#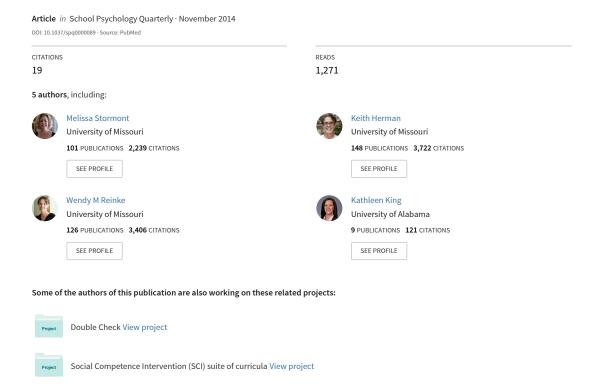 The Kindergarten Academic and Behavior Readiness Screener: The Utility of Single-Item Teacher Ratings of Kindergarten Readiness

5 authors, including:

Melissa Stormont
University of Missouri
101 PUBLICATIONS **2,239** CITATIONS

SEE PROFILE

Keith Herman
University of Missouri
148 PUBLICATIONS **3,722** CITATIONS

SEE PROFILE

Wendy M Reinke
University of Missouri
126 PUBLICATIONS **3,406** CITATIONS

SEE PROFILE

Kathleen King
University of Alabama
9 PUBLICATIONS **121** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   Double Check View project

Project   Social Competence Intervention (SCI) suite of curricula View project

The Kindergarten Academic and Behavior Readiness Screener: The Utility of Single-Item Teacher Ratings of Kindergarten Readiness

Melissa Stormont, Keith C. Herman,
and Wendy M. Reinke
University of Missouri

Kathleen R. King
East Carolina University

Sarah Owens
University of Missouri

The purpose of the study was to explore the effectiveness of a brief, feasible, and cost-effective universal screener for kindergarten readiness. The study examined whether teacher ratings of kindergarteners' academic, behavioral, and overall readiness at the beginning of the year were predictive of academic, emotional, and behavioral outcomes at the end of the year. Participants included 19 kindergarten teachers and their students ($n = 350$) from 6 urban elementary schools; all teachers were female and the majority of children were African American (74%) or White (23%). Thirty-six percent of children qualified for free or reduced lunch. Teachers completed single-item ratings of student readiness as well as full scale ratings of student prosocial skills, disruptive behaviors, and academic competence. Students also completed a standardized academic achievement test. Independent observers rated disruptive behaviors in the classroom. Readiness items had statistically significant relations with a range of academic, emotional, and behavior indicators. Hierarchical linear regression analyses found that readiness items predicted end-of-year outcomes when controlling for baseline covariates. Items also predicted higher likelihood of negative academic and behavior categorical outcomes and demonstrated classification utility. Schools need universal screening options that are feasible and easy to implement school-wide. The screening tool presented in this study offers a viable, psychometrically strong option for school teams and professionals interested in universal screening.

Keywords: mental health problems, screening, tiered models, transition

Given the importance of using prevention models to identify and intervene with children, more research is needed to guide problem solving efforts among school staff (Stormont,

Reinke, Herman, & Lembke, 2012). It is imperative to utilize effective, efficient, and feasible methods to identify and intervene as early as possible (Kilgus, Riley-Tillman, Chafouleas, Christ, & Welsh, 2014; Racz, King, Wu, Witkiewitz, & McMahon, 2013). Many children have mental health issues including behavior and academic problems that impact their ability to be successful in school (Greenwood, Kratochwill, & Clements, 2008; Stormont et al., 2012). Furthermore, many children manifest these problems at school entry, so methods are needed to identify and serve children who are at risk for negative academic and/or social behavioral outcomes.

In particular, universal screening tools are needed to detect risk for academic and behavior problems during the transition to kindergarten. Research has documented that the transition to

kindergarten is challenging for many children (Bowman, Donovan, & Burns, 2000). For a large number of students, the transition to kindergarten denotes the first formal schooling experience. For some the challenge is temporary; for others it is the beginning of a stable pattern of problems (Racz et al., 2013).

Teachers can play a key role in the universal screening process to try to identify children at risk for more persistent problems (Racz et al., 2013). Children's academic, social, and emotional skills provide primary indicators for future school success (Rimm-Kaufman, Pianta, & Cox, 2000). The use of universal screening is even more important in schools with high numbers of children in poverty given their increased risk for failure (e.g., Child Trends, 2010; Racz et al.). Teacher judgments of potential adjustment problems among children provides an opportunity to identify successful and unsuccessful transitions (Rimm-Kaufman et al., 2000). This is a particularly important area for children with internalizing problems as they are not accurately identified using traditional school screening methods. Therefore, the primary purpose of this study was to examine the utility of teacher ratings of kindergarteners' academic, behavioral, and overall readiness at the beginning of the year in predicting academic, emotional, and behavioral outcomes at the end of the year. Although screening measures with strong psychometric properties are available many have limited usability (feasibility, acceptability) because of the resources needed to utilize them. Little research has been conducted on the feasibility of brief universal screening items that require little time and resources (Distefano & Kamphaus, 2007). Within a problem solving framework, school teams need feasible, universal methods to allow early identification of children in need of additional support.

