furlong, 2004). Recently developed strength-based

approaches to assessment, however, focus on social and emotional skills, competencies, and

resources within and around the child (Beaver, 2008; Epstein, 1999). Epstein and Sharma's

(1998) commonly-cited definition of strength-based assessment is: " the measurement of those

emotional and behavioral skills, competencies, and characteristics that create a sense of

personal accomplishment, contribute to satisfying relationships with family members, peers,

and adults; enhance one's ability to deal with adversity, and promote one's personal, social,

and academic development" (p. 3). Strength-based assessment practices have evolved from

the field of positive psychology, which focuses on the development of human strengths and

virtues, as well as the prevention of problems (Seligman & Csikszentmihalyi, 2000). These

approaches are aligned with the current shift in best practice towards models that emphasize

solutions to problems rather than assessment for the purpose of identification or eligibility

(Merrell, Ervin, & Gimpel, 2006; Tilly, 2008).

     These measures provide parents and educators with information on student skills,

knowledge, and competence, which may be better suited for evaluating the outcomes of

intervention programs and universal screening. In addition, a strength-based approach to

mental health screening may be more socially acceptable and increase motivation for change

compared to problem focused approaches to assessment (Jimerson, Sharkey, Nyborg, &

Furlong, 2004). A strength-based approach to mental health screening provides an approach

that is focused on the contextual factors related to solving problems (i.e., teaching skills and

changing contingencies) that are aligned with the aims of school-based mental-health

promotion, and thus, may be more socially valid and less constrained by ethical and legal

considerations than traditional diagnostic measures associated with mental illness.

Schools must consider how to effectively address concerns related to ethical dilemmas and social validity by working with families and demonstrating how mental health screening procedures improve access to supports for students and improved student outcomes. A strength-based approach to mental health screening may not only reduce some of these barriers, but may shift the focus away from metal health problems and towards solutions that schools are more motivated and capable of addressing (Dowdy, et al., 2010).

Study Purpose

The purpose of this study was to evaluate the classification accuracy and other forms of validity of a universal screening approach using a very brief (12 items) strength-based self-report measure, the Social-Emotional Assets and Resilience Scales (student short forms) (SEARS; Merrell, 2011), in identifying middle school students who are at-risk or in need of intensive social-behavioral and mental health intervention. Specifically, this study addressed the following research questions:

1. Do reported levels of internalizing and externalizing symptoms on the Youth Self-Report (YRS; Achenbach & Rescorla, 2001) effectively and accurately discriminate between students identified as not at risk and at-risk based on the proposed strength-based approach to mental health screening?

2. Do scores from the Behavioral and Emotional Screening System Teacher Version (BESS; Kamphaus & Reynolds, 2007) effectively and accurately discriminate between students identified as not at risk and at-risk based on the proposed strength-based approach to mental health screening?

3. What is the degree of similarity between student and teacher ratings of a given

   student using respective versions of a strength-based rating scale (SEARS-Short

   Forms; Merrell, 2011)?

4. What is the degree of difference between strength-based rating scale scores

   (SEARS-Short Forms; Merrell, 2011) and primarily problem-based rating scale scores

   (BESS; Kamphaus & Reynolds, 2007; YSR; Achenbach & Rescorla, 2001)?

5. Are there significant group differences between the At-Risk and Not-At Risk group

   according to academic functioning, number of absences, number of office discipline

   referrals (ODR), gender, and disability status?

Phase I Method

Phase I Setting

Data were collected in a mid-sized (approximately 6,000 students), suburban school district surrounding the University of Oregon (NCES, 2010).   After receiving approval from both the University of Oregon's Internal Review Board and the district in which this study took place, the four middle schools in this district volunteered to participate.

Common Core Data from the National Center for Education Statistics indicates that 54.5% of the total student population across these four schools were eligible for free or reduced priced lunch. Student attendance rates at the participating schools during the 2010-2011 school year ranged from 93.3% to 95.9%, according to publicly available data through the Oregon Department of Education. All schools met AYP criteria for student participation and attendance. All subgroups represented at School 1 met state standards for annual yearly progress (AYP) based on the Oregon Statewide Assessment for English Language Arts (ELA). One of the five subgroups represented at School 1 met state standards for AYP criteria based on the Oregon Statewide Assessment for math. With the exception of Students with Disabilities, all subgroups represented at School 2 met state standards for AYP in both ELA and math. At Schools 3 and 4, all represented subgroups met state standards for AYP in ELA and all represented subgroups, with the exception of Students with Disabilities, met state standards for AYP in Math.

The four participating schools had been implementing Tier I of school wide positive behavior support (SWPBS) for at least two years, as evidenced by scores greater than 80% on the Benchmarks of Quality scores (Kincaid, Childs, & George, 2005). The Benchmarks of Quality

is a valid and reliable instrument (Kincaid, et al., 2005) and scores of 70% or higher are indicative of full implementation of Tier I of SWPBS (Cohen, Kincaid, & Childs, 2007). In these schools, Tier I of SWPBS consisted of school-wide expectations that were defined operationally for specific settings. Expectations were visibly posted throughout each building and taught to students throughout the school year. School staff members acknowledged students demonstrating prosocial behaviors aligned with their school's expectations with praise and "tokens" that were entered into school raffles or could be exchanged for prizes. Each of these middle schools has also developed a continuum of consequences for various types of problem behaviors. Office discipline referral data were collected through the School-Wide Information System (SWIS; May et al., 2006). SWIS and other data are reviewed regularity by a school-based team and a district-level team and used to refine Tier I intervention.

Phase I Participants

Participants were 6th, 7th, and 8th grade students and their teachers. Demographic information provided by Phase I participants is provided in Table 1.

Table 1

Student Characteristics as a Percentage of Each School's Population

Variable	School 1 (*n* = 417)	School 2 (*n* = 318)	School 3 (*n* = 200)	School 4 (*n* = 241)
Gender				
Female	48.0	54.7	45.5	48.5
Male	49.9	43.7	51.5	50.6
Not Reported	2.2	1.6	3.0	0.8
Grade				
Sixth	33.3	32.7	28.5	32.4
Seventh	33.3	37.4	37.5	32.0
Eight	30.7	29.6	32.0	35.3
Not Reported	2.6	0.3	2.0	0.4
Age				
Ten	0.0	0.0	0.0	0.4
Eleven	28.5	27.4	22.0	26.1
Twelve	31.7	34.9	40.0	33.2
Thirteen	32.1	31.8	32.5	31.5
Fourteen	6.0	5.3	4.5	7.5
Fifteen	0.0	0.0	0.0	0.8
Not Reported	1.7	0.6	1.5	0.4
Ethnicity				
White/Caucasian	65.7	56.0	77.5	62.7
Hispanic/Latino	9.4	14.2	5.5	17.0
Black/African American	1.7	2.2	1.5	3.3
Asian/Pacific Islander	3.4	0.9	4.5	2.5
American Indian/ Native American	2.9	3.1	1.5	2.9
Multiracial (2+)	10.3	20.1	3.5	9.5
Other	3.1	2.5	4.0	1.7
Not Reported	3.6	0.9	2.0	0.4

Phase I Instrumentation

Social-emotional assets and resilience scales - short form. The Social-Emotional Assets and Resilience Scales (SEARS; Merrell, 2011) is a strength-based assessment system designed to measure positive social-emotional attributes and skills (e.g., self-regulation, social and emotional knowledge and competence, problem solving skills, empathy) of children and adolescents in grades kindergarten to 12. The SEARS multi-informant system includes short-form versions (SEARS-SF) consisting of 12 items for each of the respective versions (Merrell, 2011). In this study the child and adolescent short form versions of the SEARS were used as the mental health screener. In addition, teachers completed the teacher version of the SEARS-SF on the subsample of students who participated in Phase II. SEARS are scored using a 4-point scale from 0 (never) to 4 (almost always). A higher score indicates a higher level of perceived social emotional strength and resilience.

The SEARS assessment system was developed with a sample selected to approximate 2009 U.S. Census distributions for ethnic groups with some oversampling of minority groups (Merrell, 2011). SEARS short forms are highly correlated with each of their respective full-length versions of the SEARS (SEARS-C = .93, SEARS-A = .94, SEARS-T = .98; Merrell, 2011; Nese et al., in press). SEARS short forms have adequate internal consistency (α = .85 for SEARS-C, .83 for SEARS-A, and .93 for SEARS-T) and two week test re-test reliability (r = .74 for the SEARS-C, .84 for the SEARS-A, and .91 for the SEARS-T).

Pearson product-moment correlations between SEARS short forms and the other strength-based rating scales indicate that SEARS short forms are measuring the social and emotional constructs that the measure was designed to assess (Nese, et al., in press). The

SEARS-T has been shown to be significantly correlated to with the Social Skills Rating Scale (SSRS; Gresham & Elliott, 1990) and the Peer Relations scale of the School Social Behavior Scales (SSBS-2; Merrell, 2002) with Pearson correlation coefficients ranging from.79 and .89. The SEARS-C is significantly correlated with the SSRS subscale and the total scores (Pearson correlation coefficients: .62 to .78) and significantly negatively correlated (-.47) with the Internalizing Symptoms Scale for Children Positive Affect subscale (Merrell & Walters, 1998). The SEARS-A SF is significantly correlated with the SSRS subscale and total scores (Pearson correlation coefficients: .67 to .72) and moderately correlated (.44) with the Students' Life Satisfaction Scale (Huebner, 1991; .44).

SEARS short forms take three to five minutes to complete and require about a third grade reading level. During phase one of the study the self-report versions of the SEARS short-forms were used as the Tier I mental health screener; the child form (SEARS-C) was completed by sixth grade students and the adolescent version (SEARS-A) was completed by seventh and eighth grade students. During phase two, participating students and their teachers completed respective versions of the SEARS short forms.

Phase I Procedures

An overview of the study and brief training on the administration procedure for the mental health screener was presented by the researchers at each of the participating schools' faculty meetings, at which time teachers also consented to participate. Across the four schools, 49 teachers administered the mental health screener to the students in their classes.

Administration packets containing a copy of the administration procedures, a script to read to students, and copies of the mental health screener (i.e., SEARS-SF) were prepared at the

University of Oregon for each classroom. Students who were absent on the day that the mental

health screener was administered at their school were not included in the study, as there were

no make-up days.

