

UC Riverside

UC Riverside Previously Published Works

Title

Comparing Informants for Mental Health Screening at the Preschool Level

Permalink

<https://escholarship.org/uc/item/0zs0b1r3>

Journal

School Psychology Review, 51(5)

ISSN

0279-6015

Authors

Moore, Stephanie A
Dowdy, Erin
Fleury, Isabelle
[et al.](#)

Publication Date

2022-09-03

DOI

10.1080/2372966x.2020.1841546

Peer reviewed



Published in final edited form as:

School Psych Rev. 2022 ; 51(5): 589–608. doi:10.1080/2372966x.2020.1841546.

Comparing Informants for Mental Health Screening at the Preschool Level

Stephanie A. Moore¹, Erin Dowdy², Isabelle Fleury², Christine DiStefano³, Fred W. Greer³

¹Graduate School of Education, University of California Riverside

²Counseling, Clinical, School Psychology, University of California Santa Barbara

³Education Studies, College of Education, University of South Carolina

Abstract

Universal screening for mental health in preschools provides the opportunity for early identification and early intervention, but guidance regarding which informants to use is needed. Preschoolers' ($N = 535$) parent and teacher reports across two screening forms were analyzed to determine similarities and discrepancies for classification results and screener scores. The analyses also examined if an additional rater provided incrementally valid information to the prediction of longitudinal kindergarten outcomes. Parents' and teachers' screening scores were significantly correlated across forms by rater and across raters. However, categorical classification results indicated that teachers were more likely than parents to rate preschoolers in at-risk ranges across forms. Finally, hierarchical regression analyses revealed teacher ratings were predictive of kindergarten social-emotional, cognitive, and academic outcomes, and that the addition of parent ratings did not significantly improve prediction of outcomes. Implications are discussed in the context of multiple raters within multiple-gating screening procedures.

Keywords

Screening; Informants; Mental Health; Preschool

In an effort to create safe and supportive schools for all students, multi-tiered systems of support (MTSS) are being considered and widely adopted in schools across the United States (Charlton et al., 2018). Within MTSS frameworks there is an emphasis on prevention, early intervention, and data-based decision-making to ensure that all students are provided with the appropriate levels of support for their academic and social-emotional development (Jimerson et al., 2016). A recent meta-analysis on the use of MTSS found strong evidence that MTSS is effective for improving social-emotional outcomes in preschoolers; MTSS programs targeting social-emotional development were more effective than programs targeting literacy or language outcomes (Shepley & Grisham-Brown, 2019). As such,

Correspondence concerning this article should be addressed to Stephanie A. Moore, Graduate School of Education, 900 University Ave, 1207 Sproul Hall, Riverside, CA 92521. stephanie.moore@ucr.edu.

The Version of Record of this manuscript has been published and is available in *School Psychology Review*, 2021, <https://www.tandfonline.com/10.1080/2372966X.2020.1841546>.

there is increased likelihood for the continued adoption of MTSS for promoting the social-emotional development of preschoolers.

Universal screening is an initial step designed to inform service delivery, with effective MTSS frameworks using screening data to make decisions about universal, targeted, and intensive supports (von der Embse et al., 2016). Universal screening for social-emotional risk is especially important at the preschool level as it provides a critical opportunity for early intervention when problems may be just starting to arise (Dowdy et al., 2013). In fact, research has demonstrated that without early intervention, behavioral and emotional problems that begin in early childhood are relatively stable and predictive of later negative social and educational outcomes (Lane et al., 2010). Considering the increasing numbers of children attending preschools (Barnett et al., 2013), coupled with knowledge that early prevention and intervention work is effective (Durlak et al., 2011), it is prudent to determine how to best conduct universal screenings which can lead to critical early intervention services in early childhood.

Current Identification Methods in Preschool

Methods available to identify children with social-emotional risk at the preschool level include referral methods (e.g., pediatric referrals, teacher nomination, parent referral to school personnel) and schoolwide universal screening (Kamphaus et al., 2014). School-wide universal screening that systematically gathers the same information on the behavioral and emotional functioning of each child has been recommended for use in preschools as part of MTSS (Kamphaus et al., 2014). This screening is most often accomplished via brief rating scales completed by key informants (Dowdy et al., 2019).

At the preschool level, both parents and teachers have been recommended as ideal informants as they are knowledgeable about the child's functioning (Smith, 2007). Teachers have been recognized as useful informants for a variety of childhood behavioral and emotional problems, particularly externalizing problems (Taylor et al., 2000). Additionally, teacher ratings have been found to be more reliable and similar across different teachers, when compared to parent ratings (Reynolds & Kamphaus, 1992). Within school-based screening, teachers are often considered to be an initial informant choice since they are the primary source of information regarding behavioral and emotional functioning within the school context.

There is also tremendous value in gathering parent ratings, with early childhood education standards (e.g., Friedman-Krauss, A. H., et al., 2020; National Association for the Education of Young Children, 2019) stressing the importance of acquiring parental input for children. Gathering information from parents within a screening program can provide information about a child's behavior over longer periods of time and across a variety of settings (McConaughy, 1993). While teachers are regarded as having excellent information on the externalizing functioning of their students, as well as a referent classroom group from which to compare students, parents are regarded as being particularly informative when reporting on internalizing symptoms in young children (Loeber et al., 1990; Smith, 2007). In addition, rating scales generally require four to six weeks of interaction with the child to be able

to accurately assess their social-emotional functioning; thus, parents are able to provide rating scale information earlier to trigger needed early intervention supports within an MTSS system (Dowdy et al., 2019).

Correspondence across Multiple Informants

Multiple informants may be useful with preschoolers, given that children in this age range are undergoing many developmental changes in a short period of time (Nagle, 2000) and their social-emotional behavior may manifest itself differently depending upon the setting (home vs. school). Multi-informant approaches including both parents and teachers can reflect variations across contexts when examining preschooler's behaviors (De Los Reyes et al., 2015). Indeed, the current standard of practice within child behavior assessment has been to collect rating scales across multiple informants (Frick et al., 2009). However, there are often time and resource constraints which may limit the ability to collect information from multiple informants within the universal screening context (Dowdy & Kim, 2012). Additionally, some research suggests that having more informants within a screening system is not always better (Biederman et al., 1990; Jones et al., 2002) and discrepancies among raters can complicate decision-making.

A robust literature base indicates that agreement between informants often varies, with inconsistent practice and research conclusions being drawn due to informant discrepancies (Achenbach et al., 1987). Differences are found depending on the nature of the problem being assessed and also the age of the child, with greater correspondence for externalizing problems and for younger children (Achenbach et al., 1987). A recent meta-analysis conducted by De Los Reyes and colleagues (2015) examined the multi-informant approach to assessment by analyzing 341 studies published between 1989 and 2014 which reported on cross-informant correspondence. Low correspondence was found between parents and teachers when rating on internalizing ($r = .21$) and externalizing ($r = .28$) problems, which is similar to findings of past meta-analytic multi-informant studies (Achenbach et al., 1987).

Perhaps due to the consistent findings of low agreement across raters, and the practicalities of needing an efficient screening system to inform prevention services within an MTSS system, recommendations regarding informants are often made based on practical realities. Little empirical research has systematically examined which informants may be most useful within a screening system (Dowdy et al., 2016). One recent preliminary examination into the added value of informants in screening for behavioral and emotional risk found there was little to no added value of including multiple informants for high school students (Dowdy et al., 2016).

Overall, additional research is needed to better understand which informant(s) to use within universal school-based screening efforts (Dowdy & Kim, 2012), as there is a significant dearth of research examining the incremental validity of multiple informant approaches for specific constructs and across different development time periods (De Los Reyes et al., 2015). Specifically, there is a need to examine if multi-informant approaches consisting of parent and teacher reports provide incrementally valid information relative to the use of any one informant's report (De Los Reyes et al., 2015).

Examining the incremental validity of multiple informants is aligned with a multiple-gating approach to screening. In brief, multiple-gating procedures involve providing additional assessments, or gates, to successively narrow down the population of interest; the desired outcome is that by the final assessment gate there is a strong likelihood that the students identified as being at risk are, in fact, true positive cases (Walker et al., 2014). Research supports multiple-gating approaches to be cost-effective and diagnostically accurate (Hill et al., 2004), and there are several multiple-gating screening procedures available for use at the preschool level (e.g., Kettler et al., 2017). Generally, an initial first gate screening assessment is provided to all students universally, and then the smaller group of students who are screened to be at risk are provided an additional second gate assessment. Although some tools use both teachers and parents as informants at the first gate, others often rely on a single informant for an initial screening gate (Feeney-Kettler et al., 2010). The second gate assessment may be: a) a more comprehensive assessment (i.e., full behavioral rating scale) via the same informant as the one used in the first gate, or b) an additional screening tool completed by a different informant (Dowdy et al., 2016). While either approach for a second gate assessment may be beneficial, this study will examine the use of teachers and parents as different informants within a multiple-gating framework.

Current Study

This study was designed to examine the use of informants at the preschool level when universally screening for behavioral and emotional risk. Specifically, we were interested in determining the similarities and discrepancies for classification results and screener scores from two different informants viewing the same child. Given that universal screening measures yield continuous (e.g., *T*-score, sum score) and categorical (i.e., risk-classification) indicators of emotional and behavioral risk, and that stakeholders can use both types of data to inform decisions about student needs and follow-up actions, consistency in informant ratings of risk were examined for each score type (i.e., continuous and categorical). Additionally, we aimed to examine if parents or teachers can provide incrementally valid information to the prediction of longitudinal kindergarten academic and social emotional functioning. As the literature is not conclusive regarding which informant, if any, is superior when screening for behavioral and emotional risk in preschoolers, we used a multiple-gating context to examine both teachers and parents as an initial gate informant, followed by the other informant providing second gate assessment information. Multiple screening forms (see measures description below) were employed to generalize results beyond the use of one specific screener form. Additionally, we used an independent measure of academic performance not rated by parents or teachers as an outcome variable to test these relations without the confound of shared method variance between the predictor and the outcome (e.g., criterion contamination; De Los Reyes et al., 2015). Consistent with a research agenda recently proposed for informant discrepancies in school-based research (De Los Reyes et al., 2019), this paper examines cross-informant correspondence in preschools with the goal of informing practice on how to use multi-informant data to inform school-based service delivery. The following research questions were addressed:

1. How similar are continuous teacher and parent ratings of preschool children's behavioral and emotional risk?

2. To what degree is there overlap in the children categorically identified as at-risk via teacher and parent ratings?
3. To what extent do teacher and parent ratings of preschool children's social-emotional risk incrementally predict kindergarten academic and social-emotional outcomes?