School Readiness

Teachers report that as many as half of all children have difficulty in their transition to kindergarten (Athanasiou, 2006). Further compounding children's problems, kindergarten teachers may believe children should have already learned many essential social skills prior to entering kindergarten (Stormont, Beckner, Mitchell, & Richter, 2005). More specifically, according to a national survey, 46% of kinder-

garten teachers indicated that half or more of their class demonstrated an inability to follow directions when entering school (Wesley & Buysse, 2003), which has been cited by teachers as an indicator of being successful in kindergarten (Stormont et al., 2005). According to Rimm-Kaufman and colleagues, (2000), kindergarten teachers identify weaknesses in academic skills, social skills, ability to follow directions, and difficulty working independently or in a group as significant factors predicting difficult transitions to formal schooling. Research has underscored the importance of teacher ratings in predicting children at risk for failure in school (Racz et al., 2013). For children at risk, recent research supported the need to identify child-level characteristics in kindergarten as they were more predictive of child outcomes than school and classroom-level characteristics (Ball, Finch, & Gettinger, 2014).

Existing literature suggests that early academic and behavior problems are readily identifiable and reliable markers of risk for future academic and social failure. A recent study of 752 students found that teacher ratings of student social behaviors in kindergarten predicted student outcomes over time (Racz et al., 2013). Specifically, high levels of aggression and disruptiveness in kindergarten predicted student adjustment in 6th grade and ADHD diagnosis, externalizing symptoms, and substance use in high school. Other research has documented the significance of co-occurring academic and behavior problems at the start of schooling (Reinke, Herman, Petras, & Ialongo, 2008). In a longitudinal community sample of 678 low-income children, Reinke and colleagues (2008) found that students with co-occurring problems had the poorest long-term outcomes in 6th grade, including academic failure, receipt of special education services, affiliation with deviant peers, suspension from school, and elevated risk for conduct problems. A subsequent study (Darney, Reinke, Herman, Stormont, & Ialongo, 2013) found these problems persisted through 12th grade. Early academic and social problems also serve as independent risk factors for internalizing symptoms (Herman et al., 2009; Herman, Lambert, Ialongo, & Ostrander, 2007; Herman, Lambert, Reinke, & Ialongo, 2008). For instance, Herman and colleagues (2007) found that low academic competence in a community sample of first grade students pre-

dicted depressive symptoms in middle school. Therefore, early screening measures across academic and social domains may help identify children at risk for both behavioral and emotional problems.

Universal Screening for Academic and Behavior Problems

Although many universal academic and behavior screeners exist, each has one or more limitations. Schools have to administer many assessments a year and, although longer measures yield important information, their utility in school settings is dubious (Distefano & Kamphaus, 2007). Thus, there is a need for more efficient, user-friendly screening items that can accurately predict a range of student outcomes over time.

Several universal academic and behavior screening measures have been designed to identify students who would benefit from supplemental and intensive interventions (i.e., Tier 2 or 3 supports). Universal screening in the area of academics has become efficient and streamlined with the growing popularity of manualized screening, benchmarking, and intervention tools such as Aimsweb (Howe & Shinn, 2002) and the Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2002). Using either of these systems, in the fall of kindergarten, students complete screening measures and results are used to assess risk status for future academic difficulties. These curriculum-based measures include 1-min, individually administered probes of key developmental skills in the progression of reading, math, and writing competence. Psychometric data support the adequacy of DIBELS early literacy measures (see Burke, Hagan-Burke, Kwok, & Parker, 2009) and scores are predictive of later academic outcomes (see Goffreda, Diperna, & Pedersen, 2009). Likewise, Aimsweb reading and math measures have evidence to support their quality, and have been found to predict other achievement measures and general school success (see Baglici, Codding, & Tryon, 2010; Clarke, Baker, Smolkowski, & Chard, 2008; Clarke & Shinn, 2004; Martinez, Missall, Graney, Aricak, & Clarke, 2009; National Reading Panel, 2000).

Behavior screening, and universal behavior screening in particular, has lagged behind aca-demic screening. Although many technically adequate behavior screeners have been developed, few districts use them to on a systematic basis to inform problem solving practices, at least in part because existing measures often lack usability or contextual appropriateness (Kilgus et al., 2014). Well-established tools such as the Systematic Screening of Behavior Disorders (SSBD; Walker & Severson, 1992), the Behavioral and Emotional Screening System (BESS; Kamphaus & Reynolds, 2007), the Student Risk Screening Scale (SRSS; Drummond, 1994), and the Teacher Observation of Classroom Adaptation-Revised (TOCA-R; Werthamer-Larsson, Kellam, & Ovesen-McGregor, 1990) all have strong psychometric properties but have limited usability (feasibility, acceptability) because they are time-consuming, and some are expensive.