One thousand, one hundred and ninety-two students (90%) of the 1,324 enrolled

students completed the SEARS-SF. These were the students who were present, eligible to

participate (i.e., their parents had not called the school to opt them out of the study), and

agreed to complete the mental health screener on the administration day. Participants were

excluded from the study if they left more than one of the 12 items on the mental health

screener blank or there was a clearly visible response pattern suggesting that the form was

invalid (e.g., the same response was circled for every item). The final sample for Phase I

included 1,176 students (response rate = 88.8%).

Trained research assistants entered the data from the mental health screener into SPSS.

To ensure accuracy of data entry, 20% of the participants' data were randomly selected and

checked for accuracy by having a second coder independently enter the data and checking for

agreement. Agreement was over 99% across all variables entered.

Once the data from Phase I were entered, students were assigned to one of three tiers

(Tier I, Tier II and Tier III) using cut scores from the SEARS assessment system's national

norming sample (Merrell, 2011). Students whose scores fell in the top 80%, the middle 15%,

and the bottom 5% were assigned to Tiers I, II, and III respectively (i.e., students were assigned

to tiers based on the percentile rank of their score on the SEARS-SF). Tiers II and III were then

collapsed to create one at-risk group, which is aligned with prior research indicating that about

20% of children in the U.S. have a behavioral and/or emotional problem and are in need of

services (Costello, Mustillo, Erkanli, Keeler, & Angold, 2003). The percentage and number of

students from the participating district assigned to each tier are presented in Figure 1.

Figure 1. Sample Selection Procedure and Size for Each Phase of the Study
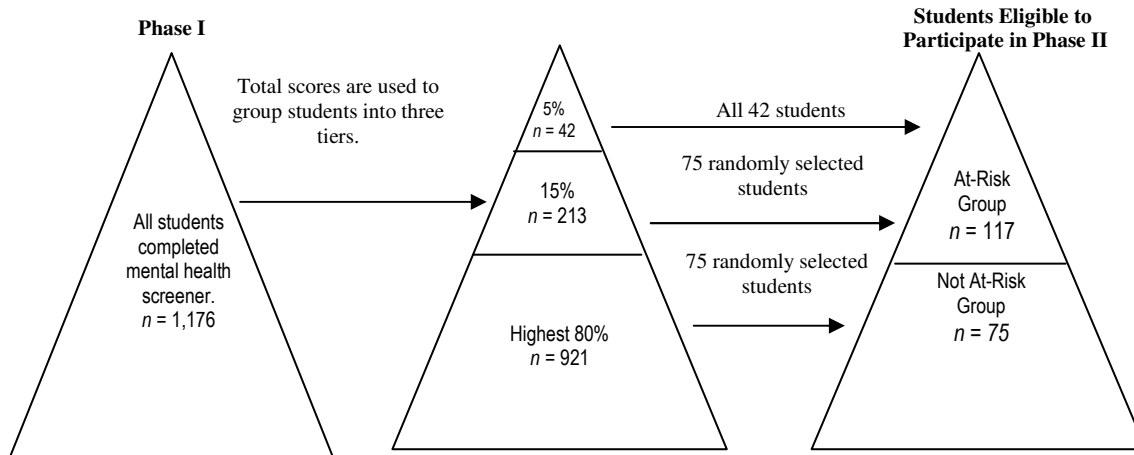
Figure 1. Sample selection procedure used to group students into two groups (At-Risk and Not
At-Risk) for Phase II of the study.

        Table 2 depicts the number and percentage of students assigned to each tier, how many

students were selected to be invited to participate in Phase II of the study, and group

classifications.

Table 2
Participants Assigned to Each Tier

Group	Tier I	Tier II	Tier III
Phase I participants (*n* = 1,176)	921 (78.3%)	213 (18.1%)	42 (3.6%)
Students selected for Phase II	75	75	42
Group classification	Not At-Risk	At-Risk	At-Risk

Phase I Results

Missing data. SEARS-SFs with more than one of 12 items incomplete were considered invalid and not included in the final sample. Of the 1,176 student participants, 42 students left one item blank. Missing data appeared to be random and missing data were replaced with mode imputations. Though imputations can falsely increase or decrease the sample mean, this method is commonly used in research (Chen & Astebro, 2003), including during the development of the SEARS-SF (Merrell, 2011).

Descriptive statistics. Mean total scores and standard deviations across tiers are presented in Table 3. Mean scores demonstrated a downward trend across tiers as would be expected given that tiers were assigned using total scores.

Table 3
Descriptive Statistics for the SEARS-C and SEARS-A Short Forms from Phase I

	M	*SD*	*Min*	*Max*
SEARS-C-SF				
Tier III (*n* = 19)	8.95	2.01	4	11
Tier II (*n* =65)	14.54	1.28	12	16
Tier I (*n* = 300)	22.79	3.74	17	33
SEARS-A-SF				
Tier III (*n* = 23)	9.09	2.30	6	12
Tier II (*n* = 148)	15.51	1.30	13	17
Tier I (*n* = 621)	23.27	3.72	18	32

Note. Min = Minimum reported score; *Max* = Maximum reported score.

Data presented in Table 4 show the percentage and number of students from each

school assigned to the three tiers. Table 4 also demonstrates that the participating schools

had a similar percentage of students assigned to each tier as the national norming sample of

the SEARS.

Table 4
Percentage of Students Assigned to Each Tier

	School 1 (n = 417)	School 2 (n = 318)	School 3 (n= 200)	*School 4 (n = 241)	Total (n = 1,176)
Tier III	2.4% (n = 10)	4.1% (n = 13)	5.5% (n = 11)	3.3% (n = 8)	3.6% (n = 42)
Tier II	17.5% (73)	20.1% (n = 64)	16.5% (n = 33)	17.8% (n = 43)	18.1% (n = 213)
Tier I	80.1% (n = 334)	75.8% (n = 241)	78.0% (n = 156)	78.8% (n = 190)	78.3% (n = 921)

Note. *School 4 was not included in the final sample.

Internal consistency reliability. We used Cronbach's alpha procedure to calculate

internal consistency on the total score of the SEARS-SF. Reliability coefficients are presented in

Table 5. Alpha levels of internal consistency of this sample were similar to alpha coefficients of

the national norming sample (i.e., SEARS-C = .85 and SEARS-A = .83; Nese et al., in press).

Table 5
Chronbach's Alpha Coefficients of the SEARS-C and SEARS-A Short-Forms

Scale	# of items	α
SEARS-C-SF	12	.81
SEARS-A-SF	12	.80

Phase II Method

Phase II Participants

A total of 192 students, comprised of 75 randomly selected students from Tiers I and II,

plus all 42 students from Tier III, were invited to participate in Phase II of the study. Table 6

summarizes the number of participants who participated in Phase I and the subsample of

participants who participated in Phase II.

Table 6
Distribution of Participants across Tiers and Risk Classification

	Tier I	Tier II	Tier II
Participants in Phase I (n)	921	231	41
Group classification	Not at-risk		At-risk
Participants in Phase II (n)	61		45

Phase II of the study required active consent from eligible student participants'

parents. A variety of strategies were employed from November to late March to secure a

sufficient sample size. After three mailings using a variety of approaches to recruit participants

for Phase II, researchers and district administrators agreed to try a more targeted approach to

recruiting Phase II participants at one of the four middle schools. 104 students from the

selected middle school were resampled and invited to participate in Phase II of the study (i.e.,

52 students from Tier I and 52 students from Tiers II and III).

In the spring, the fourth school chose not to participate in Phase II due to other

commitments on teacher time. See Figure 2 for a summary of how participants were selected
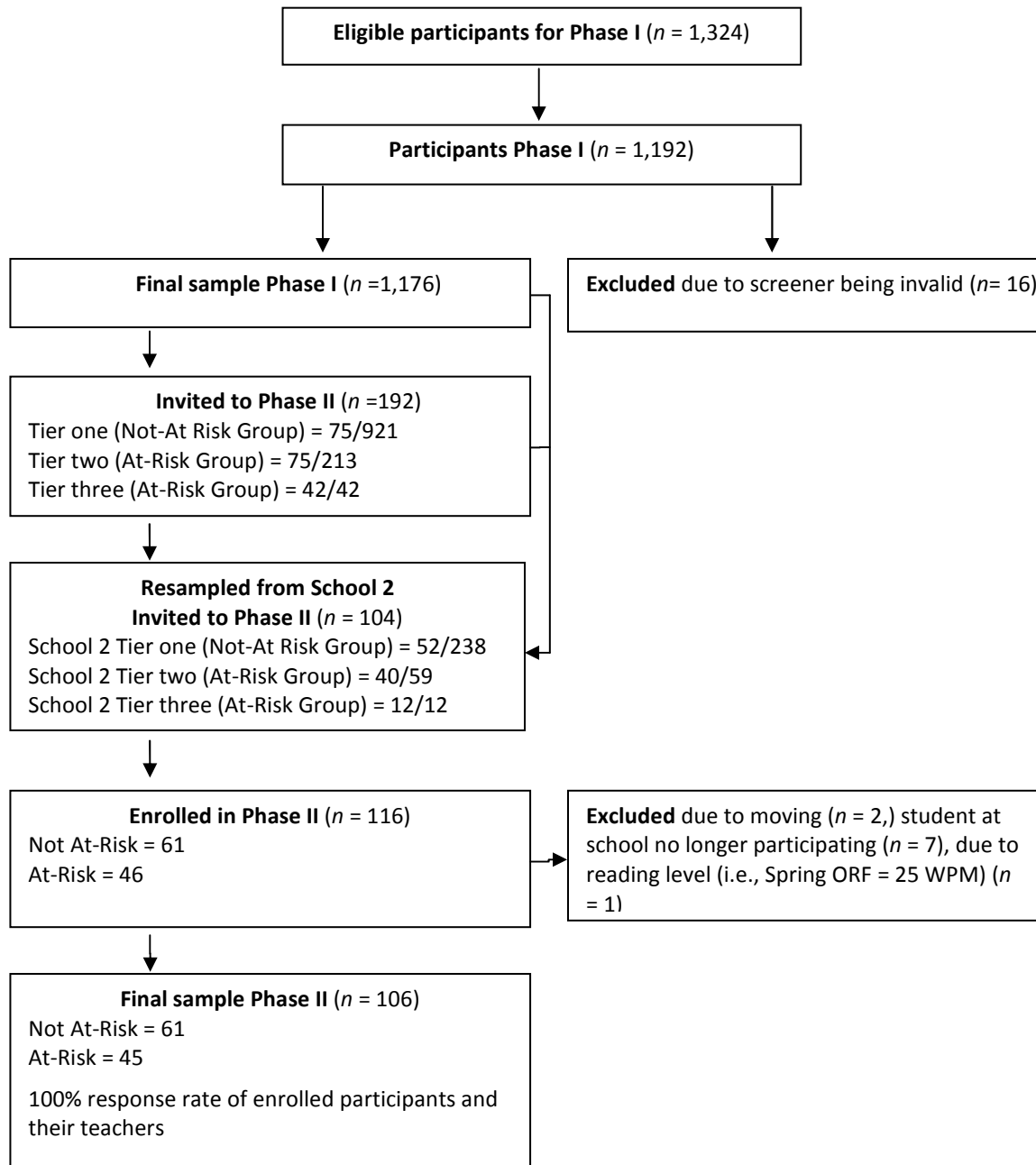
and excluded from participating.