Method

Participants

Participants were preschool children, parents, and teachers from five schools in a Title 1 school district in central California serving a predominately Latinx community. All preschools were state-funded and intended to serve families experiencing low socio-economic circumstances. First, a cross-sectional dataset consisting of teacher and parent ratings of preschoolers across three years was created. Overall, there were $N = 535$ preschool children and parents included in the cross-sectional sample, with varying subsample combinations for analyses depending on the number of forms and raters. Most parent participants were female (84.9%), Latinx/Hispanic (82.5%), and held a high school diploma or higher-education degree (65.1%). About half of parents reported speaking Spanish at home (51.5%) and more than half of the parent packets were completed in Spanish (52.6%). Parents reported demographic information on behalf of preschool-aged girls (48.3%) and boys (44.0%) who participated in this study. Parents reported that most children were Latinx/Hispanic (78.4%), and that 37.3% of children spoke English, 33.8% spoke Spanish, and 19.8% spoke both languages when at home. Participating preschool teachers ($N = 14$) were mostly female (92.9%), held at least a college degree (71.4%), and had taught for more than 10 years (64.3%). Half of teachers identified as White/European (50.0%), 28.6% as Latinx/Hispanic, and 14.3% with another race/ethnicity.

Across three years, parents completed Behavioral and Emotional Screening System (BESS; Kamphaus & Reynolds, 2015) forms ($n = 492$) and Pediatric Symptom Checklist-17 (PSC-17; Jellinek et al., 1988) forms ($n = 496$). Teachers also completed BESS ($n = 522$) and PSC-17 ($n = 518$) forms. A small group of students completed two years of preschool; for these students ($n = 85$) only ratings from their second year of preschool (i.e., the year prior to kindergarten) were used for the current study. Preschool children included in the cross-sectional sample were aged 3–5 years old ($M = 4.32$) when they were rated by parents and teachers. Multi-informant (i.e., both parent and teacher) ratings of $n = 476$ preschoolers were available for the BESS and $n = 475$ for the PSC-17. Parents completed both the BESS and PSC-17 for $n = 488$ preschool students; teachers returned both forms for $n = 512$ preschool students.

Next, a smaller longitudinal subsample of participants with multi-informant ratings was created for whom kindergarten outcomes were available. Of the $N = 535$ preschoolers in the cross-sectional sample, longitudinal kindergarten readiness outcomes were available for $n = 298$ students, and longitudinal early literacy outcomes were available for $n = 216$ students, resulting in two longitudinal subsamples. Child participants were 4–5 years old when evaluated for kindergarten readiness ($M = 5.37$) and early literacy ($M = 5.36$). We

did not assess reasons for student attrition during the study. Reduced participation in the longitudinal study was likely due to families enrolling their preschool children in elementary schools in another district or at a nonparticipating school within the same district. See Table 1 for detailed parent, teacher, and student demographic information.

Procedure

Following Institutional Review Board and school board approval, all parents and teachers at participating preschools were invited to participate. Data collected for this study were part of a larger federally funded project examining screening instruments for use in preschools. Beginning in 2016, preschool children were recruited in the Fall of three consecutive academic years (i.e., 2016–17, 2017–18, and 2018–19). Parents and teachers completed screening forms twice per year (i.e., Fall and Spring). School archival data were requested for all participating preschool students' kindergarten outcomes at the conclusion of the study (Spring 2019). See Figure 1 for a timeline describing data collection and study enrollment. Following active consent procedures, parents and teachers were provided with a packet of information to be completed. Parent packets were made available in the language (English or Spanish) reported by school district records. Parents and teachers of preschoolers were asked to complete two brief social and emotional screening forms (described below) and provide demographic information. Researchers were available to assist with the completion of packets. Although participation in the project was voluntary, incentives were provided for participation. Upon completion of the data packets, books were distributed to participating students and their families were entered into a raffle for a zoo or museum membership. Amazon gift cards were distributed to teachers upon completion of forms each year and the participating school district was provided a monetary stipend in recognition of their support of the research project.

Data from screening forms completed by teachers and parents in the Fall of participating students' preschool year were the focus of this study. To answer the third research question, Fall preschool screening data were compared to kindergarten outcomes. Kindergarten outcome data were collected in the Fall of the students' kindergarten year and provided to researchers at the end of this study from archival school district records. Information on kindergarten readiness was obtained from the students' kindergarten teacher within the first two months of school (i.e., late August to October). Information on students' early literacy was obtained via adaptive computerized testing administered at students' schools and by school staff. Early literacy data used in this study represent students' first assessment administered in their kindergarten year according to each school's computerized assessment schedules (first administration varied by school between August and February each year).

Measures

BESS Parent and Teacher Preschool—The Behavior Assessment System for Children-3 (BASC-3) Behavioral and Emotional Screening System (BESS) Parent and Teacher Preschool forms (Kamphaus & Reynolds, 2015) are nationally standardized brief rating scales designed to assess for behavioral and emotional risk. Parent and teacher report forms evaluate adaptive skills, internalizing problems, and externalizing problems. On the BESS Parent Preschool form, parents respond to 29 items on a four-point response scale

describing the frequency that the behavior is observed (i.e., *Never, Sometimes, Often, Almost Always*). On the BESS Teacher Preschool form, teachers respond to 25 items using the same response scale. Forms are stated for use with children between the ages of 2 ½ through 5 years, 11 months. For both forms, a Behavioral and Emotional Risk Index (BERI) is provided as an index (*T*-score) of the overall level of problem behaviors, with higher scores indicative of more emotional and behavioral risk. Consistent with the BESS scoring guidelines, forms with five or more items with missing or multiple responses were not scored. The BERI was calculated when four or fewer items were omitted, and externalizing and internalizing subscale scores were computed only when one or fewer responses on relevant items were omitted (Kamphaus & Reynolds, 2015). *T*-scores of 71 or higher represent the highest level of *Extremely Elevated* risk, *T*-scores between 61–70 represent *Elevated* risk, and *T*-scores of 60 or below are classified as having *Normal* levels of risk. Additionally, three sub-index scores are provided (i.e., Externalizing Risk Index, Internalizing Risk Index, and Adaptive Skills Risk Index). Continuous *T*-scores computed for the BERI and mean scores¹ for the Externalizing Risk and Internalizing Risk indices were the focus of this study. There was evidence of acceptable internal consistency of the BESS Preschool with parents ($\alpha = .87$) and teachers ($\alpha = .83$) in this study. The BESS manual provides indication of internal consistency and test-retest reliability and of the measure's validity, with moderate correlations reported between BESS screening scores and criterion measures of symptoms of emotional and behavioral problems (Kamphaus & Reynolds, 2015).

Pediatric Symptom Checklist-17—The Pediatric Symptom Checklist-17 (PSC-17; Jellinek et al., 1988) is a brief measure of internalizing, externalizing, and attention problems in children aged 4 to 17 years. Although originally validated with parents and in primary care settings (Murphy et al., 2016), recent evidence supports use of the PSC-17 with preschool teachers (DiStefano et al., 2017; Liu et al., 2019). Parents and teachers respond to 17 items on a 3-point scale describing the frequency that the behavior is observed (i.e., 0 = *Never*, 1 = *Sometimes*, 2 = *Often*). A total score, representing overall psychosocial functioning, is provided by summing item responses across all 17 items. Scores of 15 or above indicate that the child is “at-risk”, whereas scores below 15 are indicative of normative levels of social emotional functioning (Gardner et al., 2007). A total score was computed for all forms for which fewer than 4 items were omitted. Scores for the internalizing (5 items), attention (5 items), and externalizing (7 items) subscales are computed by summing across items within a given scale. Scores of 5 or above are indicative of internalizing risk whereas scores of 7 or above indicate risk for each of the attention and externalizing scales (Murphy et al, 2016). Internalizing and externalizing sum scores were computed for all forms for which fewer than 2 or 3 items were omitted, respectively. Sum scores for the total, internalizing, and externalizing scales were used in this study. Internal consistency for the PSC-17 was acceptable (i.e., parents $\alpha = .78$, teachers $\alpha = .92$) in this study. Previous research in preschool samples has supported the predictive and

¹Formulas and norming information for internalizing and externalizing risk subscales are not available in the BESS manual, thus *T*-score conversion of these raw scores was not possible. After manually computing raw internalizing and externalizing sum scores per instructions in the BESS manual, mean scores were calculated to allow for comparison of teacher and parent responses on these subscales.

criterion-related validity of the PSC-17 with emotional and behavioral diagnostic measures (e.g., Liu et al., 2019).

Kindergarten Student Entrance Profile—We examined the ability of screeners to predict kindergarten readiness by using the Kindergarten Student Entrance Profile (KSEP; Quirk et al., 2010). The KSEP is a 12-item teacher rating scale that assesses social-emotional (6 items) and cognitive (6 items) readiness based on observed behaviors (Lilles et al., 2009). A rating rubric for each item indicates the child’s level of mastery on a 4-point response scale (i.e., *not yet*, *emerging*, *almost mastered*, *mastered*). Children can demonstrate readiness in any language or form of communication. Psychometric investigations have found evidence supporting the factor structure of the social-emotional and cognitive scales of the KSEP, indicating that items on these scales measure two distinct, yet related aspects of school readiness with reliability estimates ranging from $\alpha = .81$ – $.92$ (Quirk et al., 2014; Quirk et al., 2016). Additionally, the KSEP has been found to be a significant predictor of children’s reading fluency and social-emotional wellbeing in Grade 5 (Quirk et al., 2017). Based on their direct observations of students, teachers completed the KSEP for each participant in this study in the fall of their kindergarten year. School personnel used summed item ratings to create categorical social-emotional and cognitive readiness domain scores (1 = *not yet* to 4 = *mastered*). Consistent with prior research (e.g., Quirk et al., 2017), domain scores were treated as continuous despite their ordinal nature.

STAR Early Literacy—The Renaissance STAR (STAR; <http://www.renlearn.com/sr/>) are standardized assessments to assess early literacy. These computer adaptive assessments were rated favorably by the National Center for Response to Intervention and the STAR Early Literacy assessment was used as an independent criterion to examine the ability of the BESS and P SC-17 screeners to predict academic achievement in the area of literacy. A meta-analysis found STAR Early Literacy had an average correlation of .60 to other early literacy assessments (Renaissance Learning, 2009). Results from a study conducted by Clemens et al. (2015) support the predictive utility of STAR Early Literacy for children’s reading skills in kindergarten and Grade 1. Scaled scores were used as indicators of kindergarten academic achievement in the present study.

Analysis

Continuous Teacher and Parent Ratings—The strength of the relation between teacher and parent screening scores was examined through Pearson correlations between teacher and parent report on each of the BESS and PSC-17 BERI/total score, internalizing subscale, and externalizing subscale. Strength of observed relations can be described using Cohen’s (1988) guidelines: “small” – $r = .10$, “medium” – $r = .30$, and “large” – $r = .50$. Cohen (1988) and other scholars, however, emphasize the importance of interpreting observed effects within the context of previous research (Hill et al., 2008). Thus, results will also be interpreted within the context of recent meta-analytic research that found an overall cross-informant correlation of .28 (95% CI [.22, .33]), and an effect size of .25 and .30 for internalizing and externalizing symptoms, respectively (De Los Reyes et al., 2015). Although correlations are frequently presented in multi-informant examinations, correlation is necessary but insufficient in informing concordance between raters, as correlations inform

on strength of association but not degree of agreement (Stolarova et al., 2014). Thus, in the present study, dependent samples *t*-tests were also performed to determine whether teacher and parent ratings significantly differ, on average, for each of the BESS and PSC-17 BERI/total score, internalizing subscale, and externalizing subscale. All comparisons were performed using pairwise deletion for missing data. A conservative *p* value of .01 was used to correct for multiple *t*-tests.