The Direct Behavior Rating (DBR; Chafouleas, Riley-Tillman, & Christ, 2009) is a recently developed tool designed to address some of the usability and contextual fit concerns of existing behavior screeners (Kilgus et al., 2014). The DBR is a brief measure that utilizes teacher ratings of student behavior after an observation period. After the observation, which can occur concurrently with instruction, teachers are asked to rate student behavior with 10-point, Likert-type items. Teachers typically complete DBR ratings of an entire classroom in about 15 minutes. When the behaviors of academically engaged, disruptive, and respectful were measured with the single-item DBR (DBR-SIS) results were significantly correlated with both the BESS and the SRSS (Chafouleas et al., 2013). The DBR provides support for the technical adequacy of even single-item scales for predicting important school outcomes.

Another widely used source of data with technical adequacy, usability, and contextual fit relies on the systematic collection and use of office discipline referral (ODR) data (McIntosh, Reinke, & Herman, 2010). ODRs are routinely collected by schools, and structured interventions such as Positive Behavior Interventions and Support (PBIS) have programs and materials to help schools collect and use this information in a decision-making framework. This method of identification is appealing given the feasibility of collecting the data as it requires little additional resources and time. Although

ODRs have evidence to support their use of identifying students in need of behavior support, and possibly academic support, little evidence supports their use with internalization symptoms. Further, given that ODRs occur after a problem incident has occurred they are not particularly proactive in finding student who need supports. Moreover, given that ODR rates tend to be very low in kindergarten (McIntosh et al., 2006), it is likely they would be useful only in identifying students with the most extreme risk.

In addition to the screening instruments listed above, other measures of school readiness are often used to determine risk-status upon school entry. The Kindergarten Student Entrance Profile (KSEP) is a school readiness tool that assesses children's readiness according to the areas of physical, health, social emotional, and overall knowledge (Lilles et al., 2009; Lilles, Quirk, & Furlong, 2011). Teachers complete sections of this measure after observing children in their classrooms over a period of time. If children have elevated risk on this tool then other supplemental measures are recommended as well as increased monitoring depending on the level of risk. Studies on the utility of the KSEP have documented that it has concurrent and predictive validity for academic and social emotional readiness (Lilles et al., 2009; Lilles et al., 2011; Quirk, Nylund-Gibson, & Furlong, 2013; Quirk, Rebelez, & Furlong, 2014).

Another tool is the Brigance Comprehensive Inventory of Basic Skills-Revised (CIBS-R; Curriculum Associates, 2010), which is a norm-referenced measure of academic readiness and language skills. Other tools include the Developmental Indicators for the Assessment of Learning, Third Edition (Dial-3; Mardell-Czudnowski & Goldenberg, 1998), and the Early Screening Inventory—Revised (ESI-R; Meisels, Marsden, Wiske, & Henderson, 1997). These measures are all individually administered, often requiring observation of the child in the classroom setting, and, therefore, can be time-consuming. Hence, these measures are typically utilized by school personnel only after children demonstrate some level of academic risk or failure, as a way to determine the necessity of further evaluation.

The Role of Sociodemographic Characteristics

Sex and socioeconomic status are related to academic, emotional, and behavior risk and warrant consideration in any study of screening effectiveness. Boys are at increased risk for having difficulty related to school readiness and school success in general. Zill (1999) indicated that boys are at a disadvantage cognitively and in regards to social/emotional/behavior skills upon school entry. Boys are therefore more likely than girls to earn poorer grades, demonstrate more school-related difficulties related to retention, receive special education services, and drop out of school (Dauber, Alexander, & Entwisle, 1993; McCoy & Reynolds, 1999). These differences are more pronounced in children from low-income families (Hinshaw, 1992).

Approximately 11.7 million low-income children in the United States are impacted by risk factors of poverty suggesting the importance of identifying proximal protective factors to mitigate risk and support the development of skills required for success in school (McWayne, Hahs-Vaughn, Cheung, & Wright, 2012; Wight, Thampi, & Briggs, 2010). Children from low-income families do not fare as well academically as children from more advantaged families and are more likely to arrive in kindergarten underprepared for school success (Bradley & Corwyn, 2002; Hart & Risley, 1995). For instance, young children living in poverty are much less likely than nonpoor children to be able to recognize letters of the alphabet, count to 20 or higher, or write their names, all of which are primary building blocks for continued literacy affecting future school success (Child Trends, 2010).

Purpose

Identifying risks during critical junctures of development (e.g., as children enter school) and intervening to prevent or ameliorate risks can alter the life trajectories of children (Coie & Jacobs, 1993; Kellam & Rebok, 1992). Screening children's overall, behavioral, and academic readiness during their entry into kindergarten presents an opportunity for identifying children unprepared for this significant

life event. Although screeners are available for young children, most focus on either academic risk or behavior risk but not both. Furthermore, many require resources or training and are time-intensive for teachers. It is important to investigate other options for early screeners that are quick, easy to administer, and free.