Figure 2. Flow of Participants in Mental Health Screening Study

```
                    ┌──────────────────────────────────────────┐
                    │ Eligible participants for Phase I         │
                    │ (n = 1,324)                               │
                    └──────────────────────────────────────────┘
                                        │
                                        ▼
                    ┌──────────────────────────────────────────┐
                    │ Participants Phase I (n = 1,192)          │
                    └──────────────────────────────────────────┘
                          │                          │
                          ▼                          ▼
    ┌──────────────────────────────┐   ┌──────────────────────────────────────┐
    │ Final sample Phase I         │   │ Excluded due to screener being        │
    │ (n =1,176)                   │   │ invalid (n= 16)                       │
    └──────────────────────────────┘   └──────────────────────────────────────┘
                │
                ▼
    ┌──────────────────────────────────────────┐
    │ Invited to Phase II (n =192)              │
    │ Tier one (Not-At Risk Group) = 75/921     │
    │ Tier two (At-Risk Group) = 75/213         │
    │ Tier three (At-Risk Group) = 42/42        │
    └──────────────────────────────────────────┘
                │
                ▼
    ┌──────────────────────────────────────────────────┐
    │ Resampled from School 2                           │
    │ Invited to Phase II (n = 104)                     │
    │ School 2 Tier one (Not-At Risk Group) = 52/238    │
    │ School 2 Tier two (At-Risk Group) = 40/59         │
    │ School 2 Tier three (At-Risk Group) = 12/12       │
    └──────────────────────────────────────────────────┘
                │
                ▼
    ┌──────────────────────────────┐   ┌──────────────────────────────────────┐
    │ Enrolled in Phase II         │   │ Excluded due to moving (n = 2,) student│
    │ (n = 116)                    │   │ at school no longer participating      │
    │ Not At-Risk = 61             │──▶│ (n = 7), due to reading level (i.e.,   │
    │ At-Risk = 46                 │   │ Spring ORF = 25 WPM) (n = 1)           │
    └──────────────────────────────┘   └──────────────────────────────────────┘
                │
                ▼
    ┌──────────────────────────────────────────┐
    │ Final sample Phase II (n = 106)           │
    │ Not At-Risk = 61                          │
    │ At-Risk = 45                              │
    │                                           │
    │ 100% response rate of enrolled            │
    │ participants and their teachers           │
    └──────────────────────────────────────────┘
```

Figure 2. Participant flow chart and outline of the sampling procedures used to obtain final study sample.

Participants in Phase II included a sub-sample of 106 students and their teachers. Of the

106 students that participated in Phase II, 79 were from School 2, 17 from School 1, and 10

from School 3. The demographic characteristics of the student participants from Phase II of the

study are summarized in Table 7.

Table 7
Characteristics of Phase II Participants

Variable	% of Sample
Risk status	
At-risk (Tiers II and III)	42.5
Not at-risk (Tier I)	57.5
Gender	
Female	46.2
Male	53.8
Grade	
Sixth	34.0
Seventh	37.7
Eight	28.3
Special education status	
No special education	84.0
Special education	16.0
English as a second language	
Not eligible	100.0
Eligible	0.0
Free and reduced lunch eligible	
Eligible	66.0
Not eligible	34.0
Ethnicity	
White	69.8
Hispanic	20.8
Black/African American	4.7
American Indian/ Alaskan Native	3.9
Multiple	0.9

During Phase II of the study, teachers completed behavior rating scales for participating students. Participating teachers' years of experience in education ranged from one to 33 years of experience (*M* = 13). Fifty-six percent of the behavior rating scales were completed by female teachers and 44% by male teachers. The majority of teachers reported teaching in general education classrooms (89.7%), followed by other setting (2.8%), special education classroom (0.9%), and another setting (e.g., reading, gifted program) (0.9%). Of the participating teachers, 5.6% did not specify a setting.

Phase II Instrumentation

Youth self-report form. The Youth Self-Report (YSR; Achenbach & Rescorla, 2001) is one component of a multi-axial behavioral assessment procedure, the Achenbach System of Empirically Based Assessment. The YSR is a self-report measure for children and adolescents between the ages of 11 – 18 and consists of 112 items rated on a 3-point Likert scale (0 = Not True, 1 = Somewhat or Sometimes True, and 2 = Very True or Often True). Reading the items on the YSR requires a fifth grade reading level.

The structure of the YSR includes two broadband scales: Externalizing and Internalizing, and eight narrow-band subscales (i.e., syndromes). In addition, a Total Problem score can be computed. The Externalizing Problems broadband scale consists of the Rule-Breaking Behavior and Aggressive Behavior subscales. The Internalizing Problems broadband scale is comprised of the Anxiety/Depressed, Withdrawn/Depressed, and Somatic Complaints subscales.

The YSR was standardized with a population that closely reflects U.S. population estimates for ethnicity, region, and socioeconomic status. The YSR has adequate internal consistency (α = .91 for the Internalizing scale and .92 for the Externalizing scale) and test re-

test reliability *(r* = .91 for the Internalizing scale and .92 for the Externalizing scale; Achenbach &

Rescorla, 2001). The ASEBA system has also proven to have acceptable convergent validity with

other measures of emotional and behavioral functioning such as the parent and teacher

versions of the Behavioral Assessment System for Children (Pearson correlation coefficients: .75

to .83 for Internalizing Scales and .74 to .88 for Externalizing scales) and the Conners Parent and

Teacher Rating Scales-Revised (.71 to .89; Achenbach & Rescorla, 2001).

Behavioral and emotional screening system. The Behavioral and Emotional Screening

System is a multi-disorder screening system (BESS; Kamphaus & Reynolds, 2007) designed to

identify emotional and behavioral strengths and weaknesses in students from preschool to high

school. Three versions of the report form are available: student, parent, and teacher. This

study used the teacher form, which consists of 27-items and takes approximately five to ten

minutes to complete. The BESS uses a 4-point scale (never, sometimes, often, and almost

always). The teacher-form produces a single score and provides a risk-level classification for

emotional and behavioral problems that can fall within the range of one of three categories or

levels of risk: normal, elevated, or extremely elevated.

The BESS was normed on a representative sample that closely matches recent U.S.

Census population characteristics. The teacher version of the BESS has a test-retest reliability

estimate of .91, inter-rater reliability estimate of .71, and split-half reliability estimate of .96.

The BESS teacher form has been shown to correlate moderately to strongly with other

measures of emotional and behavioral functioning such as the ASEBA (Externalizing Composite

= .68, Internalizing Composite = .28, and Total Problems = .75), Conners' Rating Scale Revised

(Global Index = .73, ADHD Index = .79, DSM IV Symptoms =.78), and Vineland II Teacher Rating

Form, Child Adolescent Version (Adaptive Behavior Composite = -.69, Communication = -.63,

Daily Living Skills= -.63, Motor Skills = -.55; Kamphaus & Reynolds, 2007). Students' risk-level

classifications have also been shown to be significantly related to school-based outcome

criterion (Renshaw et al., 2009).

Oregon assessment of knowledge and skills. Student data on the Oregon statewide

assessment system was obtained from student records. Oregon Assessment of Knowledge and

Skills (OAKS) scores in the areas of reading and math were recorded for participating sixth,

seventh, and eighth grade students. The OAKS is considered a psychometrically sound measure

and is correlated with other measures of achievement (California Achievement Test, Iowa Test

of Basic Skills, NWEA Subject Tests, and Lexile Scale), with correlations ranging from .73 to.84

(ODE, 2007).

Oral reading fluency. The participating district uses reading fluency scores as

benchmarks of student performance (AIMSweb, 2002).   Students were assessed using oral

reading fluency (ORF) passages three times over the course of the school year - fall, winter, and

spring. ORF passages assess a student's accuracy and rate in connected text. The AIMSWeb

benchmark scores have been shown to have good reliability (Christ & Silberglitt, 2007; Howe &

Shinn, 2002).

Cumulative grade point average.   Grades in this district are reported on a 4-point scale,

with higher grades indicating better academic performance. The cumulative grade point

average takes into account student grades earned over all grading periods in the 2010-2011

school year.

Absences. Absences were recorded as the total number of school days missed across the school year in which the study was conducted (2010-2011). Excused and unexcused absences were combined to obtain a total number of days absent. The participating district reported absences by the half day (e.g., 2.5 days).

Office discipline referrals. Student office discipline referral (ODR) data were retrieved by the district from the School-Wide Information System (SWIS; May et al., 2006). The SWIS is a secure web-based data-management system that allows schools to enter and monitor individual students, and to review school wide trends in ODR data across student groups, locations, times, and behaviors. The SWIS can be used to produce discipline reports that schools use for data-based decision making across all tiers of SWPBS. In the SWIS, two types of ODR data exist: minor (e.g., low-intensity disruption, inappropriate language) and major (e.g., abusive language, fighting). Type and frequency of ODRs can be used as screening measures for additional behavior support in middle schools (e.g., Tobin, Sugai, & Colvin, 1996; Tobin & Sugai, 1999).