Categorical Teacher and Parent Ratings—Agreement between categorical classification results for teacher and parent screening ratings were further examined. Following recommendations to focus on indicators of elevated risk for general mental health conditions in universal screening practice (e.g., Levitt et al., 2007), a total risk score was used to examine overlap in the children categorically identified as at-risk via teacher and parent ratings. Consistent with schools' use of screening results, students in the elevated or extremely elevated groups on the BESS were combined to represent one group indicative of being "at-risk" (Eklund & Dowdy, 2014). Teacher and parent BESS scores for the BERI were coded such that 1 = risk present ($T \geq 61$), and 0 = no risk present ($T < 60$). Similarly, teacher and parent PSC-17 total scores were coded such that 1 = risk present (total score ≥ 15) and 0 = no risk present (total score < 14). Participant scores on the teacher and parent screening measures were used to place children into groups to compare differences between identification methods (i.e., by rater and by instrument). Children were placed into groups according to their risk-status for each measure and across informants (i.e., BESS parent and teacher, PSC-17 parent and teacher). All groups were formed using pairwise deletion for missing data. Cohen's Kappa (κ) was then calculated to measure inter-rater agreement in risk classification for teachers compared to parents for each of the BESS BERI and PSC-17 total scores. Strength of agreement was rated as: poor ($\kappa < .00$), slight ($\kappa = .00-.20$), fair ($\kappa = .21-.40$), moderate ($\kappa = .41-.60$), substantial ($\kappa = .61-.80$), or almost perfect ($\kappa = .81-1.00$; Landis & Koch, 1977).

Predictive Utility—To test the utility of teacher and parent screening scores to predict kindergarten outcomes (i.e., STAR Early Literacy scale score, KSEP cognitive and social-emotional domain scores), a series of hierarchical regression analyses were specified using IBM SPSS Statistics 26. Hierarchical regression analysis was chosen to determine the strength of a given rater's (i.e., parent or teacher) screening scores on kindergarten academic and social-emotional outcomes before examining the incremental contribution of the second rater. Models were specified separately for each screening tool (i.e., BESS or PSC-17). Demographic variables were entered into the first block of all models to account for the relation between these indicators and children's social-emotional and academic performance (e.g., Edyburn et al., 2020; Denham et al., 2012; Halle et al., 2012; Hammer et al., 2011; Romer et al., 2011). Demographic variables included: language of parents' rating forms (1 = Spanish, 0 = English), parent-reported child's home language preference [dummy coded to be Spanish (1 = Spanish, 0 = else) and bilingual (1 = Spanish and English, 0 = else)], and child gender (1 = female, 0 = male). The second block was then varied to capture contributions of each rater, and to mimic a multiple-gating system in which one informant provides the first-gate screening information. That is, models were specified such that block 2 included either a teacher reported total problems score *or* a parent reported total problems

score for a given measure (e.g., teacher-reported BESS BERI *T* or parent-reported BESS BERI *T*). Mimicking a second gate, block 3 then included the score for the alternate rater of the same measure (i.e., if block 2 included teacher-reported BESS BERI *T*, block 3 included parent-reported BESS BERI *T*). This was conducted to determine if adding information from a second informant significantly aided in the prediction of outcomes. For example, when parents were entered in the first block and teachers entered in the second block, this was designed to mimic utilizing parents as a first gate informant followed by teachers as a second gate informant. All analyses were performed using listwise deletion form missing data. A conservative *p* value of .01 was used for model statistics to correct for multiple tests.

Results

Continuous Teacher and Parent Ratings

Correlations for parent and teacher scores on the PSC-17 (i.e., total, internalizing, externalizing) and BESS (i.e., BERI, internalizing, externalizing) are presented in Table 2. Parent and teacher screening scores were significantly correlated across forms by rater as well as across raters. Total symptom ratings were more strongly correlated across measures for the same informant ($r = .69 - .89$), than across informants for the same measure ($r = .22 - .26$). Teacher ratings were more strongly correlated across measures than parent ratings. Teachers' ratings of students' internalizing and externalizing symptoms were moderately correlated across measures ($r = .45 - .43$). The relation between parents' ratings of their children's internalizing and externalizing symptoms were more strongly associated on the BESS ($r = .64$) than on the PSC-17 ($r = .26$). Further, associations between teacher and parent ratings of children's internalizing and externalizing symptoms, using the same measure, were stronger for externalizing ($r = .18 - .26$) than internalizing symptoms ($r = .04_{NS} - .18$).

Means, standard deviations, and results of dependent-samples *t*-tests are presented in Table 3. Significant differences for parent and teacher total symptom ratings were observed for the PSC-17 but not the BESS BERI. On average, parents' PSC-17 ratings indicated children to have significantly more problems than did teacher PSC-17 ratings. Across measures, teacher and parent ratings of children's internalizing symptoms significantly differed. Parents rated their children to have significantly more internalizing problems than did their teachers using the BESS whereas teachers rated children to have significantly more internalizing problems than did parents using the PSC-17. Children's parents rated them to have significantly more externalizing problems on the PSC-17 than did their teachers, with BESS results also trending in this direction.

Categorical Teacher and Parent Ratings

The percentage of students identified as at-risk or not at-risk by informant and measure is displayed in Table 4 A–D. Overall, most (84% – 94%) preschool children were rated within the “normal” range by their parents and teachers on both measures. Teachers were more likely to rate children within the “at-risk” range on both the BESS and PSC-17 than were parents. Further, more children were rated within the “at-risk” range on the BESS than on the PSC-17. When using the same measure, fewer than 2% of children were classified as “at-

risk” by both teacher and parent ratings (see Table 4C and 4D). Cohen’s κ was calculated as indication of the strength of inter-rater agreement between parents and teachers for each of the BESS BERI and PSC-17 total symptom scores. Although statistically significantly greater than chance, strength of agreement for both measures was “slight,” $\kappa(\text{BESS}) = .08$ (95% CI [-.02, .19]), $p = .046$, $\kappa(\text{PSC-17}) = .09$ (95% CI [-.01, .21]), $p = .022$ (Landis & Koch, 1977). Results in Table 4 indicate that parents and teachers are more likely to agree on “normal” than on “at-risk” classifications.

Predictive Utility

Correlation coefficients for kindergarten outcomes with parent and teacher PSC-17 and BESS screening scores are presented in Table 2. Overall, preschool teacher ratings tended to be more strongly correlated with kindergarten outcomes than were parent ratings, with the strongest correlations observed for teacher PSC-17 total and BESS BERI scores with kindergarten-teacher rated KSEP social-emotional readiness ($r = -.44$ and $-.47$, respectively). Parent BESS BERI scores, but not PSC-17 scores, were significantly correlated with kindergarten outcomes. Results of hierarchical regression analyses testing whether teacher and parent PSC-17 and BESS BERI T ratings incrementally predict kindergarten academic and social-emotional outcomes are displayed in Tables 5 through 8. In our presentation of results below, we first describe block 1 demographic results, followed by results for blocks 2 and 3. We first present results for the models including BESS BERI predictors then for the models including PSC-17 predictors.

Early Literacy—The block 1 model with demographic characteristics predicting early literacy was not significant at $p < .01$, although parent-reported children’s preference for speaking Spanish at home was associated with lower STAR early literacy scores in kindergarten ($\beta = -.28$). In subsequent models, child’s Spanish language preference was a significant predictor of early literacy, with the exception of the model with parent BESS entered in block 2 ($p = .011$). Language of parent rating forms, child preference for both English and Spanish, and child’s gender were not significantly associated with kindergarten early literacy.

We then entered teacher-rated information in block 2, followed by parent-rated information in block 3 as a proxy for a multiple gating system with teachers as the first gate informant and parents as the second gate informant. Entered in block 2, teacher BESS BERI T scores were significantly predictive of early literacy in kindergarten above and beyond the significant association of early literacy with child’s Spanish language preference ($\beta = -.27$; see Table 5). Parent BESS BERI T scores included in block 3 were not significantly associated with kindergarten early literacy and did not significantly improve model prediction ($\beta = -.07$; $R^2 = 0$). Next, to mimic a multiple gating system in which parents were the first gate informant followed by teachers as a second gate informant, a model including parent BESS BERI T in block 2 with teacher BESS BERI T added in block 3 was tested (Table 6). The overall model for parent BESS BERI T entered in block 2 was not statistically significant. The model became significant when teacher ratings were included in block 3, with teacher BESS BERI T being significantly associated with kindergarten early literacy ($\beta = -.25$). Fewer teacher-rated social-emotional problems

reported on the BESS in preschool were significantly associated with children's higher early literacy scores in kindergarten.

We then replicated this same analysis substituting BESS BERI *T* scores for PSC-17 total scores, mimicking a multiple gating system with an alternate measure. Teacher PSC-17 total scores entered in block 2 were not significantly predictive of early literacy in kindergarten ($\beta = -.16, p = .032$) above and beyond the significant association of early literacy with child's Spanish language preference (see Table 7). When parent PSC-17 total scores were entered in block 3, the overall model fell just above the threshold for statistical significance ($p = .01$). Neither parent PSC-17 total scores nor teacher PSC-17 total scores were significantly associated with kindergarten early literacy in this model. The model that included parent PSC-17 ratings as a first gate in block 2 (see Table 8) was also not significant. This model remained nonsignificant when teacher PSC-17 ratings were subsequently entered in block 3. Demographic and screening scores included in tested models were associated with between 6% and 10% of the variance in kindergarten early literacy scores.

Kindergarten Cognitive Readiness—Similar models, mimicking multiple informants across various screening forms, were then tested examining kindergarten cognitive readiness as the outcome measure. The block 1 model with demographic characteristics was significant; Spanish language of parent rating forms was associated with children's lower kindergarten cognitive readiness ($\beta = -.19$). The *p* value for this variable was above our a priori criteria for statistical significance for BESS models, likely due to sample variation. Spanish language of parent rating forms remained a predictor of kindergarten cognitive readiness after including parent and teacher screening results, with β achieving statistical significance only for PSC-17 models.

Entered in block 2, teacher BESS BERI *T* was significantly predictive of kindergarten cognitive readiness ($\beta = -.22$; Table 5), such that more teacher-reported social-emotional problems were associated with lower cognitive readiness in kindergarten. Including parent BESS BERI *T* in block 3 did not significantly improve prediction of kindergarten cognitive readiness ($R^2 = 0$); parent BESS BERI *T* was not significantly associated with kindergarten cognitive readiness ($\beta = -.06$). In the model with parent BESS BERI *T* entered as a first gate in block 2 (Table 6), parent BESS BERI *T* was not significantly associated with kindergarten cognitive readiness ($\beta = -.12, p = .049$). Teacher BESS BERI *T* added in block 3 was significantly associated with kindergarten cognitive readiness ($\beta = -.20$) and resulted in significant improvement in variance associated with model predictors ($R^2 = .04$).