The purpose of this study is to investigate the predictive utility of a brief transition readiness screener, the Kindergarten Academic Behavior Readiness (K-ABR; Stormont, Reinke, & Herman, 2011) screener. The K-ABR consists of three items that teachers can use to identify children in need of academic and/or behavior support. This study investigates whether the K-ABR items have promise as a screener for identifying children at risk in their transition to kindergarten. Finally, given the co-occurrence of academic and behavior problems with internalizing problems (see Herman et al., 2009), we were interested in determining how well the K-ABR identified students with internalizing symptoms. We began with the following research questions: (a) What percentage of children was rated as ready for kindergarten overall, academically, and behaviorally? (b) Are there group differences on readiness items according to sex and free and reduced lunch status? (c) Are the K-ABR items collected at the beginning of the academic year significantly correlated with other measures of children's readiness for kindergarten (including teacher-rated disruptive problems, concentration problems, prosocial behaviors, and internalizing symptoms; math and reading achievement; and independent observations of disruptive behaviors) gathered at the same time? (d) Do the K-ABR items collected at the beginning of the academic year predict academic and social behavioral outcomes at the end of the year controlling for baseline covariates? (e) Are children rated as poor in terms of their overall, academic, or behavior readiness at the beginning of the year at risk for negative outcomes at the end of the year? And (f) What are the classification patterns for each readiness item for correctly identifying end-of-the-year student academic and behavior outcomes?

Method

Participants

The participants included 19 kindergarten teachers and their students ($n = 350$) from six urban elementary schools serving primarily African American students. All teacher participants and parents of student participants provided written consent, and students provided written assent to participate in the study. All teacher participants were female (100%); 36% reported their race/ethnicity as African American and 64% identified as White. The student sample was composed of 54% males and 46% females. The majority of students were African American (74%) or White (23%). Thirty-six percent of students qualified for free or reduced lunch. Participants were part of a larger sample (in K-3 grades) that participated in a group randomized trial. In this study we focused only on the kindergarten subsample of the teachers who completed kindergarten readiness measures on their students at the start of the trial.

Measures

Demographics. Free and reduced lunch status (FRL) and sex were obtained from the school district for all participating students. Intervention status was coded as 0 or 1 for control and intervention classrooms, respectively.

Kindergarten Academic and Behavior Readiness Screener (K-ABR; Stormont et al., 2011). Teachers completed the K-ABR at the beginning of the academic year (October). The K-ABR asks the teachers to rate each student's academic, behavioral, and overall readiness for kindergarten on a 5-point scale (*poor*, *fair*, *good*, *very good*, and *excellent*). The specific items include: (a) *Compared with other students in this school, how was this child's readiness for the behavioral expectations of kindergarten*, (b) *Compared with other students in this school, how was this child's academic readiness for kindergarten*, and (c) *Compared with other students in this school, how was this child's overall readiness for kindergarten*. The three screening items had high internal consistency (Cronbach's alpha = .86) and intercorrelations.

Teacher ratings of student behaviors. Teachers also rated student behaviors on the *Teacher Observation of Classroom Adaptation-Checklist* (TOCA-C; Koth, Bradshaw, & Leaf,

2009) in October and April of the academic year. The TOCA-C is a 24-item, teacher-completed checklist of student classroom behaviors. The TOCA-C subscales utilized in this study included disruptive behavior, concentration problems, and internalizing problems. The TOCA-C is scored using a 6-point, Likert-type scale with item responses ranging from 1 (*almost never*) to 6 (*almost always*). Data yielded from other studies of the TOCA-C have supported that it has strong evidence of high psychometric functioning. Previous research of the TOCA-C has found internal consistency estimates ranging from .86 to .96. For the current study, the internal consistency (computed using Cronbach alphas) for each subscale ranged from .82 to .96. Moreover, a recent factor analytic study confirmed the factor structure of the TOCA-C in our sample and found it to have strong psychometric properties in terms of goodness of fit and measurement invariance (Wang et al., in press).

Teachers also rated student pro-social behavior using the *Social Competence Scale-Teacher version* (T-COMP; CPPRG, 1995) in October and April. This measure includes a 6-point, Likert-type scale with item responses ranging from 0 (*almost never*) to 5 (*almost always*). Previous research has found adequate psychometric properties for the prosocial subscale (CPPRG, 1995; Gouley, Brotman, Huang, & Shrout, 2008) including moderate correlations with the other prosocial measures, peer relations, and internalizing and externalizing symptoms.

Academic achievement. The *Woodcock-Johnson III Normative Update Tests of Achievement* (WJ; Woodcock, McGrew, & Mather, 2007) is an assessment of student academic achievement. The present study included two subscales, Broad Reading and Broad Math. The WJ is a well-established and widely used standardized achievement test with past studies indicating test-retest reliabilities estimates exceeding .80 and strong intercorrelations with other achievement measures. Teachers also rated student academic competence using the *Social Competence Scale–Teacher version* (T-COMP; CPPRG, 1995). The item responses range from 0 (*almost never*) to 4 (*almost always*). Previous research has found adequate psychometric properties for the T-COMP (CPPRG, 1995; Gouley et al., 2008).

Direct observations of students. Independent observers conducted direct observations of student disruptive behaviors using the Multi-Option Observation System for Experimental Studies (MOOSES; Tapp, Wehby, & Ellis, 1995) interface for hand held computers to gather real time data using the *Student Teacher Classroom Interaction Observation* code (ST-CIO; Reinke & Newcomer, 2010). Before data collection, observers were trained for 2 weeks using videos and practice sessions to an 85% reliability with a master coder. Reliability checks were conducted on 30% of observations and observers received continuing supervision to ensure against observer drift. All observations were conducted early in the school year (October) and at near the end of the school year (April) during academic instruction. Each student was observed for five minutes at each time point. The overall mean percentage agreement across raters for first time point (October) was 88%, and 93% for the second time point (April). A recent study using a subsample of youth with behavior problems (drawn the larger group randomized trial described in the method section of the present study) found that rate of disruptions was sensitive to intervention effects of behavior support planning (Reinke et al., in press). Student disruptive behavior was operationalized as any behavior that interrupts instruction and was recorded as a frequency count. A detailed coding manual with explicit examples and nonexamples was developed and used for training.