Procedures

During Phase II, students completed self-report measures (i.e., YSR, a social validity question, and the SEARS-SF), teachers completed behavior rating scales on participating students (i.e., BESS, SEARS-SF), and data were collected from student records (i.e., ODRs, number of absences, etc.).

All 107 students completed both the SEARS-SF and YSR. At each group administration, the lead researcher reviewed the information in student assent form and provided students with directions for completing the YSR and SEARS-SF. Most students completed the YSR and

SEARS-SF in approximately 20 to 25 minutes. For their participation, students received $5.00

gift card and pencil.

For each participating student, a teacher completed the BESS and teacher version of the

SEARS-SF. All participating schools were comprised of teams of teachers who taught core

subject areas (e.g., Math, Science, Language Arts, and Social Studies) and had been teaching

participating students for at least one period per day since the beginning of the school year.

The return rate for teachers completing the behavior rating scales was 100%. Teachers

received a $5.00 gift card or credit to purchase school supplies for each student for whom they

completed behavior rating scales.

Trained research assistants entered student and teacher data converting data into

electronic files using SPSS. A second researcher scored twenty percent of randomly selected

participants' measures and agreement was 99%.

Phase II Results

Results from Phase II are presented and arranged in sections by analyses and research

questions. All analyses were conducted using SPSS 18.0 Grad Pack (SPSS, 2009).

Of the 106 Phase II student participants, 45 were identified to be At-Risk and 61 Not At-

Risk. G*Power 3.1 was used to establish whether statistical power was sufficient to address the

primary research question. A sensitivity analysis was performed to determine the effect size

that would be detectable with 80% power. A two-tailed independent groups t-test with alpha

of .05 and two groups of size 45 and 61 has 80% power to detect an effect size of d = .56, which

is a medium or moderate effect size (Cohen, 1992). Therefore, the sample size was deemed

sufficient to address the primary aims of the study.

Missing data. For the published norm-references measures (BESS, YSR) missing data at the item level were treated in accordance with the publishers' manuals. All forms were sufficiently complete so that they could be scored (Achenbach & Rescorla, 2001; Kamphaus & Reynolds, 2007; Merrell, 2011).

Data on the norm referenced measures (BESS, YSR, and SEARS-SF) were screened for normality, range restriction, outliers, and missing data for each of the two classification groups. Within the At-Risk group, three outliers were identified on the student version of the SEARS-SF administered in the fall. These three outliers represented students with the lowest scores on the SEARS-SF, who had been assigned to Tier III based on of the study's research design. No other severe outliers were identified within in the At-Risk group on any other measure.

Within the Not At-Risk group, one severe outlier (i.e., the highest score) was identified on the student version of the SEARS-SF administered in the fall. On the YSR, the Not At-Risk group had one severe outlier on the Internalizing Problems scale, two on the Externalizing Problems scale, and three on the Total Problems scale. The severe outliers on the Total Problems Scale were the same three outliers from the Internalizing and Externalizing Problems scale. Prior to the logistic regression analyses, outliers were further analyzed using Cook's D, leverage scores, and DFBETA, which established that outliers would not cause an undue effect on the equation nor would they be potentially influential (i.e., significantly change the intercept as a function of deleting the outlier). Thus all cases were retained for all analyses. No other severe outliers were identified within the Not At-Risk group on any other measure.

Descriptive statistics. Descriptive statistics for each measure and group are provided in Tables 8 and 9. Mean scores and standard deviations follow the expected trend with the Not

At-Risk group having a higher mean score on the SEARS (i.e., strength-based measure) and lower scores on the YSR and BESS (i.e., primarily deficit-based measures) than the At-Risk group. Mean scores on the SEARS-SF also increased from when students completed the form in the fall and again in the spring. Analyses were conducted using raw scores [note that correlations between T scores and raw scores were very high (YSR Total = .975, YSR Internalizing = .954; YSR Externalizing = .977, SEARS-SF Fall = .951, SEARS-SF Spring = .998, BESS = 1.000, and SEARS-T = .999), and results essentially remained the same when using T scores and raw scores].

Table 8
Descriptive Statistics for Student Completed Forms

Measure	Not at-risk group n = 61				At-risk group n = 45			
	Min	*Max*	*M*	*SD*	*Min*	*Max*	*M*	*SD*
SEARS-SF								
Fall	17	32	21.93	3.62	6	17	14.24	2.86
Spring	14	36	24.36	5.04	12	30	19.51	4.41
YSR								
Internalizing	0	43	11.62	8.30	0	40	15.22	10.48
YSR								
Externalizing	0	38	8.97	7.26	3	33	12.98	7.27
YSR								
Total	0	113	39.57	24.07	10	131	53.40	28.32

Note. Min = Minimum reported score; *Max* = Maximum reported score.

Table 9
Descriptive Statistics for Teacher Completed Forms

Measure	Not at-risk group $n = 61$				At-risk group $n = 45$			
	Min	*Max*	*M*	*SD*	*Min*	*Max*	*M*	*SD*
SEARS-SF	3	636	20.69	7.70	4	36	19.33	8.25
BESS	0	47	21.20	12.39	0	61	27.07	14.73

Note. Min = Minimum reported score; *Max* = Maximum reported score.

Logistic regression analyses. Four separate logistic regression analyses and independent groups t-tests were used to answer the research questions - *Do reported levels of internalizing and externalizing symptoms on the Youth Self-Report (YSR; Achenbach & Rescorla, 2001) effectively and accurately discriminate between students identified as not at risk and at-risk based on the proposed strength-based approach to mental health screening? Do scores from the Behavioral and Emotional Screening System Teacher Version (BESS; Kamphaus & Reynolds, 2007) effectively and accurately discriminate between students identified as not at risk and at-risk based on the proposed strength-based approach to mental health screening?*

Four one-predictor logistic models were fitted to the data to examine the relationship between risk classification (i.e., whether a student was identified as being At-Risk or Not At-Risk using the strength-based approach to mental health screening under investigation) and YSR (Achenbach & Rescorla, 2001) and BESS scores (Kamphaus & Reynolds, 2007). The dependent variable was risk classification with two levels: Not At-Risk and At-Risk. Four predictor (independent) variables were analyzed: Externalizing, Internalizing, and Total Problems scores from the YSR and the Total score from the teacher version of the BESS.

To determine if any of the independent variables were predictors of risk-classification, separate logistic regression analyses were used in order to meet the assumption of multicollinearity (Pedhazur, 1997). Given that predictor variables were expected to be correlated with one another, multicollinearity was investigated by examining zero-order correlations among independent variables and Variance Inflation Factor (VIF) values. Correlations between YSR scores ranged from .67 to .91 ($p < .001$). Average VIF values ranged from 8.65 to 18.52 with an average VIF value of 10.94 suggesting that multicollinearity could bias the regression model (Bowerman & O'Connel, 1990; Myers, 1990).

Four independent logistic regression analyses revealed that all YSR scores (Internalizing, Externalizing, and Total) and the BESS score were significant or nearly significant ($p = .055$ for YSR Internalizing Problems Scale Score) predictors of students risk classifications (see Table 10). The Hosmer-Lemeshow test was insignificant for all predictor variable ($p > .05$) suggesting that each model fit the data. According to the models, the odds of a child being identified at-risk was positively related to all predictor variables. The higher the score on the YSR or BESS the more likely the student was to be identified at-risk.

Table 10

Logistic Regression Analyses of Risk Classification

Predictor	β	SE β	df	Wald's χ^2	P	e^{β} (odds ratio)
YSR Internalizing	.042	.022	1	3.672	.055	1.043
YSR Externalizing	.076	.029	1	6.821	.009	1.079
YSR Total	.020	.008	1	6.492	.011	1.021
BESS	.033	.015	1	4.634	.031	1.033

Because classification was a goal of these analyses, the classification tables for each of the predictor variables is provided in Tables 11 to 14. Predicted probabilities of at-risk status were retained from the logistic regression analyses. Dichotomized probabilities (below or above .50) were used to place individuals in their most likely category. Tables 11 to 14 provide information regarding the validity of the predicted probabilities using a cutoff set at 0.50. The percentage listed in the first row of each of the classification tables indicates the magnitude of specificity or proportion of correctly non-identified students (i.e., not at-risk). The percentage listed in the second row indicates the sensitivity or proportion of correctly identified students (i.e., at-risk). The overall classification accuracy for the predictor variables in the logistic regression analyses ranged from 60.4 to 64.2, which is an improvement to the chance level, but low.

Table 11
The Observed and Predicted Frequencies for Risk Classification by Logistic Regression for the YSR Internalizing Problems Scale

	Predicted		
Observed	No	Yes	% Correct
No (not at-risk)	51	10	83.6
Yes (at-risk)	32	13	28.9
Overall % correct			60.4

Note. Cutoff = 0.50.

Table 12
The Observed and Predicted Frequencies for Risk Classification by Logistic Regression for the YSR Externalizing Problems Scale

	Predicted		
Observed	No	Yes	% Correct
No (not at-risk)	49	12	80.3
Yes (at-risk)	26	19	42.2
Overall % correct			64.2

Note. Cutoff = 0.50.

Table 13
The Observed and Predicted Frequencies for Risk Classification by Logistic Regression for the YSR Total Problem Scale

	Predicted		
Observed	No	Yes	% Correct
No (not at-risk)	51	10	83.6
Yes (at-risk)	29	16	35.6
Overall % correct			63.2

Note. Cutoff = 0.50.

Table 14
The Observed and Predicted Frequencies for Risk Classification by Logistic Regression for the BESS Teacher Report

	Predicted		
Observed	No	Yes	% Correct
No (not at-risk)	50	11	82.0
Yes (at-risk)	30	15	33.3
Overall % correct			61.3

Note. Cutoff = 0.50.

To determine significant group differences between YSR and BESS scores, independent group *t*-tests were conducted. Results are presented on Table 15 and indicate significant differences between the Not At-Risk and At-Risk group across all scales. Cohen's d effect sizes demonstrate overall moderate differences between groups.