Paralleling results for models with BESS teacher ratings entered first in block 2, teacher PSC-17 total score entered in block 2 was significantly predictive of cognitive readiness ($\beta = -.22$; Table 7). Parent-rated social-emotional problems included in block 3 were not significantly associated with kindergarten cognitive readiness and did not result in significant improvements in variance associated with the model ($\beta = -.06; R^2 = .0$). Similarly, the overall model with parent PSC-17 scores entered in block 2 was significant (Table 8), but parent PSC-17 total score ratings did not significantly predict kindergarten cognitive readiness ($\beta = -.04, p = .50$). The observed change in variance explained when adding the parent PSC-17 total score predictor in block 2 of the model was also

negligible ($R^2 = 0$). The addition of teacher-rated PSC-17 total scores in block 3 resulted in a significant improvement in variance explained ($R^2 = .03$); higher teacher PSC-17 total scores were significantly associated with lower kindergarten cognitive readiness ($\beta = -.18$). Overall, models tested were associated with 10%–11% of the variance in children's kindergarten cognitive readiness.

Kindergarten Social-Emotional Readiness—Finally, models were replicated with kindergarten social-emotional readiness as the outcome. The block 1 model with demographic characteristics predicting social-emotional readiness was not significant at $p < .01$. In the model including parent BESS screening scores in block 2, child's female gender was significantly associated with more social-emotional readiness ($\beta = .17$). However, child's gender was not a significant predictor when teacher BESS BERI or teacher PSC-17 total scores were included in either of block 2 or 3 of their respective models.

Teacher BESS BERI T entered in block 2 was significantly associated with kindergarten social-emotional readiness ($\beta = -.46$), such that fewer preschool teacher rated social-emotional problems was associated with greater kindergarten social-emotional readiness (Table 5). Including parent BESS BERI T in block 3 did not significantly improve prediction of kindergarten social-emotional readiness; parent BESS BERI T was not significantly associated with kindergarten social-emotional readiness above and beyond demographic indicators and teacher BESS BERI T ($\beta = -.07$). In contrast, when parent BESS BERI T was entered as a first gate in block 2 (Table 6), parent BESS BERI T was significantly associated with kindergarten social-emotional readiness ($\beta = -.19$). However, when teacher BESS BERI ratings were subsequently entered in block 3, teacher BESS BERI T was significantly associated with social-emotional readiness ($\beta = -.44$), but parent BESS BERI ratings were no longer significantly related to social-emotional readiness ($\beta = -.07, p = .26$).

Teacher PSC-17 total score entered in block 2 was also significantly associated with kindergarten social-emotional readiness ($\beta = -.43$), with fewer teacher-rated social-emotional problems associated with greater kindergarten social-emotional readiness (Table 7). Parent PSC-17 total score included in block 3 was not significantly associated with kindergarten social-emotional readiness and did not result in significant improvements in variance associated with the model ($\beta = .05, R^2 = 0$). The model with parent PSC-17 ratings entered as a first gate (Table 8) was not statistically significant. When teacher PSC-17 ratings were included in block 3 the overall model was statistically significant and teacher PSC-17 ratings were significantly associated with kindergarten social-emotional readiness ($\beta = -.44$). Overall, significant models accounted for between 20% and 22% of the variance in social-emotional readiness.

Discussion

Accurate and efficient early identification of behavioral and emotional difficulties at an early age is critical for supporting prevention and early intervention efforts. Overall, consistent with prior literature, results of continuous score comparisons indicated that parent and teacher ratings were significantly correlated across forms and raters. However, several nuanced results were of interest. When compared to previous meta-analytic research that

examined cross-informant correspondence, the observed correlations between parent and teacher screening scores in this study tended to be weaker ($r_{PSC-17} = .04 - .24$, $r_{BESS} = .17 - .26$) than the overall effects reported in a meta-analysis by De Los Reyes and colleagues (2015; $r = .25 - .30$). Although the strength of the relation between teacher and parent total symptom ratings on each of the BESS and PSC-17 fell within the 95% confidence interval for overall effects estimated by De Los Reyes et al. (2015; [.22, .33]), correspondence between parents' and teachers' reports of internalizing and externalizing risk was weaker in the present study. Further, when comparing internalizing problems across raters and forms, parents tended to endorse more problems than teachers on the BESS whereas teachers tended to endorse more problems than parents on the PSC-17. However, when comparing externalizing problems, parents tended to endorse more problems than teachers across both forms.

Categorical results showed that teachers were more likely to identify students as "at-risk" across both forms, which is important to consider within the context of multiple-gating screening approaches. In multiple-gating approaches, high sensitivity is often prioritized over specificity in an initial screening gate, which suggests erring on the side of increased false positives (Glover & Albers, 2007). Given this, it may be prudent to use teachers as a first gate assessment, as they are effectively identifying more children, with a more balanced approach to sensitivity and specificity in subsequent gates. However, results also highlight that teachers and parents are more likely to agree when students are placed in the "normal" classification area. Thus, if multiple informants are used in an initial screening gate, it is recommended that further assessment include students who are identified as "at risk" by either parents or teachers. Consistent with prior research (e.g., De Los Reyes et al., 2015), the overall strength of agreement between raters' risk classifications across forms was small, suggesting that schools will get different results depending on the measures and raters used.

Predictive utility results using analytic techniques designed to mimic a multiple-gating procedure highlighted the superiority of teachers as first gate informants when compared to parents with respect to predicting early literacy and kindergarten cognitive and social emotional readiness. Although order of entry of informant screening scores into regression models was arbitrary and replicated with both informants entered first (i.e., teachers as first gate informant followed by parents as a second gate informant; parents as first gate informant followed by teachers) for each measure (i.e., PSC-17 or BESS), teacher-rated screening data were consistently predictive of kindergarten social-emotional and cognitive readiness. Teacher-rated BESS, but not PSC-17, scores were able to significantly predict early literacy. Findings build upon previous research on an earlier version of the teacher-rated BESS which showed moderate to large associations with measures of kindergarten readiness, social emotional development, and receptive vocabulary (Dowdy et al., 2013), as well as moderate sensitivity and high specificity, positive predictive value, and negative predictive value when predicting emotional and behavioral problems (Kettler et al., 2017). Together, these results highlight that the teacher-rated BESS is predictive of a range of important educational and social-emotional outcomes.

Preschool teacher screening scores were also more strongly correlated with kindergarten outcomes, specifically kindergarten-teacher rated readiness, than were parent screening

scores. This is consistent with previous research indicating that predictor and criterion measures completed by the same or similar informants (i.e., both teachers) are more strongly related than are predictors with objective or alternate-rater criterion variables (De Los Reyes et al., 2015). Thus, although parents as a second gate informant did not significantly improve model prediction relative to the use of teachers alone when predicting later outcomes, it is important to recognize the potential impact of criterion contamination in that both kindergarten social-emotional and cognitive readiness outcomes were rated by teachers.

When examining the use of parents as a first gate informant, results highlighted that although parent ratings using the BESS were predictive of kindergarten social-emotional readiness, parent ratings using the BESS or the PSC-17 were not predictive of kindergarten early literacy or kindergarten cognitive readiness. In social-emotional readiness models, when teachers were added as a second gate informant, parent-rated information was no longer significantly associated with later outcomes. This further emphasizes the value of teacher-rated information for predicting kindergarten outcomes. However, other research examining the value of parent-rated information has shown that, at kindergarten entry, parent rated screening data can predict later social, behavioral, and academic functioning beyond information obtained by academic screeners alone (Owens et al., 2015). Further, in a review of four preschool social and emotional screening systems for use by both parent and teacher raters, Feeney-Kettler et al. (2010) found significant variance among informants with none of the four measures reviewed reporting high predictive validity across all validity indices when examining later social emotional outcomes; we found similar results for the BESS and PSC-17. Unfortunately, there is still a dearth of information on the ability of preschool parent and teacher rated screeners to predict a variety of later educational and behavioral outcomes. In light of the lack of a preponderance of research evidence, and considering that the behaviors of young children differ significantly across informants and settings, it is still often advised to gather information from multiple informants (Feeney-Kettler et al., 2010). In fact, in some research examining the ability of parents and teachers to predict later outcomes, informant ratings have been averaged to conform to clinical practice recommendations to collect information from multiple informants (Essex et al., 2009). Specifically, when mother and teacher social-emotional ratings from children in early elementary school were combined, they were able to significantly predict academic, social, and global impairment outcomes, physical health problems, and mental health service use in Grade 5 (Essex et al., 2009). This further emphasizes the potential value of early screening for childhood mental health problems.

Child and parent demographic characteristics also demonstrated significant associations with outcomes of interest in this study, with relevant demographic characteristics varying by kindergarten outcome and screening measure. Child language preference was associated with kindergarten early literacy, such that children who preferred speaking Spanish at home were likely to have lower early literacy scores in kindergarten than were their peers with English or both English and Spanish language preference; this effect was in the small to medium range (Cohen, 1988). This finding is consistent with previous research indicating that children's home language experiences impact their early language skills, including early literacy (Halle et al., 2012; Lewis et al., 2016). Similarly, language in which parents completed this study's measures was associated with kindergarten cognitive readiness in

PSC-17 models, such that completion of measures in Spanish was associated with children's lower cognitive readiness in kindergarten; this effect was small. Finally, and also consistent with prior literature, child's gender was associated with kindergarten social-emotional readiness, such that girls were more likely to have greater social-emotional readiness than their male peers (Romer et al., 2011). This effect was small and was not robust to the inclusion of teacher screening scores in regression models. Additional research is needed to further understand the demographic variants of early literacy and kindergarten readiness, as well as to disentangle potential moderators of these effects.

Implications

A primary goal of this study was to inform practice recommendations regarding the use of parent and teacher informants in preschool settings. Results from this study are important in informing selection of informants and measures in universal screening for behavioral and emotional risk in a preschool population; however, the results do not provide a definitive solution to inform a standard of practice with results varying depending on the cross-sectional or longitudinal nature of the questions examined. As such, it is likely that an accumulation of evidence is still needed and replication is warranted. In the current sample, it is clear that a majority of preschoolers did not demonstrate significant emotional and behavioral risk, and that these results were consistent across forms and raters. Practitioners may take comfort in the knowledge that regardless of the measure or informant used, only a small seemingly-manageable percentage of preschoolers may need more intensive follow-up assessment or intervention. However, classification into a "risk" category was more unreliable across forms and raters. In the absence of additional information, it is still recommended that practitioners follow up with these students, regardless of if they were only identified by one measure or one form. Future studies in applied practice will be critical to further inform if these students classified as "at-risk" are significantly different from those identified as within normal ranges.

It may be important to note that the BESS, both teacher and parent rated forms, identified more students as "at-risk" than the PSC-17. There are a range of potential explanations for this, but it may mean that the BESS items are more sensitive to problems among preschoolers or that the BESS items are easier to understand. Of course, without information on "true" positives, it could also be that the BESS is identifying more students as at risk than is actually the case. Recognizing that the same PSC-17 form is used across a variety of different developmental levels (i.e., ages 4 to 17), revisions to certain items may be needed to ensure developmental appropriateness. For example, it may be difficult for parents and teachers of preschoolers to understand what it behaviorally looks like when a child "feels hopeless."