Procedures

The study complied with Institutional Review Board policies of the University of Missouri. Teachers and students were recruited in September. Teachers and parents of participating students provided consent, and students provided assent for participation. Data were collected approximately one month into the school year (early October) and near the end of the school year (late April). Teachers completed rating packets on each student. The packet included the scales described above as well as several other assessments. The WJ was administered by personnel trained and supervised by doctoral students in school psychology and counseling psychology. The direct observations were gathered by independent observers blind to study condition.

Analysis

First, descriptive statistics were calculated to determine mean scores on all variables and numbers of students identified within each readiness (K-ABR) domain. Second, intercorrelations and internal consistency of the three K-ABR readiness items were calculated to determine overlap and consistency among the items. Third, intercorrelations of the three K-ABR items with other indicators of readiness in fall and spring of kindergarten were calculated. Fourth, we conducted chi square tests of independence to examine the hypothesis that boys and children on free and reduced lunch would be more likely to be rated as having poor readiness compared with girls and those not on free and reduced lunch, respectively. Fifth, evidence supporting the predictive validity of each K-ABR item was determined through a series of three-step hierarchical regressions with end-of-year outcomes as the dependent variable: (a) baseline covariates [intervention status, sex, and lunch status] were entered first; (b) baseline K-ABR items were added as continuous predictors to determine the unique contributions of each item to explaining outcome variance; (c) baseline scores on the dependent variable were added as a covariate to determine if K-ABR items predicted change over time in the dependent variable. All regression analyses used adjusted standard errors and accounted for clustering of students in classrooms.

Sixth, we calculated the odds ratio of negative end-of-year binary outcomes (15th percentile or worse) given poor readiness ratings at baseline. We used binary outcomes to try to emulate how the screening items would be used in practice to make categorical decisions about who should receive services. Odds ratios are derived from contingency tables and reflect the likelihood of an adverse outcome (e.g., low reading skills at the end of the year) based on the presence of a risk factor (e.g., poor readiness at the start of the year). These analyses quantified the risk of adverse outcomes given a rating of 'poor' readiness in kindergarten and provided another indicator of the practical utility of these items in identifying students at risk. We chose 15th percentile cut scores given that this is a common definition of risk status in school-based tiered models of support (e.g., Stormont et al., 2012). Finally, we examined classification patterns from the contingency tables for

each readiness item to determine overall correct classification rates, sensitivity (probability that the screener correctly identified students with adverse outcomes), and specificity (probability the screener correctly identified students who did not have adverse outcomes) of student academic and behavior end-of-year status. Sensitivity values were calculated by dividing the number of true positives (correctly identified students who had adverse outcomes) by the total number of students who had adverse outcomes. Specificity was calculated by dividing the number of true negatives (correctly identified students without adverse outcomes) by the total number of students who did not have adverse outcomes. We also computed the positive predictive value (probability that all students rated as having poor readiness actually had adverse outcomes) and negative predictive value (probability that all students rated as ready did not have adverse outcomes) for the readiness items. Finally, we calculated overall classification rates by dividing the total number of true positives and true negatives by the total number of students.

Results

Descriptive and Psychometric Data

Overall, teachers rated the overall readiness of 15.7% of kindergarteners as poor. Teachers also gave poor ratings to 20.1% of students on the academic readiness item and to 13.1% on the behavior readiness item. These ratings align well with a tiered approach to service delivery given that most models suggest 15–20% of students might benefit from additional academic or behavioral supports (Stormont et al., 2012).

Assumptions.   For chi square analyses, the key assumptions are independent observations and expected counts above 5 (Yates, Moore, & McCabe, 1999). The design ensured independent observations and data analyses indicated expected counts above 5 for each analysis. Regression analyses also assume independence (again evidenced by the present design) as well as linearity, independence of error variances, and nonmulticollinearity. Scatterplots indicated linear relations between continuous dependent and independent variables. The Durbin-Watson statistic fell within the acceptable range of 1.5–2.5 for nearly all continuous variable compari-

sons with transition items (most falling between 1.9 and 2.1) suggesting that error variances were independent. The one exception concerned the relation between internalizing ratings and transition ratings, which fell at 1.20. To address the problem of potential overlapping error variances we treated internalizing ratings as a categorical outcome (cut at the upper 15th percentile) and conducted logistic regression analyses. Additionally, frequencies of observed disruptive behaviors are most accurately treated as "count" data; thus, for analyses with observed disruptions as an outcome we conducted Poisson regression analyses. Finally, multicollinearity diagnostics for the multiple regression analyses were acceptable as indicated by the Variance Inflation Index less than 3.