Table 15
Group Differences in YSR and BESS Scores

Measure	Classification (n)	M(SD)	t	p	ES
YSR	Not at-risk (61)	11.62 (8.30)	1.97	.051	.38
Internalizing	At-risk (45)	15.22 (10.48)			
YSR	Not at-risk (61)	8.97 (7.26)	2.81	.006	.55
Externalizing	At-risk (45)	12.98 (7.27)			
YSR	Not at-risk (61)	39.57 (24.07)	2.71	.008	.53
Total	At-risk (45)	53.40 (28.32)			
BESS	Not at-risk (61)	21.20 (12.39)	2.22	.028	.43
	At-risk (45)	27.07 (14.73)			

Cross-informant reliability. Data from the matched teacher and student reports on the SEARS-SF were used to answer the research question - *"What is the degree of similarity*

between student and teacher ratings of a given student using respective versions of a strength-

based rating scale (SEARS-Short Forms; Merrell, 2011) ?" Cross-informant reliability between

the total score on the student version and the teacher version of the SEARS-SF were analyzed

using Pearson's product-moment correlations. Pearson product-moment correlations for the

teacher (*M* = 20.11, *SD* = 7.93) and student (M = 22.30, SD = 5.34) forms were statistically

significant *p* = .001, *r* = .33. This finding indicates a moderate to weak correlation of teacher

and student ratings on student's social-emotional assets and strengths, and that there is

considerable variance across raters.

      Discriminant validity. Teacher and student data were used to answer the research

question -"What *is the degree of difference between strength-based rating scale scores (SEARS-*

Short Forms; Merrell, 2011) and primarily problem-based rating scale scores (BESS; Kamphaus &

*Reynolds, 2007; YSR; Achenbach & Rescorla, 2001)?" S*eparate correlation coefficients were

calculated to analyze the relationship between (1) the teacher version of the SEARS-SF and the

BESS, and (2) the student version of the SEARS-SF and the YSR composite scales. Descriptive

information is depicted in Table 16.

Table 16
Descriptive Statistics of Student and Teacher Scores on the YSR and BESS

Scale	M	SD
BESS	23.69	13.68
YSR		
Internalizing problems	13.15	9.41
Externalizing problems	10.67	7.50
Total problems	45.44	26.73

Pearson product-moment correlations are presented in Table 17 and indicate moderate negative correlations between student and teacher ratings of students' social and emotional strengths and assets and ratings of student social, emotional, and behavioral problems ($p <$ .001).

Table 17
Pearson Product-Moment Correlations for the BESS Teacher Form, YSR, and SEARS-SF Scores

Scale	SEARS-SF student version	SEARS-SF teacher version
YSR		
Internalizing composite	-.48*	
Externalizing composite	-.50*	
Total score	-.55*	
BESS		-.70*

*$p < .001$

Group Comparisons. A series of analyses were conducted to answer the research question, "Are there significant group differences between the At-Risk and Not-At Risk group according to academic functioning, number of absences, number of office discipline referrals, gender, and disability status?" Results are presented by dependent variable.

Student gender. A two-variable chi-square test was used to evaluate differences in the proportion of male and female students between the Not At-Risk and At-Risk group (see Table 18). The difference in the number of male and female students in the Not At-Risk and At-Risk group was not significant, χ^2 ($N = 106, 1) = 3.58$, $p = .058$.

Table 18
Percentages and Counts of Female and Male Students by Risk Classification Group

Group	Not at-risk (n = 61)	At-risk (n = 45)
Female (n = 49)	54.1% (33)	35.6% (16)
Male (n = 57)	45.9% (28)	64.4% (29)

Disability status. A two-variable chi-square test was used to evaluate if there was a difference in the proportion of students receiving special education services (i.e., students with an IEP) between the Not At-Risk and At-Risk group (see Table 19). Results revealed no significant difference in the proportion of students receiving special education services in the Not At-Risk and At-Risk group, χ^2 (N = 106, 1) = 0.18, p = .675.

Table 19
Percentages and Counts of Students Receiving Special Education Services by Risk Classification Group

Group	Not At-Risk (n = 61)	At-Risk (n = 45)
IEP (n = 17)	14.8% (9)	17.8% (8)
No IEP (n = 89)	85.2% (52)	82.2% (37)

Office discipline referrals. A two-variable chi-square test was used to evaluate if there was a difference in the proportion of students identified as being at At-Risk or Not At-Risk based on ODRs. Chi-square analyses were conducted utilizing several different groupings of major and minor ODRs that students received over the course of the school year (Horner, Sugai, Todd, & Lewis-Palmer, 2005).   Combining major and minor ODRs, three groups were created based on the following cut points: six or more ODRs = Tier III, two to five ODRs – Tier II, and zero to one ODRs = Tier I. Groupings based on Major ODRs were also created: three or more

major ODRs = Tier III, one to two major ODRs = Tier II, and zero major ODRs = Tier I. Regardless

of how students were grouped, there was no significant difference in the proportion of

students identified at-risk, (p > .10; see Tables 20 to 23).

Table 20
Percentages and Counts of Students by Risk Classification Group and Major ODRs based on Two Tiers

Group	Not At-Risk (n = 61)	At-Risk (n = 45)
> 1 Major ODRs (n = 35)	27.9% (17)	40.0% (18)
0 Major ODRs (n = 71)	72.1% (44)	60.0% (27)

p = .189.

Table 21
Percentages and Counts of Students by Risk Classification Group and Major ODRs based on Three Tiers

Group	Not At-Risk (n = 61)	At-Risk (n = 45)
> 3 Major ODRs (n =15)	13.1% (8)	15.6% (7)
1-2 Major ODRs (n =20)	14.8% (9)	24.4% (11)
0 Major ODRs (n = 71)	72.1% (44)	60.0% (27)

p = .374.

Table 22
Percentages and Counts of Students by Risk Classification Group and Major and Minor ODRs based on Two Tiers

Group	Not At-Risk (n = 61)	At-Risk (n = 45)
> 6 Minor and major ODRs (n = 29)	23.0% (14)	33.3% (15)
0-5 Minor and major ODRs (n = 77)	77.0% (47)	66.7% (30)

p = .236.

Table 23

Percentages and Counts of Students by Risk Classification Group and Major ODRs based on Three Tiers

Group	Not At-Risk (*n* = 61)	At-Risk (*n* = 45)
> 3 Minor and major ODRs (*n* =13)	13.1% (8)	11.1% (5)
1-2 Minor and major ODRs (*n* = 16)	9.8% (6)	22.2% (10)
0 Minor and major ODRs (*n* = 77)	77.0% (47)	66.7% (30)

p = .212.

Academic functioning. Reading and math OAKS scores, Oral Reading Fluency (ORF), and cumulative Grade Point Average (GPA) for At-Risk and Not At-Risk students were evaluated with independent groups *t*-tests. Results are presented on Table 24 and the only significant difference between the Not At-Risk and At-Risk group was indicated for cumulative GPA.

Table 24

Differences in Academic Indicators by Risk Classification Group

	Gender (n)	M (SD)	t	p	ES
Math	Not at-risk (60)	231.30 (10.72)	0.70	.485	.14
	At-risk (44)	229.89 (9.32)			
Reading	Not at-risk (60)	231.02 (7.85)	0.11	.915	.02
	At-risk (44)	230.84 (8.85)			
ORF	Not at-risk (60)	161.45 (42.96)	0.94	.349	.19
	At-risk (45)	153.82(38.46)			
Cumulative GPA	Not at-risk (61)	3.28 (0.63)	3.22	.002	.63
	At-risk (45)	2.86 (0.71)			

Note. All distributions of scores were roughly symmetrical for the At-Risk and Not At-Risk groups. One severe outlier based on OAKS math scores was identified in the At-Risk group. The distribution of OAKS reading scores revealed one severe outlier in the Not At-Risk group and six severe outliers in the At-Risk group.   One severe outlier based on cumulative GPA was identified in the Not At-Risk group. All outliers were retained.

Absences. The number of absences for At-Risk and Not At-Risk students was evaluated using a Mann-Whitney test, as distributions for the At-Risk (skew = 1.48) and Not At-Risk (skew = 1.25) groups had a strong positive skew. Results revealed that the total number of absences of Not At-Risk students (*Mdn* = 5.0) did not differ significantly from At-Risk students (*Mdn* = 5.5), *U* = 1338.00, *z* = -0.22, *p* = .825.

Discussion

Overall, results revealed that the odds of a child being identified as at-risk using the strength-based approach under investigation was positively related to well-established measures of social-behavioral problems. Students identified as being At-Risk differed from Not At-Risk students on grade point average and teacher and self-report measures of social, behavioral, and emotional functioning (YSR; Achenbach & Rescorla, 2001; BESS; Kamphaus & Reynolds, 2007). The At-Risk and Not At-Risk groups did not significantly differ on disability status, ODRs, gender, absences, and standardized measures of academic performance. Cross-informant reliability and discriminant validity were analyzed, and results were consistent with previous research providing additional support that the SEARS-SF meets these standards of validity and reliability (AERA, APA, & NCME, 1999).

Discussion of the Findings

Research Questions 1 and 2:

1. Do reported levels of internalizing and externalizing symptoms on the Youth Self-Report (YRS; Achenbach & Rescorla, 2001) effectively and accurately discriminate between students identified as not at risk and at-risk based on the proposed strength-based approach to mental health screening?

2. Do scores from the Behavioral and Emotional Screening System Teacher Version

(BESS; Kamphaus & Reynolds, 2007) effectively and accurately discriminate

between students identified as not at risk and at-risk based on the proposed

strength-based approach to mental health screening?

Measures of internalizing and externalizing behaviors (Achenbach & Rescorla, 2001;

Kamphaus & Reynolds, 2007) discriminated between students identified as Not At-Risk and At-

Risk using the proposed strength-based approach to mental health screening. BESS and YSR

(Internalizing, Externalizing, and Total) scores were significant or nearly significant predictors of

whether students were grouped into the At-Risk or Not At-Risk group. The overall classification

accuracy for the predictor variables in the logistic regression analyses ranged from 60.4 to 64.2,

which is an improvement to the chance level, yet low. These results were likely impacted by

the 6-month delay between Phase I and Phase II. Furthermore, students were classified as At-

Risk and Not At-Risk using a strength-based measure, while predictor variables were broad-

band measures of primarily problem behaviors. Given that positive and negative indicators of

mental health are not necessarily at the opposite ends of the same continuum (Greenspoon &

Saklofske, 2001; Suldo & Shaffer, 2008), one might expect predictive validity to be moderate at

best.