As results may differ depending on the nature of the areas assessed, it will be important for practitioners to have a shared understanding with their school teams regarding "what" to screen for (Dowdy et al., 2010). Although parents were more likely to rate children as having internalizing problems than teachers on the BESS, the opposite was true for the PSC-17. However, considering recommendations to use broad instruments in universal screening (Levitt et al., 2007) and to interpret an overall score prior to subscale scores (e.g.,

Rodriguez et al., 2015), a measure that assesses multiple areas of risk may be preferred to ensure that students with a variety of risk indicators are considered for additional assessment or intervention (Levitt et al., 2007). Additionally, given that significant differences were present when examining categorical but not continuous scores, it is important to remember the limitations of relying on single cut scores when making diagnostic or placement decisions. Although cut scores are routinely used in practice to make decisions regarding which students receive services, cut scores always lead to some misidentification, can produce over-simplified results, and do not capture the full variation present in children's emotional and behavioral functioning (Moore et al., 2019). Examining continuous scores, utilizing more person-centered approaches to classification, and combining screening data with additional information (e.g., behavioral observations, teacher and parent interviews, grades and standardized achievement data, discipline referrals), as opposed to solely relying on screening cut scores may help to ensure that all students receive needed services (Dowdy et al., 2014; Stiffler & Dever, 2015).

Questions regarding which informant to use within a multiple-gating screening procedure became clearer when examining the ability of screening scores to predict later outcomes in kindergarten. Generalizing across all models and combinations of informants at different gates, it is clear that if the goal is the prediction of later kindergarten readiness and early literacy outcomes, teachers are an ideal informant and the addition of parents as a second gate may not be necessary. Of course, parents provide valuable and needed information about their children's social-emotional functioning. In reality, however, given resource constraints, it is highly likely that practitioners will only give one screening measure and perhaps only to one informant. Within a school-based screening program, results from this study suggest that it is advisable to gather screening information from teachers and that information from a single informant is sufficient to predict later kindergarten outcomes. These multiple gating results mirror similar research at the high school level which found little to no added value of multiple informants or multiple gates (Dowdy et al., 2016). However, prior research investigating multiple- and single-gate measures with preschoolers found informant results to differ depending on the screening measures used and the outcomes assessed (Kettler et al., 2017). Based on their results, Kettler et al. (2017) recommended either multiple- or single-gate screening with teacher informants and concluded that single-gate measures (e.g., BESS) may be most appropriate for integrating parent perspectives into identification of preschool children's emotional and behavioral risk. Overall, further research is needed to provide a more nuanced understanding regarding under which circumstances it is most prudent to gather information from multiple informants and use multiple gates.

Limitations and Future Directions

The sample characteristics consisting of a largely Latinx population of children and parents may be viewed as both a strength and limitation. Despite the need for culturally-responsive practices, and the increase in research on Latinx children (Perez Huber et al., 2015), there continues to be a critical demand for research on samples historically underrepresented in research. This may be particularly important considering the growing proportion of Latinx children attending schools in the United States (U.S. Census Bureau, 2017), and that

few screening measures for use within MTSS frameworks have been studied extensively within the Latinx preschool population (Rodriguez et al., 2017). Considering contextual circumstances that can place Latinx students at higher risk for social emotional difficulties, an understanding that Latinx students are less likely to receive needed mental health services, and a history of discriminatory practices in schools, it is even more essential that research focuses on mental health screening for Latinx students (Bertone et al., 2019; Gudiño et al., 2009). Despite the focus on Latinx preschoolers, it should be noted that the sample in this study was limited to one school district in central California. Considering that Latinx children are a heterogeneous population, results may not be generalizable to other Latinx children in other parts of the country, in addition to a lack of generalizability to other preschools. Additionally, parents completing the BESS and PSC-17 measures in Spanish and English were combined; future examinations of the invariance of the present study's findings across different language versions of the measures are needed (Milfont & Fischer, 2010).

Although common among school-based studies to find that the vast majority of students are within normal risk ranges, this sample was noticeably at low risk. Regardless of the measure or rater, a small percentage (less than 12%) of the preschoolers were rated as at-risk, certainly impacting the ability to detect differences among risk status across raters and measures. Moreover, although results of screening data were provided to teachers following each data collection period, information was not available regarding whether participants with teacher or parent screening scores in the "at-risk" range received intervention following universal screening. Thus, we are unable to identify the extent to which intervention may have impacted observed predictive utility results; additional research that aims to replicate this study's findings while also controlling for intervention exposure is needed. The reasons for attrition in the longitudinal sample were not available, but in future research warrant investigation to determine how attrition may impact results. Overall, replication is needed across more diverse samples, including among more population-based samples with more children who are identified as at risk.

The outcomes in this study were also limited, with only one outcome examining early literacy truly independent of rater effects. Measures of kindergarten readiness in this study were completed by children's kindergarten teachers and were more strongly correlated with preschool teacher screening scores than with children's parents scores. Results with respect to the early literacy outcome should not have been impacted by shared method variance as the evaluation of early literacy skills was computer based. However, it should be noted that there was the potential for criterion contamination with both kindergarten social-emotional and cognitive readiness outcomes being teacher-rated. Future research is needed to examine the predictive validity of parent and teacher ratings across more comprehensive and rater-independent measures of social-emotional and academic functioning. In particular, it will be critical to examine independent measures from sources that do not overlap with the parents and teachers completing rating scales (e.g., as completed by another school staff member or observational data collected by research team members; De Los Reyes et al., 2019). Finally, the analyses used in this study did not account for potential variation in outcomes of interest at the teacher or school level. It will be important for future studies that compare informants for universal screening to account for potential variation at the teacher and/or school levels via multi-level approaches.

Conclusion

Recently, a research agenda examining multi-informant research has been outlined with the aim of further understanding informant discrepancies and how to interpret these discrepancies for use in school-based research and practice (De Los Reyes et al., 2019). New studies, similar to this one, were solicited to further an understanding of informant discrepancies. Given the current state of the field with a strong reliance on rating scales for use in school-based assessment practices, it is likely that teacher and parent rating scales will still be a popular tool used to help direct prevention and intervention services within MTSS systems. Results contribute to a larger body of research examining the use of multi-informant assessments, with this study specifically examining the use of screening instruments for use in preschools. Based on current results, parents and teachers are both viewed as valuable sources of information to consider when screening for emotional and behavioral risk among preschoolers. However, if the goal of screening is to predict later kindergarten outcomes, teacher ratings may be the most informative. With a continued goal of further understanding which informant(s) to use and under what circumstances in school-based screening, it will become increasingly easier to identify students in need and more efficiently provide early intervention services.

Acknowledgments

We would like to thank the families, teachers, and schools who generously contributed their time to this project. We have no known conflict of interest to disclose. The data that support the findings of this study are available from the corresponding author, SM, upon reasonable request. This work was supported by the Institute of Education Sciences, U.S. Department of Education, under Grant R305A150152 to South Carolina Research Foundation and National Institute of Mental Health under Grant T32MH109436 02. The opinions expressed are those of the authors and do not represent views of the Institute of Education Sciences, the U.S. Department of Education, or the National Institutes of Health.

References

- Achenbach TM, McConaughy SH, & Howell CT (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, 101, 213–232. 10.1037/0033-2909.101.2.213 [PubMed: 3562706]
- Barnett SW, Carolan ME, Squires JH, & Clarke Brown K. (2013). *The state of preschool 2013*. New Brunswick, NJ: National Institute for Early Education Research.
- Bertone A, Moffa K, Wagle R, Fleury I, & Dowdy E. (2019). Considerations for mental health screening with Latinx dual language learners. *Contemporary School Psychology*, 23, 20–30. 10.1007/s40688-018-0205-y
- Biederman J, Keenan K, & Faraone SV (1990). Parent-based diagnosis of attention deficit disorder predicts a diagnosis based on teacher report. *Journal of the American Academy of Child and Adolescent Psychiatry*, 29, 698–701. 10.1097/00004583-199009000-00004 [PubMed: 2228921]
- Clemens NH, Hagan-Burke S, Wen Luo, Cerda C, Blakely A, Frosch J, Gamez-Patience B, Jones M, Betts J, & VanDerHeyden A. (2015). The predictive validity of a computer-adaptive assessment of kindergarten and first-grade reading skills. *School Psychology Review*, 44(1), 76–97. 10.17105/SPR44-1.76-97
- Cohen J. (1988). *Statistical power analysis for the behavioral sciences*. New York, NY: Routledge.
- Charlton CT, Sabey CV, Dawson MR, Pyle D, Lund EM, & Ross SW (2018). Critical incidents in the scale-up of state multitiered systems of supports. *Journal of Positive Behavior Interventions*, 20(4), 191–202. 10.1177/1098300718770804
- Denham SA, Bassett HH, Thayer SK, Mincic MS, Sirotkin YS, & Zinsser K. (2012). Observing preschoolers' social-emotional behavior: Structure, foundations, and prediction of early school

- success. *The Journal of Genetic Psychology*, 173(3), 246–278. 10.1080/00221325.2011.597457 [PubMed: 22919891]
- De Los Reyes A, Augenstein TM, Wang M, Thomas SA, Drabick DAG, Burgers DE, & Rabinowitz J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*, 141(4), 858–900. 10.1037/a0038498 [PubMed: 25915035]
- De Los Reyes A, Cook CR, Gresham FM, Makol BA, & Wang M. (2019). Informant discrepancies in assessments of psychosocial functioning in school-based services and research: Review and directions for future research. *Journal of School Psychology*, 74, 74–89. 10.1016/j.jsp.2019.05.005 [PubMed: 31213233]
- DiStefano C, Liu J, & Burgess Y. (2017). Investigating the structure of the Pediatric Symptoms Checklist in the preschool setting. *Journal of Psychoeducational Assessment*, 35, 494–505. 10.1177/0734282916647648
- Dowdy E, Chin J, & Quirk M. (2013). Preschool screening: An examination of the Behavioral and Emotional Screening System Preschool Teacher Form (BESS Preschool). *Journal of Psychoeducational Assessment*, 31, 578–584. 10.1177/0734282913475779
- Dowdy E, Dever BV, Raines T, & Moffa K. (2016). A preliminary investigation into the added value of multiple gates and informants in universal screening for behavioral and emotional risk. *Journal of Applied School Psychology*, 32, 178–198. 10.1080/15377903.2016.1165327
- Dowdy E, DiStefano C, Greer F, Moore S, & Pompey K. (2019). Latent structure examination of the BASC3 BESS Parent Preschool Form. *Journal of Psychoeducational Assessment*, 181–193. 10.1177/0734282917739109
- 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 Doll B, Pfohl W, & Yoon J. (Eds.). *Handbook of youth prevention science*. New York: Routledge. 10.4324/9780203866412.ch4
- Dowdy E, Kamphaus R, Twyford J, & Dever BD (2014). Culturally competent emotional and behavioral screening. In Weist M, Lever N, Bradshaw C, & Owens J. (Eds.), *Handbook of school mental health* (pp. 311–322). New York, NY: Springer. 10.1007/978-1-4614-7624-5_23
- Dowdy E, & Kim E. (2012). Choosing informants when conducting a universal screening for behavioral and emotional risk. *School Psychology Forum*, 6(4), 1–10.
- Durlak JA, Weissberg RP, Dymnicki AB, Taylor RD, & Schellinger KB (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82(1), 405–432. 10.1111/j.1467-8624.2010.01564.x [PubMed: 21291449]
- Edyburn K, Dowdy E, DiStefano C, Bertone A, & Greer F. (2020). Measurement invariance of the English and Spanish BASC-3 Behavioral and Emotional Screening System Parent Preschool forms. *Early Childhood Research Quarterly*, 51(2), 307–316. 10.1016/j.ecresq.2019.12.002
- Eklund K, & Dowdy E. (2014). Screening for behavioral and emotional risk versus traditional school identification methods. *School Mental Health: A Multidisciplinary Research and Practice Journal*, 6, 40–49. 10.1007/s12310-013-9109-1
- Essex MJ, Kraemer HC, Slattery MJ, Burk LR, Boyce WT, Woodward HR, & Kupfer DJ (2009). Screening for childhood mental health problems: outcomes and early identification, 50, 562–570. 10.1111/j.1469-7610.2008.02015.x
- Frick PJ, Burns C, & Kamphaus RW (2009). *Clinical assessment of child and adolescent personality and behavior* (2nd ed.). New York, NY: Springer. 10.1007/978-1-4419-0641-0
- Feeney-Kettler KA, Kratochwill TR, Kaiser AP, Hemmeter ML, & Kettler RJ (2010). Screening young children's risk for mental health problems: a review of four measures. *Assessment for Effective Intervention*, 35(4), 218–230. 10.1177/1534508410380557
- Friedman-Krauss AH, Barnett WS, Garver KA, Hodges KS, Weisenfeld GG, & Gardiner BA (2020). The state of preschool 2019. Retrieved July 24, 2020, from http://nieer.org/wp-content/uploads/2020/07/YB2019_Full_Report.pdf
- Gardner W, Lucas A, Kolko DJ, & Campo JV (2007). Comparison of the PSC-17 and Alternative Mental Health Screens in an At-Risk Primary Care Sample. *Journal of the American Academy*