Intercorrelations. Although our primary interest was in determining the predictive utility of each item, we wanted to explore the relations among the items as a further indication of psychometric fit and overlap. Table 1 provides intercorrelations among the readiness items and with Time 1 and Time 4 readiness indicators and outcomes. Notably, the correlation between academic and overall readiness had a near perfect correlation ($r = .91$), suggesting that teachers were primarily relying on their perceptions of a student's academic background in their ratings of overall preparedness. Behavior readiness had more moderate correlations with academic ($r = .50$) and overall readiness ($r = .59$), suggesting this was a related but distinct construct.

Further evidence of the convergent and discriminant validity of the academic and behavior items was found when examining unique correlates of the items. For instance, the behavior readiness item had much stronger correlations with baseline disruptive ($r = -.60$) and prosocial behaviors ($r = .74$) compared with the academic item ($r = -.12$ and .34, respectively). On the other hand, the academic item had stronger correlations with baseline reading ($r = .47$), math ($r = .46$), and teacher-rated achievement scores ($r = .79$) compared with the behavior item ($r = .31$, .29, and .58, respectively). Both behavior and academic items had small but significant correlations with internalizing scores ($r = -.21$ and $-.19$, respectively). Notably, academic readiness was unrelated to Time 1 concentration problems ($-.05$), whereas behavior readiness had

Table 1
Intercorrelations Among Study Variables at Baseline and Follow-Up

	1	2	3	4	5	6	7	8	9	10	11
									Time 1 with Time 4		
1. Academic Readiness	—				-.11**	-.54**	-.23**	.32**	.54**	.40**	-.18**
2. Behavior Readiness	.50**	—			-.48**	-.66**	-.21**	.65**	.32**	.27**	-.21**
3. Overall Readiness	.91**	.59**	—		-.16**	-.54**	-.32**	.35**	.50**	.41**	-.19**
4. T-COMP Academic	.79**	.58**	.77**		-.19**	-.66**	-.23**	.44**	.54**	.40**	-.19**
5. TOCA Disruptive	-.12*	-.60**	-.21**	-.34**	(-.78)	.48**	.21**	-.58**	-.09	-.06	.14*
6. TOCA Concentrate	-.05	-.19**	-.07	-.11*	.09	(-.22)	.13*	-.19**	-.03	-.03	.03
7. TOCA Internalizing	-.21**	-.19**	-.24**	-.24**	.31**	.14*	(.64)	-.24**	-.06**	-.03*	.07
8. T-COMP Prosocial	.34**	.74**	.40**	.55**	-.62**	-.23**	-.28**	(.73)	.28**	.21**	-.22**
9. WJ Reading	.47**	.31**	.47**	.47**	.11*	-.09	.05	.26**	(.67)	.52**	-.10
10. WJ Math	.46**	.29**	.45**	.45**	-.09	-.08	.08	.24**	.64**	(.71)	-.08
11. Observed Disruptions	-.18**	-.31**	-.22**	-.25**	.28**	.08	.06	-.30**	-.18**	-.12*	(.04)
						Time 1 with Time 1					

Note. Correlations in parentheses indicate the correlation of a variable at Time 1 with the same variable at Time 4 (e.g., the correlation between Time 1 T-COMP Academic with Time 4 T-COMP Academic is .75). WJ = Woodcock Johnson; TCOMP = Teacher Social Competence; TOCA = Teacher Observation of Classroom Adaptation.
* $p < .01$. ** $p < .05$.

a small and significant relationship with concentration problems ($r = -.19$).

Examination of readiness correlates with end-of-year social and academic outcomes provided further support for the utility of the single items. The behavior item had moderate to strong correlations with behavior outcomes including teacher rated disruptions ($r = -.48$), concentration problems (r $-.66$), and prosocial behaviors ($r = -.65$). The academic item had moderate to strong and statistically significant correlations with reading ($r = .54$) and math outcomes ($r = .40$) as well as concentration problems ($r = -.54$). Both the behavior and the academic readiness items also had small and significant correlations with end-of-year internalizing scores ($r = -21$ and $-.23$) and direct observations of student disruptions ($r = -.21$ and $-.18$).

Sociodemographics. As predicted, sex and free and reduced lunch status were related to readiness items. Boys were significantly more likely to be rated as having poor overall, $\chi^2[1] = 7.49$; $phi = .15$; $p < .05$, academic, $\chi^2[1] = 5.11$; $phi = .12$; $p < .05$, and behavior readiness, $\chi^2[1] = 10.38$; $phi = .17$; $p < .05$, compared with girls. Similarly, children who qualified for free and reduced lunch versus those who did not were more likely to be rated as having poor academic, $\chi^2[1] = 5.36$; $phi = .13$; $p < .05$, and overall readiness, $\chi^2[1] = 6.81$; $phi = .13$; $p < .05$, but not behavior readiness. In most cases, the overall readiness item mimicked the relations observed between the academic readiness item and other variables. Given the high degree of overlap between these items, we dropped the overall readiness item from subsequent analyses in favor of parsimony.