More research is needed to improve classification accuracy of this approach to mental

health screening and to determine if the differences between the At-Risk and Not At-Risk group

are meaningful in relation to students' actual social and emotional functioning and mental

wellness. These results should be considered a conservative measure of predictive validity

given the approximately six month delay between the strength-based screening procedure

used to group students and the administration of the criterion measures. In fact, Pearson

product-moment correlations between SEARS-SF scores obtained during Phase I and Phase II of

the study were only moderately correlated (r = .56, p < .01) indicating that students self-reports

of social emotional strengths and assets changed from the time they completed the SEARS-SF in

the fall to when they completed the SEARS in the spring. Thus, although findings of the logistic

regression analyses were overall significant, a stronger relationship may have been identified

had all of the measures been administered in the fall.

Of the four predictor variables, the Internalizing scale score on the YSR scores was the

least significant predictor of risk classification when compared to the other three predictor

variables (YSR Externalizing, YSR Total, and BESS scores). Early identification of internalizing

problems is particularly important as these types of problems can cause severe impairment, yet

often go undetected unless children and adolescents are directly asked about their internalizing

behaviors (Pagano, et al., 2000). Students with internalizing problems also receive fewer

services than those students with externalizing symptoms (Bradshaw, et al., 2008), and many of

the behaviors associated with disorders such as anxiety and depression are very difficult to

observe, particularly at the subclinical level. Furthermore, students coping with internalizing

problems are generally not disruptive and, therefore, potentially less likely to be noticed by

their teachers. Teachers may also be less motivated to intervene if a student is not disruptive.

Finally, while schools have considerable amounts of student data associated with externalizing

behavior problems (e.g., ODRs) available, school-based mental health promotion requires a

strategic and systematic approach of assessing students' social and emotional functioning, and

behavior rating scales provide a way for schools to identify students at-risk for internalizing

problems.

Research questions 3 and 4.

3. What is the degree of similarity between student and teacher ratings of a given

student using respective versions of a strength-based rating scale (SEARS-Short

Forms; Merrell, 2011)?

4. What is the degree of difference between strength-based rating scale scores

(SEARS-Short Forms; Merrell, 2011) and primarily problem-based rating scale

scores (BESS; Kamphaus & Reynolds, 2007; YSR; Achenbach & Rescorla, 2001)?

Cross-informant reliability of the SEARS-SF indicated relatively weak correlations

between teacher reports and child self-reports based on Pearson-product moment correlations

(r = .32). Discriminant validity between the SEARS-SF (student, self-report) and YSR (student,

self-report) and BESS (teacher report) indicated moderate negative correlations (r = -.48 to -

.70). These results are consistent with other research validating the SEARS assessment system

(Merrell, 2011, Merrell, Cohn, & Tom, 2011; Nese et al., in press) and meet standards for

reliability and validity (AERA, APA, NCME, 1999).

Research question 5.

5. Are there significant group differences between the At-Risk and Not-At Risk

group according to academic functioning, number of absences, number of

office discipline referrals (ODR), gender, and disability status?

The At-Risk and Not At-Risk groups did not differ on disability status, ODRs, gender,

absences, and standardized measures of academic performance. These results are somewhat

surprising as these outcomes have been associated with students' social and emotional skills and assets (Fleming et al., 2005; Zins, et al. 2004; Burns, et al., 2004; Knoff, 2004). However, these results may have been confounded for several reasons that will be discussed on the limitations section. More research is needed to determine the relationship between strength-based screening and student outcomes.

As noted previously, significant group differences between the At-Risk and Not At-Risk group were found on the YSR composite scales and the BESS teacher report scores. Also, the At-Risk group had significantly lower grade point averages at the end of the school year compared to students in the Not At-Risk group. This finding is particularly interesting in light of SEARS-SF self-report scores from the start of the school year being used to group students (At-Risk and Not At-Risk groups) and that significant differences between these two groups were based on grade point averages at the end of the academic school year. Moreover, this was the case despite 26 participants who had been classified At-Risk in the fall, no longer being classified as At-Risk in the spring (based on a second administration of the SEARS-SF). This finding may be explained by grades accounting for not only academic performance, but also social and emotional behaviors (e.g., interpersonal relationships, problem solving and self-regulation skills) related to the constructs assessed by the SEARS-SF. Replication of this finding is needed before drawing any conclusions; however, the link between the strength-based approach to mental health screening used in this study and grade point averages at the end of the school year may have particular relevance to schools focused on coordinated academic and mental health supports as grade point average and problem behavior for this age group have been shown to be early indicators for high school dropout (Suh & Suh, 2007; Tobin & Sugai,

1999) and speaks to the importance of systematically monitoring and supporting student behavior and academic performance.

Limitations

When evaluating the findings of this study, it is important to consider possible confounding influences on the results. First, no procedural integrity data were collected to measure the consistency of the administration of the screener during Phase I of the study. Teachers received a brief training on how to administer the screener but the extent to which they adhered to the protocol is unknown. Also during Phase I of the study, researchers scored and analyzed the data, all of which were obtained via pen-and-paper measures. This procedure likely is overly cumbersome for a school to implement independently. Thus, there is a need for more efficient data collection and analysis, for example by computerized test administration and analysis.

The limited sample size resulted in underrepresentation of certain demographic groups and regional sample bias. Although a sample of 106 students was deemed adequate, a larger sample would have increased statistical power. A larger sample would have also made it feasible to recruit three groups (Tier I, Tier II, and Tier III) and allow for comparison between students in Tier I and Tier III. The sample of students was also relatively homogeneous and nested within one district. A large portion of the sample was from one school in the Pacific Northwest region on the United States, predominantly White/Caucasian, which has been implementing a three-tiered model of behavioral and academic supports. This school also has a history of collaboration with researchers from the local university. These nesting effects could further decrease the generalizability of the results. Finally, the approach to mental health

screening in this study involved using cut scores based on the national norming sample and it is unrealistic to assume that single cut scores can adequately detect risk for all subgroups of students represented in other schools.

Another limitation to be considered is the time delay Phase I and Phase II of the study. Based on the second administration of the SEARS-SF in the spring, which was administered at the same time as the criterion measures, 26 fewer students would have been identified At-Risk compared to the beginning of the school year. This suggests that, over the course of the year, this sample had fewer students whose scores fell within the At-Risk range and that results may have been confounded by changes in behavioral, social, and emotional behavior over time (Merrell, 2008). In addition, because the participating schools were all implementing SWPBS with good integrity, at minimum all students were receiving Tier I interventions aimed at supporting prosocial behavior during those six months. These schools were implementing Tiers II and III supports within Intensive Positive Behavior Support and used school data to select students in need of further intervention so it is likely that at least some of those students identified via the SEARS-SF in the fall but not the spring received intervention in the interim that reduced their risk status.

Although Phase I of the study was completed quickly and seamlessly in part because active consent from parents was not necessary, the process of gathering active consent resulted in a long delay between Phase I and Phase II. The need for active consent to proceed with the second phase of the study also introduced the potential for additional sampling bias by resampling from one of the schools and parents of students identified as Not At-Risk were more likely to return their consent forms.

Future Directions and Implications for Practice

To our knowledge, this was one of if not the first study to assess the validity of a solely strength-based *approach* to mental health screening. Behavior rating scales and self-assessments have been suggested for universal screening and progress monitoring as part of multi-tiered mental health intervention models (Levitt, et al., 2007; Merrell & Gueldner, 2010). Future research should determine the utility of short form versions of strength-based rating scales (e.g., Devereux Student Strengths Assessment-mini; Naglieri, LeBuffe, & Shapiro, 2011; SEARS, Merrell, 2011) for the purpose of universal screening.

Replication of this study with a higher level of methodological rigor is needed to further investigate if this strength-based approach to mental health screening is indeed sufficient or if it can be used in conjunction with other readily available student data to identify students in need of additional supports. Additional research is also needed to investigate if this type of strength-based approach is acceptable and feasible for parents, students, and educators, and to identify proximal and distal outcomes related to using a strength-based approach to mental health screening. Further scale development (i.e., content selection) and psychometric support for strength-based measures designed specifically for the purpose of universal screening within a three-tiered model may lead to better measures for this purpose. Finally, this line of research also needs to consider how universal mental health screening can be integrated into a system of supports while taking into account available resources and other contextual variables.

Although this study investigated the validity of a solely strength-based measure as a mental health screener, the criterion measures used to validate this approach were primarily deficit-based. As such, criterion measures did not assess the social and emotional

competencies of the participating students or the school environment. It seemed logical to

have the first step of validating a solely strength-based approach to mental health screening

align with current, well-established standards of behavioral, social, and emotional assessment

(Merrell, 2008). However, the premise of using a solely strength-based approach to mental

health screening is based on the gains in prevention science that have come from a perspective

focused on systematically building competence rather than correcting weakness (Seligman &

Csikszentmihalyi, 2000), and as such, research is needed to validate a strength-based approach

to mental health screening in relation to indicators of mental wellness.

Future research may also focus on the psychometric properties of strength-based

assessment systems such as the SEARS including longitudinal test re-retest reliability analyses,

cross informant analyses with parents, teachers, and students, convergent validity with other

strength-based measures, and use as an intervention outcome measure. Research is needed to

investigate the relationship between teacher, parent, and self-report ratings, as well as direct

observations of student social and emotional skills. Studies utilizing receiver-operating

characteristic curve analysis are needed to evaluate the classification accuracy of the SEARS-SF

and identify optimal cut scores. Differential Item Functioning and Item Response Theory can be

used to determine if responses on strength-based assessment measures vary between different

demographic groups. Lastly, additional research is needed to investigate the sensitivity of the

SEARS-SF to short-term changes in student social and emotional behavior.

To determine the validity of using a strength-based measure for the purpose of mental

health screening across all grades, extensions of this type of research to preschool, elementary,

and high schools settings is needed. In younger grades, a similar screening process would likely

involve teacher ratings (Flanagan, Bierman, & Kam, 2003). Although a multimethod, multisource, multisetting assessment is the standard for behavioral, social, and emotional assessment practices (Merrell, 2008), additional research is needed to determine which source or sources (i.e., teacher, caregiver, student) are optimal informants for universal screening measures involving ratings of student behavior (Cook, Volpe, & Livanis, 2010). Cross-informant agreement on student's social emotional functioning is generally low (Achenbach, et al. 1987), and more than one rater may be needed to accurately identify students at-risk of developing mental health problems.