of Child & Adolescent Psychiatry, 46(5), 611–618. 10.1097/chi.0b013e318032384b [PubMed: 17450052]

- Glover TA, & Albers CA (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology, 45*, 117–135. 10.1016/j.jsp.2006.05.005
- Gudiño OG, Lau AS, Yeh M, McCabe KM, & Hough RL (2009). Understanding racial/ethnic disparities in youth mental health services: do disparities vary by problem type? *Journal of Emotional and Behavioral Disorders, 17*(1), 3–16. 10.1177/1063426608317710 .
- Halle T, Hair E, Wandner L, McNamara M, & Chien N. (2012). Predictors and outcomes of early versus later English language proficiency among English language learners. *Early Childhood Research Quarterly, 27*(1), 1–20. 10.1016/j.ecresq.2011.07.004 [PubMed: 22389551]
- Hammer CS, Jia G, & Uchikoshi Y. (2011). Language and literacy development of dual language learners growing up in the United States: A call for research. *Child Development Perspectives, 5*(1), 4–9. 10.1111/j.1750-8606.2010.00140.x [PubMed: 23259006]
- Hill CJ, Bloom HS, Black AR, & Lipsey MW (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives, 2*(3), 172–177.
- Hill LG, Lochman JE, Coie JD, Greenberg MT & The Conduct Problems Prevention Research Group (2004). Effectiveness of early screening for externalizing problems: Issues of screening accuracy and utility. *Journal of Consulting and Clinical Psychology, 72*, 809–820. 10.1037/0022-006x.72.5.809 [PubMed: 15482039]
- Jellinek MS, Murphy JM, Robinson J, Feins A, Lamb S, & Fenton T. (1988). Pediatric Symptom Checklist: Screening school-age children for psychosocial dysfunction. *The Journal of Pediatrics, 112*(2), 201–209. [PubMed: 3339501]
- Jimerson S, Burns M, & VanDerHeyden A. (2016). *Handbook of response to intervention: The science and practice of multi-tiered systems of support* (2nd ed.) Springer Science + Business Media, New York, NY. 10.1007/978-1-4899-7568-3
- Jones D, Dodge KA, Foster EM, Nix R, and the Conduct Problems Prevention Research Group. (2002). Early identification of children at risk for costly mental health service use. *Prevention Science, 3*, 247–256. 10.1023/a:1020896607298 [PubMed: 12458763]
- Kamphaus RW, Reynolds CR, & Dever BV (2014). Behavioral and mental health screening. In Kettler RJ, Glover TA, Albers CA, Feeney-Kettler KA, Kettler RJ, Glover TA, ... Feeney-Kettler KA (Eds.), *Universal screening in educational settings: Evidence-based decision making for schools* (pp. 249–273). Washington, DC, US. 10.1037/14316-010
- Kamphaus RW, & Reynolds CR (2015). *The Behavioral and Emotional Screening System (BESS)*. Austin, TX. Pearson, Inc.
- Kettler RJ, Feeney-Kettler KA, & Dembitzer L. (2017). Social, emotional, and behavioral screening: A comparison of two measures and two methods across informants. *Journal of School Psychology, 64*, 93–108. 10.1016/j.jsp.2017.05.002 [PubMed: 28735610]
- Landis JR, & Koch GG (1977). The measurement of observer agreement for categorical data. *International Biometric Society, 33*(1), 159–174. 10.2307/2529310
- Lane KL, Little AL, Menzies HM, Lambert W, & Wehby JH (2010). A comparison of students with behavioral challenges educated in suburban and rural settings: Academic, social and behavioral outcomes. *Journal of Emotional and Behavioral Disorders, 18*, 131–148. 10.1177/1063426609332343
- Levitt JM, Saka N, Romanelli LH, Hoagwood K, & Hunter Romanelli L. (2007). Early identification of mental health problems in schools: The status of instrumentation. *Journal of School Psychology, 45*(2), 163–191. 10.1016/j.jsp.2006.11.005
- Lewis K, Sandilos LE, Hammer CS, Sawyer BE, & Méndez LI (2016). Relations among the home language and literacy environment and children’s language abilities: A study of head start dual language learners and their mothers. *Early Education and Development, 27*(4), 478–494. 10.1080/10409289.2016.1082820 [PubMed: 27429533]
- Lilles E, Furlong M, Quirk M, Felix E, Dominguez K, & Anderson M. (2009). Preliminary Development of the Kindergarten Student Entrance Profile. *California School Psychologist, 14*, 71–80. 10.1007/BF03340952

- Liu J, Burgess Y, DiStefano C, Pan F, & Jiang N. (2019). Validating the Pediatric Symptoms Checklist-17 in the preschool environment. *Journal of Psychoeducational Assessment*, 38(4), 460–474. 10.1177/0734282919828234
- Loeber R, Green SM, Lahey BB, & Stouthamer-Loeber M. (1990). Optimal informants on childhood disruptive behaviors. *Development and Psychopathology*, 1, 317–337. 10.1017/s095457940000050x
- McConaughy SH (1993). Advances in empirically based assessment of children's behavioral and emotional problems. *School Psychology Review*, 22, 285–307.
- Milfont TL, & Fischer R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, 3(1), 111–121. 10.21500/20112084.857
- Moore SA, Dowdy E, Nylund-Gibson K, & Furlong M. (2019). An empirical approach to complete mental health classification in adolescents. *School Mental Health*, 11, 438–453. 10.1007/s12310-019-09311-7 [PubMed: 31788132]
- Murphy JM, Bergmann P, Chiang C, Sturmer R, Howard B, Abel MR, & Jellinek M. (2016). The PSC-17: Subscale scores, reliability, and factor structure in a new national sample. *Pediatrics*, 138(3). 10.1542/peds.2016-0038
- Nagle RJ (2000). Issues in preschool assessment. *The psychoeducational assessment of preschool children*, 19–32.
- National Association for the Education of Young Children (2019). NAEYC Early Learning Program Accreditation Standards and Assessment Items. Retrieved July 24, 2020, from https://www.naeyc.org/sites/default/files/globally-shared/downloads/PDFs/accreditation/early-learning/standards_and_assessment_2019.pdf
- Owens JE, Storer J, Holdaway AS, Serrano VJ, Watabe Y, Himawan LK, Krelko RE, & Vause KJ (2015). Screening for social, emotional, and behavioral problems at kindergarten entry: Utility and incremental validity of parent report. *School Psychology Review*, 44, 21–40. 10.17105/SPR44-1.21-40
- Perez Huber L, Malagón MC, Ramirez BR, Gonzalez LC, Jimenez A, & Vélez VN (2015). Still Falling through the Cracks: Revisiting the Latina/o Education Pipeline. CSRC Research Report. Number 19. UCLA Chicano Studies Research Center. Retrieved September 11, 2020, from <https://www.chicano.ucla.edu/files/RR19.pdf>
- Quirk M, Furlong MJ, & Felix E. (2010). Kindergarten Student Entrance Profile Protocol. Santa Barbara, CA: Center for School-Based Youth Development, University of California, Santa Barbara.
- Quirk M, Dowdy E, Goldstein A, & Carnazzo K. (2017). School readiness as a longitudinal predictor of social-emotional and reading performance across the elementary grades. *Assessment for Effective Intervention*, 42(4), 248–253. 10.1177/1534508417719680
- Quirk M, Rebelez J, & Furlong M. (2014). Exploring the dimensionality of a brief school readiness screener for use with Latino/a children. *Journal of Psychoeducational Assessment*, 32, 259–264. 10.1177/0734282913505994
- Quirk M, Mayworm A, Edyburn K, & Furlong M. (2016). Dimensionality and measurement invariance of a school readiness screener by ethnicity and home language. *Psychology in the Schools*, 53, 772–784. 10.1002/pits.21935
- Learning Renaissance. (2009). STAR Early Literacy: Technical manual. Wisconsin Rapids, WI: Renaissance Learning, Inc.
- Reynolds CR, & Kamphaus RW (1992). Behavior Assessment System for Children (BASC). Circle Pines, MN: AGS.
- Rodriguez A, Reise SP, & Haviland M. (2015). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment*, 98(3), 223–237. 10.1080/00223891.2015.1089249 [PubMed: 26514921]
- Rodriguez JC, Kettler RJ, & Feeney-Kettler KA (2017). Behavioral screening with a translated measure: Reliability and validity evidence for the preschool behavior screening system–parent Spanish form. *International Journal of Humanities and Social Science Invention*, 6(1), 11–21.