Predictive Utility

Hierarchical regressions. The correlations between readiness items and end-of-year outcomes provided evidence of their screening utility. To further quantify the predictive utility of the single-item measures, we examined how well they predicted continuous outcomes when controlling for baseline covariates including free and reduced lunch status, sex, and treatment status, and baseline scores of target outcome variables. The first two steps in these analyses were conducted to determine whether the readiness items predicted unique variance in outcomes over and above sociodemographic variables. Significant improvements in percentage of variance explained indicated the items contributed above known correlates in explaining outcomes thus justifying their predictive utility over other readily available and efficient indicators. The third step in these analyses added baseline scores of targeted variables to the equation as a more stringent test of utility. Significant improvements of variance explained at this step suggested that the readiness items predicted change over time and development in key outcomes (i.e., the emergence of new symptoms or problems). To reduce the number of analyses we only conducted regressions on academic and behavior readiness items and like outcomes (e.g., we only examined behavior readiness as predictors of behavior-related outcomes). We included regressions with academic readiness predicting concentration given prior literature linking these domains (e.g., Herman et al., 2007).

The results of the regression analyses are presented in Table 2. The academic readiness item significantly predicted reading and math achievement scores, teacher-rated academic competence, concentration problems, and internalizing problems over-and-above demographic covariates. The beta weights for the academic readiness item were in the moderate range ($\beta s = .37–.52$) for all analyses. In the subsequent step with baseline scores on the dependent variable as additional covariates, academic readiness explained a significant amount of variance in academic achievement measures and concentration problems but not internalizing problems.

The behavior readiness item had moderate-sized betas ($\beta s = -.46$ to $-.62$) when predicting disruptive behaviors, prosocial behaviors, and concentration problems, and smaller betas when predicting observed disruptions and internalizing symptoms ($\beta s = -.27$ to $-.31$) with demographic covariates in the model. The behavior readiness item also predicted significant amounts of variance in models with baseline scores on the dependent variables for all analyses except internalizing symptoms.

Odds ratios. Poor ratings on academic and behavior readiness also served as significant risk factors for future negative categorical outcomes (see Table 3). Children who were

Table 2
Hierarchical Regression Analyses Predicting Time 2 Outcomes From Time 1 Covariates and Readiness Items

Time 2 outcome – Time 1 predictor	Models with covariates	R^2 (ΔR^2)	Semi-partial r/OR	β	CI of β/Exp(B)
WJ Reading – Academic Readiness	1: Sex, Lunch, and Intervention	.32 (.28)	.52*	.52*	[.44–.63]
	2: Sex, Lunch, Intervention, and Baseline Score	.54 (.06)	.24*	.28*	[.20–.35]
WJ Math – Academic Readiness	1: Sex, Lunch, and Intervention	.18 (.15)	.39*	.39*	[.31–.48]
	2: Sex, Lunch, Intervention, and Baseline Score	.53 (.02)	.13*	.10*	[.02–.18]
TCOMP Academic Competence Academic Readiness	1: Sex, Lunch, and Intervention	.51 (.41)	.63*	.64*	[.56–.72]
	2: Sex, Lunch, Intervention, and Baseline Score	.62 (.01)	.10*	.17*	[.08–.26]
TOCA Concentration Problems – Academic Readiness	1: Sex, Lunch, and Intervention	.35 (.24)	−.49*	−.50*	[−.40– −.60]
	2: Sex, Lunch, Intervention, and Baseline Score	.37 (.23)	−.48*	−.49*	[−.39– −.59]
TOCA Internalizing Problems – Academic Readiness	1: Sex, Lunch, and Intervention[a]	n/a	.69[a]	−.37*	[.54–.88]
	2: Sex, Lunch, Intervention, and Baseline Score[a]	n/a	.84[a]	−.18	[.63–1.14]
TOCA Internalizing Problems – Behavior Readiness	1: Sex, Lunch, and Intervention[a]	n/a	.74[a]	−.31**	[.57–.95]
	2: Sex, Lunch, Intervention, and Baseline Score[a]	n/a	.95[a]	−.05	[.69–1.33]
TOCA Disruptive Problems – Behavior Readiness	1: Sex, Lunch, and Intervention	.25 (.20)	−.45*	−.46*	[−.33– −.59]
	2: Sex, Lunch, Intervention, and Baseline Score	.54 (.00)	−.05	−.06	[.03– −.17]
TCOMP Prosocial Behavior – Behavior Readiness	1: Sex, Lunch, and Intervention	.46 (.36)	.59*	.61*	[.52–.70]
	2: Sex, Lunch, Intervention, and Baseline Score	.59 (.02)	.11*	.17*	[.03–.31]
TOCA Concentration Problems – Behavior Readiness	1: Sex, Lunch, and Intervention	.48 (.37)	−.61*	−.62*	[−.54– −.71]
	2: Sex, Lunch, Intervention, and Baseline Score	.48 (.34)	−.58*	−.61*	[−.53– −.70]
Observed Disruptions – Behavior Readiness	1: Sex, Lunch, and Intervention[b]	n/a	n/a	−.27*	[−.46– −.07]
	2: Sex, Lunch, Intervention, and Baseline Score[b]	n/a	n/a	−.27*	[−.47– −.07]