A multiple gating approach utilizes data from multiple assessments, sources, and settings in order to identify at-risk youth (Loeber, Dishion, & Patterson, 1984; Sprague, Walker, Stieber, Simonsen, & Nishioka, 2001; Walker & Severson, 1991). A multiple gating approach is also designed to minimize false positives and negatives. For example the Systematic Screening for Behavior Disorders (SSBD; Walker & Severson, 1992) has three gates that include: (1) teacher nominations, (2) teacher completed behavior rating scales, and (3) direct observations of student behavior on the playground or in the classroom. A multiple gating procedure begins with a cost-effective and relatively easy to administer screening procedure. Those students that pass through the first gate (i.e., are identified as being at-risk using the screening procedure) are assessed further. Within a multiple gating approach, the SEARS-SF could be considered the first gate within this type of procedure. Then those students identified at-risk using the SEARS-SF would be assessed further using the teacher version of the SEARS-SF or a broad-band measure of behavior (Achenbach & Rescorla, 2001; Reynolds & Kamphaus, 2004). Universal screening of social emotional behaviors may be best accomplished by schools

assessing both risk and protective factors.

The problem-solving process requires identification of problems as well as information about the students' skills and the contingencies within the school environment maintaining student behavior. A combined approach that includes a strength-based measure may better capture not only risk factors, but protective factors such as peer friendships, engagement in productive activities, and teacher-student relationships impacting the student body as a whole and the resources of the school, integral to the development of Tier I supports (Doll & Cummings, 2008). Schools not only need to know what empirically supported screeners are available, but what combination of data provides a comprehensive, yet efficient, and cost effective means of assessing student performance and Tier I supports across a range of academic and social behavioral domains. Similarly, some of the data collected as part of this study (ODRs, attendance, grade point average) are also predictors, and could be combined to monitor the effectiveness of universal supports and overall 'health' of the school and district.

Universal school-based mental health screening is still in its early stages and only one component of a service delivery process. As such, this line of research lends itself to a collaborative effort between researchers, educators, families, community based service providers, and policy makers as school-based mental health programming needs to be aligned with values and resources of the community (Weist, et al., 2007) and ultimately lead to improved and valued outcomes. Careful consideration needs to be given to factors such as how to ensure parent and community involvement, protect students and family rights (e.g., consent and confidentiality), and proactively address legal and ethical considerations (Levitt, et al., 2007). Another common concern about universal mental health screening is that more

students will be identified as requiring additional supports than a school currently has readily available. To address this concern, researchers should work closely with schools to identify resources within the school and community at the onset of a study. Finally, mental screening measures and procedures need to be feasible for schools to administer and include considerations such as associated costs and methods used to manage and interpret data (Glover & Albers, 2007). The SEARS-SF takes only a few minutes to complete and measures student social and emotional skills and assets that are aligned with mental health programming. Despite these benefits, the cost and data management could pose potential barriers to a schools adoption of this measure as a mental health screener.

Conclusion

In conclusion, this is one of, if not the first study to investigate a solely strength-based approach to mental health screening based on a three-tiered model of service delivery. The findings of this study were mixed, but did indicate that the approach used to identify students At-Risk and Not At-Risk did result in significant differences between the groups on well-established measures of social, emotional, and behavioral functioning (Achenbach & Rescorla, 2001; Kamphaus & Reynolds, 2007) six months later and students cumulative grade point average at the end of the school year. Significant differences were not found on measures of academic skills, absences, special education status, office discipline referrals, or gender.

Advocates for the integration of school-based mental health services propose that, 'data should document the collective mental health needs of students in the district, research should guide the match between interventions and those needs, and evaluation should document the impact of the program of mental health services on children's psychological wellness and

success,' (Doll & Cummings, 2008, p. 1334). This cycle of assessment, planning, intervention,

and evaluation is aligned with three-tiered models emphasizing data-based decision making

and linking assessment to intervention (Chard, Harn, Sugai, Horner, Simmons, & Kame'enui,

2008). To this end, strength-based assessment is particularly relevant to treatment validity and

how assessment measures guide interventions aimed at improving social emotional functioning

and mental wellness. The strength-based approach used in this study shows promise as

educators strive to 'get ahead of the problem' by measuring social emotional competencies

associated with mental wellbeing and resilience, rather than the absence of mental illness or

problems.

References

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,* 213-232.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-aged forms and profiles.* Burlington: University of Vermont Research Center for Children, Youth and Families.

Albers, C. A., Glover, T. A., & Kratochwill, T. R. (2007). Where are we and where do we go now? Universal screening for enhanced educational and mental health outcomes. *Journal of School Psychology*, *45*, 257-263.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing.* Washington, DC: Author.

Adelman, H. S., & Taylor, L. (2000). Moving prevention from the fringes into the fabric of school improvement. *Journal of Educational and Psychological Consultation, 11,* 7-36.

Adelman, H. S., & Taylor, L. (2008). School-wide approaches to addressing barriers to learning and teaching. In B. Doll & J. A. Cummings (Eds.), *Transforming school mental health services: Population-based approaches to promoting the competency and wellness of children* (pp. 277-306). Thousand Oaks, CA: Corwin Press/National Association of School Psychologists.

Beaver, B. R. (2008). A positive approach to children's internalizing problems. *Professional Psychology: Research and Practice, 39,* 129 – 136.

Bowerman, B. L., & O'Connell, R. T. (1990). *Linear statistical models: An applied approach* (2nd ed.). Belmont, CA: Duxbury.

Burns, E. J., Walrath, C., Glass-Siegel, M., & Weist, M. D., (2004). School-based mental health services in Baltimore: Association with school climate and special education referrals. *Behavior Modification, 28,* 491-512.

Bradshaw, C. P., Buckley, J. A., Ialongo, N. S. (2008). School-based service delivery among urban children with early onset educational and mental health problems: The squeaky wheel phenomenon. *School Psychology Quarterly, 23,* 169-186.

Chard, D.J., Harn, B.A., Sugai, G., Horner, R.H., Simmons, D.C., & Kame'enui, E.J. (2008). Core features of multi-tiered systems of reading and behavioral support. In C.R. Greenwood,

T.R. Kratochwill, & M. Clements (Eds.), *Schoolwide prevention models: Lessons learned in elementary schools* (pp. 31-59). The Guilford Press: New York.

Chen, G. & Astebro, T. (2003). How to deal with missing categorical data: Test of a simple Bayesian method. *Organizational Research Methods, 6,* 309-327.

Christ, T. J., & Silberglitt, B. (2007). Estimates of the standard error of measurement for curriculum-based measures of oral reading fluency. *School Psychology Review, 36,* 130-146.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155-159.

Cohen, R., Kincaid, Childs, K. (2007). School-wide positive behavior support implementation: development and validation of the benchmark of quality. *Journal of Positive Behavior Intervention, 9*, 203-213.

Cook, C. R., Volpe, R. J., & Livanis, A. (2010). Constructing a roadmap for future universal screening research beyond academics. A*ssessment for Effective Intervention, 35,* 197-205.

Costello, E. J., Mustillo, S., Erkanli, A., Keeler, G, & Angold, A. (2003). Prevalence and development of psychiatric disorders in childhood and adolescence. *Archives of General Psychiatry, 60,* 837-844.

Doll, B. (1996). Prevalence of psychiatric disorders in children and youth: An agenda for advocacy by school psychology. *School Psychology Quarterly, 11,* 20-46.

Doll, B., & Cummings, J. A. (2008). Best practices in population-based school mental health services. In A. Thomas & Grimes (Eds.), *Best practices in school psychology V* (pp. 1333-1346). Bethesda, MD: National Association of School Psychologists.

Doll, B. & Cummings, J. A. (2008). Why population-based services are essential for school mental health, and how to make them happen. In B. Doll & J. A. Cummings (Eds.), *Transforming school mental health services: Population-based approaches to promoting the competency and wellness of children* (pp. 1-20). Thousand Oaks, CA: Corwin Press/National Association of School Psychologists.

Dowdy, E., Furlong, M., Eklund, K., Saeki, E., & Ritchey, K. (2010). Screening for mental health and wellness: Current school based practice and emerging possibilities. In B. Doll, W. Pfohl, & J. Yoon (Eds.). *Handbook of youth prevention science.* New York: Routledge.

Drummond, T. (1994). *The Student Risk Screening Scale (SRSS).* Grants Pass, OR: Josephine County Mental Health Program.

Epstein, M. H. (1999). The development and validation of a scale to assess the emotional and behavioral strengths of children and adolescents. *Remedial and Special Education, 20*, 258-263.

Epstein, M. H., & Sharma, H. M. (1998). *Behavioral and Emotional Rating Scale: A strength-based approach to assessment.* Austin, TX: PRO-ED.

Flanagan, K.S., Bierman, K.L., Kam, C., & CPPRG (2003) Identifying at-risk children at school entry: The usefulness of multibehavioral problem profiles. *Journal of Clinical Child andAdolescent Psychology, 32*, 396-407.

Farmer, E. M. Z., & Farmer, T. W. (1999). The roles of schools in the outcomes for youth: Implications for children's mental health services research. *Child and Family Studies, 8,* 377-396.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Fleming, C. B., Haggerty, K. P., Brown, E. C., Catalano, R. F., Harachi, T. W., Mazza, J. J., & Gruman, D. H. (2005). Do social and behavioral characteristics targeted by preventative interventions predict standardized test scores and grades? *Journal of School Health, 75,* 342-349.

Freudenberg, N., & Ruglis J. (2007). Reframing school dropout as a public health issue. *Preventing Chronic Disease, 4.* Retrieved from http://www.cdc.gov/pcd/issues/2007/oct/07_0063.htm.

Glover, T. A., & Albers, C. A. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology, 45*, 117-135.