- Romer NR, Ravitch NK, Tom K, Merrell KW, & Wesley KL (2011). Gender differences in positive social-emotional functioning. *Psychology in the Schools*, 48(10), 958–970. 10.1002/pits
- Shepley C, & Grisham-Brown J. (2019). Multi-tiered systems of support for preschool-aged children: A review and meta-analysis. *Early Childhood Research Quarterly*, 47, 296–308. 10.1016/j.ecresq.2019.01.004
- Smith SR (2007). Making sense of multiple informants in child and adolescent psychopathology: A guide for clinicians. *Journal of Psychoeducational Assessment*, 25, 139–149. 10.1177/0734282906296233
- Stiffler MC, & Dever BV (2015). *Mental health screening at school: Instrumentation, implementation, and critical issues*. New York, NY: Springer.
- Stolarova M, Wolf C, Rinker T, & Briemann A. (2014). How to assess and compare inter-rater agreement and correlation of ratings: An exemplary analysis of mother-father and parent-teacher expressive vocabulary rating pairs. *Frontiers in Psychology*, 5, 1–13. 10.3389/fpsyg.2014.00509 [PubMed: 24474945]
- Taylor HG, Anselmo M, Foreman AL, Schatschneider C, & Angelopoulos J. (2000). Utility of kindergarten teacher judgments in identifying early learning problems. *Journal of Learning Disabilities*, 33, 200–210. 10.1177/002221940003300208 [PubMed: 15505949]
- U.S. Census Bureau. (2017). Current population estimates, vintage 2016. Retrieved from <http://datacenter.kidscount.org/data/tables/103-child-population-by-race>
- von der Embse NP, Iaccarino S, Mankin A, Kilgus SP & Magen E. (2016). Development and validation of the Social, Academic, and Emotional Behavior Risk Screener – Student Rating Scale. *Assessment for Effective Intervention*, 42, 186–192. 10.1177/1534508416679410
- Walker HM, Small JW, Severson HH, Seeley JR, & Feil EG (2014). Multiple-gating approaches in universal screening within school and community settings. In Kettler RJ, Glover TA, Albers CA, & Feeney-Kettler KA (Eds.), *Universal screening in educational settings: Evidence-based decision making for schools* (pp. 47–75). Washington, DC: American Psychological Association. 10.1037/14316-003

Impact Statement

Universal screening for behavioral and emotional risk is important at the preschool level as it provides a critical opportunity for early identification and early intervention. Parents and teachers who participated in universal screening demonstrated low levels of agreement in their ratings of preschool children's emotional and behavioral risk, although both are valuable informants. If the goal of screening is to predict later kindergarten academic and social-emotional outcomes, practitioners may consider collecting screening information from teachers first.

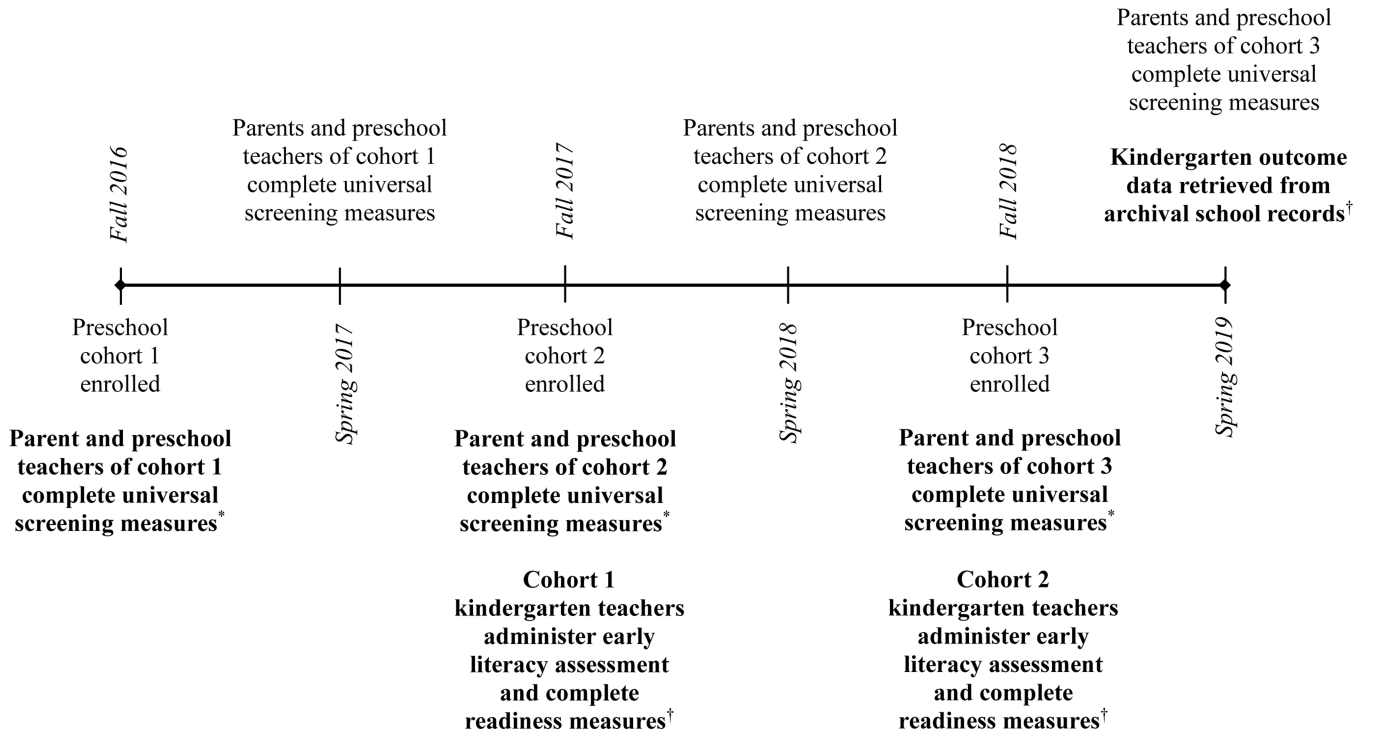


Figure 1. Data collection timeline for longitudinal screening measure study. **Bold** text denotes data that were used to answer this study’s research questions. *cross-sectional data informing parent and teacher correspondence research questions and predicting kindergarten outcomes. †longitudinal data informing predictive utility research question.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Cross-Sectional Sample Demographics

Variable	Parent	Teacher	Child
Gender			
Female	84.9%	92.9%	48.3%
Male	7.6%	7.1%	44.0%
Race/Ethnicity *			
Latinx/Hispanic	82.5%	28.6%	78.4%
White/European	7.8%	50.0%	12.1%
Other	3.9%	14.3%	3.4%
Education Level			
Less than high school	18.8%	–	–
High school diploma	37.5%	–	–
Some college/professional training	16.0%	21.4%	–
College degree	10.1%	50.0%	–
Graduate School	1.5%	21.4%	–
Language of Packet Completed			
English	40.9%	100.0%	–
Spanish	52.6%	–	–
Home Language			
English	25.4%	–	37.3%
Spanish	51.5%	–	33.8%
Both	13.1%	–	19.8%
Number of years teaching			
3–5 years	–	7.1%	–
5–10 years	–	14.3%	–
>10 years	–	64.3%	–
<i>N</i>	535	14	535

Note.

* Participants were given the option to report multiple racial/ethnic categories.

Table 2
 Bivariate Correlations for Teacher and Parent BESS and PSC-17 scores and Kindergarten Outcomes

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. PSC-17 Teacher Total	-													
2. PSC-17 Teacher INT	.66***	-												
3. PSC-17 Teacher EXT	.92***	.46***	-											
4. BESS Teacher BERI	.89***	.55***	.83***	-										
5. BESS Teacher INT	.61***	.81***	.50***	.65***	-									
6. BESS Teacher EXT	.85***	.37***	.86***	.85***	.44***	-								
7. PSC-17 Parent Total	.26***	.15***	.23***	.26***	.16***	.25***	-							
8. PSC-17 Parent INT	.06	.02	.03	.05	.06	.04	.56***	-						
9. PSC-17 Parent EXT	.22***	.13**	.21***	.21***	.14**	.22***	.84***	.24**	-					
10. BESS Parent BERI	.25***	.13**	.21***	.28***	.17***	.24***	.68***	.41***	.58***	-				
11. BESS Parent INT	.12*	.11*	.09	.13**	.16**	.10*	.58***	.48***	.45***	.73***	-			
12. BESS Parent EXT	.30***	.15**	.26***	.28***	.17**	.31***	.65***	.29***	.59***	.78***	.64***	-		
13. STAR Early Literacy	-.16*	-.05	-.17*	-.27**	-.08	-.16*	-.13	.00	-.16*	-.16*	-.05	-.09	-	
14. KSEP AC Readiness	-.19**	-.16**	-.16**	-.25***	-.10	-.14*	-.11	.02	-.09	-.15*	-.03	-.04	.44***	-
15. KSEP SE Readiness	-.44***	-.19**	-.43***	-.47***	-.22***	-.42***	-.09	.00	-.08	-.19**	-.01	-.20**	.32***	.46***

Note. INT = Internalizing, EXT = Externalizing, AC = Academic, SE = Social-Emotional

* $p < .05$

** $p < .01$

*** $p < .001$

Table 3

Results of Dependent Samples t-tests of Continuous Parent and Teacher Screening Scores

	<i>M (SD)</i>		<i>t</i>	<i>df</i>	<i>p</i>	<i>d</i>
	Teacher	Parent				
PSC-17 Total	5.4 (6.0)	6.4 (4.2)	-3.44	474	.001	0.20
PSC-17 Internalizing	1.3 (1.6)	0.7 (1.1)	6.19	473	<.001	0.40
PSC-17 Externalizing	2.0 (2.8)	3.0 (2.3)	-6.29	475	<.001	0.37
BESS BERI	47.3 (9.5)	46.4 (7.8)	1.67	475	.100	0.09
BESS Internalizing	0.4 (0.4)	0.4 (0.3)	-2.84	474	.005	0.16
BESS Externalizing	0.5 (0.6)	0.5 (0.4)	-2.08	472	.038	0.12

Note. $d = .20$ = "small," $d = .50$ = "medium," $d = .80$ = "large" (Cohen, 1988)

Table 4

Percent of Students Identified as Normal or At-Risk by Rater and Measure

A. Teacher BESS BERI and PSC-17 (N = 512)

PSC-17	BESS	
	Normal	At-Risk
Normal	85.00%	4.30%
At-Risk	2.50%	8.20%

B. Parent BESS BERI and PSC-17 (N = 488)

PSC-17	BESS	
	Normal	At-Risk
Normal	93.60%	2.70%
At-Risk	1.60%	2.00%

C. Parent and Teacher BESS BERI (N = 476)

BESS Parent	BESS Teacher	
	Normal	At-Risk
Normal	83.80%	11.30%
At-Risk	3.60%	1.30%

D. Parent and Teacher PSC-17 Total (N = 475)

PSC-17 Parent	PSC-17 Teacher	
	Normal	At-Risk
Normal	86.10%	10.10%
At-Risk	2.70%	1.10%

Summary of Hierarchical Regression Models for BESS BERI with First Gate Teacher Ratings and Second Gate Parent Ratings

Table 5

	STAR Standard Score		KSEP Cognitive Readiness		KSEP Social-Emotional Readiness				
Block 1	$F(4, 186) = 2.62$ $adj R^2 = .04$ $R^2 = .05$		$F(4, 256) = 6.17^*$ $adj R^2 = .07$ $R^2 = .09^*$		$F(4, 256) = 2.26$ $adj R^2 = .02$ $R^2 = .03$				
	B	SEB	β	B	SEB	β			
Parent Packet Completed in Spanish	9.96	15.79	.06	-.38	.16	-.19	.05	.16	.02
Child's Home Language Spanish	-46.58	17.25	-.28*	-.31	.18	-.15	-.11	.17	-.06
Child's Home Language English and Spanish	-27.35	18.18	-.14	.06	.19	.03	.11	.18	.05
Child's Gender is Female	18.20	12.10	.11	.12	.12	.06	.32	.12	.17*
Block 2	$F(5, 185) = 5.15^*$ $adj R^2 = .10$ $R^2 = .07^*$		$F(5, 255) = 7.85^*$ $adj R^2 = .12$ $R^2 = .05^*$		$F(5, 255) = 15.78^*$ $adj R^2 = .22$ $R^2 = .20^*$				
	B	SEB	β	B	SEB	β	B	SEB	β
Parent Packet Completed in Spanish	12.67	15.27	.08	-.36	.16	-.18	.10	.14	.05
Child's Home Language Spanish	-48.49	16.66	-.29*	-.32	.18	-.16	-.14	.15	-.08
Child's Home Language English and Spanish	-32.62	17.61	-.17	.04	.18	.02	.07	.16	.03
Child's Gender is Female	7.95	11.99	.05	.05	.12	.02	.18	.11	.09
BESS Total Teacher	-2.25	0.59	-.27*	-.02	.01	-.22*	-.05	.01	-.46*
Block 3	$F(6, 184) = 4.45^*$ $adj R^2 = .10$ $R^2 = .00$		$F(6, 254) = 6.72^*$ $adj R^2 = .12$ $R^2 = .00$		$F(6, 254) = 13.38^*$ $adj R^2 = .22$ $R^2 = .00$				
	B	SEB	β	B	SEB	β	B	SEB	β
Parent Packet Completed in Spanish	10.93	15.38	.07	-.37	.16	-.18	.09	.14	.05

	STAR Standard Score	KSEP Cognitive Readiness	KSEP Social-Emotional Readiness
Child's Home Language Spanish	-46.98	.18	.15
Child's Home Language English and Spanish	-33.27	.18	.16
Child's Gender is Female	8.35	.12	.11
BESS Total Teacher	-2.11	.01	.01
BESS Total Parent	-0.82	.01	.01

Note: The entries are adjusted R^2 , R^2 , B, standard error (SE) B, and β for each predictor obtained when the variable was entered into corresponding Block. A conservative p value of .01 was used for model statistics to correct for multiple tests.

* $p < .01$.

Summary of Hierarchical Regression Models for BESS BERI with First Gate Parent Ratings and Second Gate Teacher Ratings

Table 6

	STAR Standard Score		KSEP Cognitive Readiness		KSEP Social-Emotional Readiness				
Block 1	$F(4, 186) = 2.62$		$F(4, 256) = 6.17^*$		$F(4, 256) = 2.26$				
	$adj R^2 = .03$		$adj R^2 = .07$		$adj R^2 = .22$				
	$R^2 = .05$		$R^2 = .09^*$		$R^2 = .00$				
	B	SEB	β	B	SEB	β			
Parent Packet Completed in Spanish	9.96	15.79	.06	-.38	.16	-.19	.05	.16	.02
Child's Home Language Spanish	-46.58	17.25	-.28*	-.31	.18	-.15	-.11	.17	-.06
Child's Home Language English and Spanish	-27.35	18.18	-.14	.06	.19	.03	.11	.18	.05
Child's Gender is Female	18.20	12.10	.11	.12	.12	.06	.32	.12	.17*
Block 2	$F(5, 185) = 2.76$		$F(5, 255) = 5.78^*$		$F(5, 255) = 3.79^*$				
	$adj R^2 = .04$		$adj R^2 = .08$		$adj R^2 = .22$				
	$R^2 = .02$		$R^2 = .01$		$R^2 = .17^*$				
	B	SEB	β	B	SEB	β			
Parent Packet Completed in Spanish	7.04	15.79	.04	-.40	.16	-.20	.02	.16	.01
Child's Home Language Spanish	-44.00	17.21	-.26	-.29	.18	-.14	-.10	.17	-.05
Child's Home Language English and Spanish	-29.16	18.11	-.15	.04	.19	.02	.08	.18	.04
Child's Gender is Female	17.77	12.03	.11	.13	.12	.06	.33	.12	.17*
BESS Total Parent	-1.52	0.85	-.13	-.02	.01	-.12	-.03	.01	-.19*
Block 3	$F(6, 184) = 4.45^*$		$F(6, 254) = 6.72^*$		$F(6, 254) = 13.38^*$				
	$adj R^2 = .10$		$adj R^2 = .12$		$adj R^2 = .22$				
	$R^2 = .06^*$		$R^2 = .04^*$		$R^2 = .17^*$				
	B	SEB	β	B	SEB	β			
Parent Packet Completed in Spanish	10.93	15.38	.07	-.37	.16	-.18	.09	.14	.05

	STAR Standard Score	KSEP Cognitive Readiness	KSEP Social-Emotional Readiness
Child's Home Language Spanish	-46.98	.18	.15
Child's Home Language English and Spanish	-33.27	.18	.16
Child's Gender is Female	8.35	.12	.11
BESS Total Parent	-0.82	.01	.01
BESS Total Teacher	-2.11	.01	.01

Note: The entries are adjusted R^2 , R^2 , B, standard error (SE) B, and β for each predictor obtained when the variable was entered into corresponding Block. A conservative p value of .01 was used for model statistics to correct for multiple tests.

* $p < .01$.

Summary of Hierarchical Regression Models for PSC-17 with First Gate Teacher Ratings and Second Gate Parent Ratings

Table 7

	STAR Standard Score		KSEP Cognitive Readiness		KSEP Social-Emotional Readiness	
Block 1	$F(4, 186) = 2.86$		$F(4, 254) = 5.99^*$		$F(4, 254) = 2.39$	
	$adj R^2 = .04$		$adj R^2 = .07$		$adj R^2 = .02$	
	$R^2 = .06$		$R^2 = .09^*$		$R^2 = .04$	
	B	SEB	β	B	SEB	β
Parent Packet Completed in Spanish	6.94	15.38	.04	-.43	.16	-.21*
Child's Home Language Spanish	-47.47	16.83	-.29*	-.23	.17	-.12
Child's Home Language English and Spanish	-29.56	17.63	-.16	.14	.18	.06
Child's Gender is Female	15.78	11.87	.10	.09	.12	.04
				.33	.12	.18*
Block 2	$F(5, 185) = 3.27^*$		$F(5, 253) = 6.68^*$		$F(5, 253) = 13.59^*$	
	$adj R^2 = .06$		$adj R^2 = .10$		$adj R^2 = .20$	
	$R^2 = .02$		$R^2 = .03^*$		$R^2 = .18^*$	
	B	SEB	β	B	SEB	β
Parent Packet Completed in Spanish	6.85	15.23	.04	-.43	.16	-.21*
Child's Home Language Spanish	-49.11	16.69	-.30*	-.25	.17	-.12
Child's Home Language English and Spanish	-31.41	17.48	-.17	.13	.18	.05
Child's Gender is Female	9.15	12.14	.06	.01	.12	.01
PSC-17 Total Teacher	-2.10	0.97	-.16	-.03	.01	-.18*
				-.07	.01	-.43*
Block 3	$F(6, 184) = 2.89$		$F(6, 252) = 5.54^*$		$F(6, 252) = 11.44^*$	
	$adj R^2 = .06$		$adj R^2 = .10$		$adj R^2 = .20$	
	$R^2 = .01$		$R^2 = .00$		$R^2 = .00$	
	B	SEB	β	B	SEB	β
Parent Packet Completed in Spanish	7.43	15.24	.04	-.43	.16	-.21*
				-.03	.14	-.02

	STAR Standard Score	KSEP Cognitive Readiness	KSEP Social-Emotional Readiness					
Child's Home Language Spanish	-48.18	16.71	-29.*	.17	-.12	-.07	.16	-.04
Child's Home Language English and Spanish	-31.38	17.48	-.17	.13	.18	.05	.16	.07
Child's Gender is Female	7.96	12.20	.05	.01	.12	.01	.16	.08
PSC-17 Total Teacher	-1.90	0.99	-.14	-.03	.01	-.18.*	-.07	-.44.*
PSC-17 Total Parent	-1.65	1.66	-.07	.00	.02	.00	.01	.05

Note: The entries are adjusted R^2 , R^2 , B, standard error (SE) B, and β for each predictor obtained when the variable was entered into corresponding Block. A conservative p value of .01 was used for model statistics to correct for multiple tests.

* $p < .01$

Summary of Hierarchical Regression Models for PSC-17 with First-Gate Parent Ratings and Second-Gate Teacher Ratings

Table 8

	STAR Standard Score		KSEP Cognitive Readiness		KSEP Social-Emotional Readiness	
Block 1	$F(4, 186) = 2.86$		$F(4, 254) = 5.99^*$		$F(4, 254) = 2.39$	
	$adj R^2 = .04$		$adj R^2 = .07$		$adj R^2 = .02$	
	$R^2 = .06$		$R^2 = .09^*$		$R^2 = .04$	
	B	SEB	β	B	SEB	β
Parent Packet Completed in Spanish	6.94	15.38	.04	-.43	.16	-.21*
Child's Home Language Spanish	-47.47	16.83	-.29*	-.23	.17	-.12
Child's Home Language English and Spanish	-29.56	17.63	-.16	.14	.18	.06
Child's Gender is Female	15.78	11.87	.10	.09	.12	.04
				.33	.12	.18*
Block 2	$F(5, 185) = 2.70$		$F(5, 253) = 4.87^*$		$F(5, 253) = 2.09$	
	$adj R^2 = .04$		$adj R^2 = .07$		$adj R^2 = .02$	
	$R^2 = .01$		$R^2 = .00$		$R^2 = .00$	
	B	SEB	β	B	SEB	β
Parent Packet Completed in Spanish	7.73	15.35	.05	-.42	.16	-.21*
Child's Home Language Spanish	-46.39	16.81	-.28*	-.24	.18	-.12
Child's Home Language English and Spanish	-29.78	17.58	-.16	.14	.18	.06
Child's Gender is Female	13.18	11.98	.08	.08	.12	.04
PSC-17 Total Parent	-2.31	1.63	-.10	-.01	.02	-.04
				-.01	.02	-.06
Block 3	$F(6, 184) = 2.89$		$F(6, 252) = 5.54^*$		$F(6, 252) = 11.44^*$	
	$adj R^2 = .06$		$adj R^2 = .10$		$adj R^2 = .20$	
	$R^2 = .02$		$R^2 = .03^*$		$R^2 = .17^*$	
	B	SEB	β	B	SEB	β
Parent Packet Completed in Spanish	7.43	15.24	.04	-.43	.16	-.21*
				-.03	.14	-.02

	STAR Standard Score	KSEP Cognitive Readiness	KSEP Social-Emotional Readiness					
Child's Home Language Spanish	-48.18	16.71	-29* -0.25	.17	-.12	-.07	.16	-.04
Child's Home Language English and Spanish	-31.38	17.48	-17	.13	.18	.05	.16	.07
Child's Gender is Female	7.96	12.20	.05	.01	.12	.01	.16	.08
PSC-17 Total Parent	-1.65	1.66	-07	.00	.02	.00	.01	.05
PSC-17 Total Teacher	-1.90	0.99	-14	-.03	.01	-.18*	-.07	-.44*

Note: The entries are adjusted R^2 , R^2 , B, standard error (SE) B, and β for each predictor obtained when the variable was entered into corresponding Block. A conservative p value of .01 was used for model statistics to correct for multiple tests.

* $p < .01$