Greenberg, M.T., Domitrovich, C., & Bumbarger, B., (2001). The prevention of mental disorders in school-aged children: Current state of the field. *Prevention and Treatment, 4,* 1-172**.**

Greenspoon, P. J., & Saklofske, D. H. (2001). Toward an integration of subjective well-being and psychopathology. *Social Psychology International, 18,* 229-243.

Greenwood, C. R., Horner, R. H. & Kratochwill, T. R. (2004). Introduction. In C. R. Greenwood, T. R. Kratochwill, & M. Clements (Eds.), *Schoolwide Prevention Models: Lessons Learned in Elementary Schools* (p. 3). New York: Guilford Press.

Gresham, F. M. (2005). Responsive to intervention: An alternative approach to identifying students as emotionally disturbed. *Education and Treatment of Children, 28,* 328-344.

Gresham, F. M., & Elliott, S. N. (1990). *Social Skills Rating System.* Circle Pines, MN: American Guidance Service.

Horner, R. H., Sugai, G., Todd, A. W., & Lewis-Palmer, T. (2005). School-wide positive behavior support. In L. Bambara & L. Kern (Eds.), *Individualized supports for students with problems behaviors: Designing positive behavior plans* (pp. 359-390). New York: Guilford Press.

Howe, K. B., & Shinn, M. M. (2002). *Standard reading assessment passages (RAPs) for use in general outcome measurement: A manual describing development and technical features.* Eden Prairie, MN: Edformation.

Huebner, E. S. (1991). Initial development of the Students' Life Satisfaction Scale. *School Psychology International, 12,* 231-240.

Hu, T. (2004). *An international review of the economic costs of mental illness* (Working Paper 31). Washington DC: A joint project of the World Bank, the Fogerty International Center of the National Institutes of Health, The Bill and Linda Gates Foundation, and the World Health Organization.

Jimerson, S. R., Sharkey, J. D., Nyborg, V. & Furlong, M.J. (2004). Strength-based assessment and school psychology: A summary and synthesis. *The California School Psychologist, 9,* 9-19.

Kamphaus, R. W., & Reynolds, C. R. (2007). *BASC-2 Behavioral and Emotional Screening System Manual.* Circle Pines, MN: Pearson.

Kincaid, D., Childs, K., & George, H. (2005). *School-wide benchmarks of quality*. Unpublished instrument, University of South Florida, Tampa.

Knoff, H. M. (2004). Inside Project ACHIEVE: A comprehensive, research-proven whole school improvement process focused on student academic and behavioral outcomes. In K. Robinson (Ed.), *Advances in school-based mental health: Best practices and program models* (pp. 19-1 to 19-28). Kingston, NJ: Civic Research Institute, Inc.

Levitt, J. M., Saka, N., Romanelli, L. H., & Hoagwood, K. (2007). Early identification of mental health problems in schools: The status of instrumentation. *Journal of School Psychology, 45,* 163-191.

Levitt, V. H., & Merrell, K. W. (2009). Linking assessment to intervention for internalizing problems of children and adolescents. *School Psychology Forum: Research in Practice, 3,* 13-26.

Loeber, R., Dishion, T.J., & Patterson, G.R. (1984). Multiple gating: A multi-stage assessment procedure for identifying youths at risk for delinquency. *Journal of Research in Crime and Delinquency, 21*, 7–32.

May, S., Ard, W., Todd, A. W., Horner, R. H., Glasgow, A., Sugai, G., et al. (2006). *School-wide information system*. Eugene: Educational and Community Supports, University of Oregon.

Merrell, K.W. (2008*). Behavioral, social, and emotional assessment of children and adolescents* (3rd ed.). New Jersey: Lawrence Erlbaum Associates, Inc.

Merrell, K. W. (2011). *Social and emotional assets and resilience scales (SEARS).* Lutz, FL: Psychological Assessment Resources.

Merrell, K. W. (2002). *School Social Behavior Scales, Second Edition.* Baltimore: Paul H. Brookes Publishing.

Merrell, K. M., Cohn, B. & Tom, K. M. (2011). Development and validation of a teacher report measure for assessing social-emotional strengths of children and adolescents. *School Psychology Review, 40,* 226-241.

Merrell, K. W., Ervin, R. A. & Gimpel, G. A. (2006*). School psychology for the 21st century: Foundations and practices.* New York: Guilford.

Merrell, K. W., & Gueldner, B. A. (2010). *Social and emotional learning in the classroom: Promoting mental health and academic success.* New York: Guilford Press.

Merrell, K.W., & Walters, A. (1998). *Internalizing Symptoms Scale for Children.* Austin, TX: PROED**.**

Myers, R. (1990). *Classical and modern regression with applications* (2nd ed.). Boston: MA: Duxbury.

Naglieri, J. A., LeBuffe, P., & Shapiro, V. B. (2011). Universal screening for social-emotional competencies: A study of the reliability and validity of the DESSA-MINI. *Psychology in the Schools, 48,* 660-671.

Nastasi, B. K. (2004). Meeting the challenges of the future: Integrating public health and public education for mental health promotion. *Journal of Educational and Psychological Consultation, 15,* 295-312.

Nese, R. N. T., Doerner, E., Romer, N., Kaye, N. C., Merrell, K. W., & Tom, K. M. (in press). Social emotional assets and resilience scales short form version: Development of a strength-based behavior rating scale system. *Journal for Educational Research Online.*

New Freedom Commission on Mental Health (2003). *Achieving the promise: Transforming mental health care in America.* Final Report DHHS Pub. Vol. SMA—3-3832. Rockville, MD: Author retrieved: November 15, 2009 from http://www.mentalhealthcommission.gov/reports/FinalReport/toc.html.

Oregon Department of Education. (2007). 2006–2007 *Technical Report Oregon's Statewide Assessment System Reliability and Validity, Volume 4.* Retrieved from http://www.ode.state.or.us/search/page/?=1305

Osher, D., Sprague, J., Weissberg, R. P., Alxelrod, J., Keenan, S., Kendziora, K., & Zins, J. E. (2008). A comprehensive approach to promoting social, emotional, and academic growth in contemporary schools. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 1263-1278). Bethesda, MD: National Association of School Psychologists.

Pagano, M.E., Cassidy, L.J., Little, M., Murphy, J.M., & Jellinek, M.S. (2000). Identifying psychosocial dysfunction in school-age children: The Pediatric Symptom Checklist as a self-report measure. *Psychology in the Schools, 37,* 91-106.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research* (3rd ed.). Orlando, L: Harcourt Brace & Co.

Renshaw, T. L., Eklund, K., Dowdy, E. Jimerson, S. R. Shelley, R. H., Earhart, J. & Camille, N. J. (2009). Examining the relationship between scores on the behavioral and emotional screening system and student academic, behavioral, and engagement outcomes: An investigation of concurrent validity in elementary school. *The California School Psychologist,* Retrieved from http://findarticles.com/p/articles/mi_7479/is_200901/ai_n52375546/?tag=content;col1

Ries Merikangas, K., He, J., Brody, D., Fisher, P. W., Bourdon, K., & Kortez, D. S. (2010). Prevalence and treatment of mental health disorders among US children in the 2001-2004 NHANES. *Pediatrics, 125,* 75-81.

Romer, N. (2012). Mental health screening within a tiered model: Investigation of a brief strength-based approach (Unpublished doctoral dissertation). University of Oregon, Eugene, OR.

Rones, M., & Hoagwood, K. (2000). School-based mental health services: A research review. *Clinical Child and Family Psychology Review, 3,* 223-241.

Seligman, M.E.P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist, 55,* 5-14.

Sprague, J., Walker, H. M., Steiber, S., Simonsen, B., & Nishioka, V. (2001). Exloring the relationship between school discipline referrals and delinquency. *Psychology in the Schools, 38, 197-*206.

SPSS Graduate Student Version for Windows (version 18.0) [Computer software]. 2009. Chicago, IL: SPSS Inc.

Suh, S. & Suh, J. (2007). Risk factors and levels of risk for high school dropouts. *Professional School Counseling, 10,* 297-306.

Suldo, S.M., & Shaffer, E.J. (2008). Looking beyond psychopathology: The dual-factor model of mental health in youth. *School Psychology Review, 37,* 52-68.

Tilly, D. V. (2008). Best practices in school psychology as a problem-solving enterprise. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 17-36). Bethesda, MD: National Association of School Psychologists.

Tobin, T.J., Sugai, G., & Colvin, G. (1996). Patterns in middle school discipline records. *Journal of Emotional and Behavioral Disorders 4*, 82-94.

Tobin, T. J., & Sugai, G. M., (1999). Using sixth-grade school records to predict school violence, chronic discipline problems, and high school outcomes. *Journal of Emotional and Behavioral Disorders, 7,* 40-53.

U.S. Department of Education Institute of Education Sciences. (n.d.). *National Center for Education Statistics, Common Core Data Search for Public School Districts.* Retrieved from http://nces.ed.gov/ccd/schooLSearch/index.asp

U.S. Department of Health and Human Services. (1999). *Mental Health: A Report of the Surgeon General.* Rockville, MD: U.S. Department of Health and Human Services.

Walker, H. M., & Severson, H. H. (1992). *Systematic screening for behavior disorders* (SSBD)(2nd Ed.). Longmont, CO: Sopris West.

Wagner, M., Kutash, K., Duchnowski, A. J., & Eptein, M. H. (2005). The special education elementary longitudinal study and the national longitudinal transition study: Study designs and implications for children and youth with emotional disturbance. *Journal of Emotional & Behavioral Disorders, 13,* 25-41.

Wasserman, G. A., Peters, J., Ko, S. J., Cocozza, J., Trupin, E., Angold, A., Cauffman, E, & Grisso, T. (2004). Mental health assessments on juvenile justice: Report on the Consensus Conference. *Journal of the American Academy of Child and Adolescent Psychiatry, 42,* 752-761.

Weist, M. D., Rubin, M., Moore, E., Adelsheim, S., & Worbel, G. (2007). Mental health
        screening in schools. *Journal of School Health, 77*, 53-58.

Zins, J.E., Bloodworth, M.R., Weissberg, R.P., & Walberg, H.J. (2004). The scientific base for
        linking social and emotional learning to school success. In J. E. Zins, R. P Weissberg, M. C.
        Wang, & H. J. Walberg (Eds.), *Building academic success on social and emotional
        learning* (pp. 3-22). New York: Teachers College Press.