Note. All models adjusted for clustering of students in classrooms. OR = odds ratio; CI = confidence interval of beta or odds ratio; WJ = Woodcock Johnson; TCOMP = Teacher Social Competence; TOCA = Teacher Observation of Classroom Adaptation.
[a] Logistic regression.   [b] Poisson Regression.
* $p < .01$.   ** $p < .05$.

rated poor on academic readiness also had elevated end-of-year risk for being in the lowest 15th percentile on WJ Reading and Math scores (odds ratios [OR] = 7.91 and 5.64, respectively), teacher rated academic incompetence (OR = 10.63), concentration problems (OR = 5.19), and internalizing symptoms (OR = 2.19). Poor behavior readiness increased odds for being in the risk category (15th percentile) for disruptive behaviors (OR = 11.04), prosocial behaviors (OR = 13.66), concentration problems (OR = 6.37), internalizing problems (OR = 2.90), and observed disruptions (OR = 2.08) at the end of the year.

Correct classification rates. Finally, we examined classification profiles of each readiness items using several categorical outcomes (15th percentile or worse) as criterion to further characterize their screening utility. As with the risk analyses described above, we cut the readiness items into categorical ratings of poor versus ready. The academic screening item yielded high overall correct classification rates for WJ reading achievement (.82) and teacher-rated academic competence (.85) and had high specificity (.87 for both) and moderate sensitivity (.55 and .61, respectively). The behavior readiness item had similar high rates of overall correct classification of disruptive behaviors (.86) and

Table 3
Odds of Negative Outcomes (15th Percentile) Given Poor Readiness Baseline Rating

Time 2 binary outcome – Time 1 predictor	Odds ratio	Confidence interval
WJ Reading – Academic Readiness	7.91	[4.07–15.36]
WJ Math – Academic Readiness	5.64	[2.93–10.86]
TCOMP Academic Competence Academic Readiness	10.63	[5.28–21.43]
TOCA Concentration Problems – Academic Readiness	5.19	[2.66–10.11]
TOCA Internalizing Problems – Academic Readiness	2.19	[1.08–4.44]
TOCA Internalizing Problems – Behavior Readiness	2.90	[1.35–6.34]
TOCA Disruptive Problems – Behavior Readiness	11.04	[5.30–23.02]
TCOMP Prosocial Behavior – Behavior Readiness	13.66	[5.28–22.94]
TOCA Concentration Problems–-Behavior Readiness	6.37	[3.05–13.29]
Observed Disruptions – Behavior Readiness	2.08	[1.10–3.94]

Note. WJ = Woodcock Johnson; TCOMP = Teacher Social Competence; TOCA = Teacher Observation of Classroom Adaptation.

low prosocial behaviors (.87). The pattern of high specificity (.93 for both) and modest sensitivity (.44 and .49) was also observed for the behavior readiness item. Finally, we examined the utility of combining the academic and behavior readiness ratings for the purpose of screening internalizing symptoms, given that problems with either academic or behavior competence commonly precede internalizing symptoms (Herman et al., 2008). For these analyses, we assigned students a 1 if they received a poor rating on either the academic or behavior readiness rating and a 0 if they did not. The combined readiness rating resulted in an overall correct classification of internalizing risk of .73 with a sensitivity of .50 and the specificity of .78. Given the positive classification pattern we recalculated the odds ratio for identifying students with internalizing symptoms using this combined readiness item. Students rated poor on either readiness item were 3.57 times [1.85–6.76] more likely to have elevated internalizing symptoms at follow-up.

Discussion

The purpose of this study was to investigate the effectiveness of a quick universal screener administered at the beginning of kindergarten to identify students in need of academic or behavior supports. Findings suggest that single-item academic and behavior readiness screening items have concurrent and predictive utility for this purpose. Both screening items had statistically significant relations with a range of academic, emotional, and behavior indicators collected both at the same time point and also with variables collected six months later, at the end of the academic year. Moreover, each item had stronger correlations with like variables; that is, the academic readiness item had stronger relations with most academic-related variables and the behavior readiness item was more closely aligned with behavior outcomes. The correlational findings were bolstered by the use of two independent sources of data to ensure findings were not simply an artifact of source bias: academic achievement test performance and independent observations of disruptions in the classroom. Additionally, items were related to student sex and SES as expected (Bradley & Corwyn, 2002; Stormont et al., 2012; Hart & Risley, 1995).

Subsequent analyses revealed that each readiness item predicted end-of-year outcomes when controlling for baseline covariates. The academic readiness item had moderate relations with WJ reading and math performance and teacher rated academic competence and concentration problems even after controlling for sociodemographic characteristics. Moreover, the academic readiness item explained a significant amount of variance in these outcomes when controlling baseline scores on these measures suggesting readiness ratings were sensitive to the emergence of new symptoms or problems over time, particularly concentration problems. The behavior readiness item also pre-

dicted end-of-year teacher ratings of prosocial behaviors, disruptive behaviors, concentration problems, as well as observed disruptive behaviors when controlling for baseline covariates.

Odd ratios of categorical outcomes supported the predictive utility of the readiness items. Poor readiness ratings at baseline were associated with elevated risk for being in a problem group at the end of the year. Most notably, poor academic readiness increased the risk for low reading performance at the end of the year nearly eightfold, and students rated as having poor behavior readiness were nearly 14 times as likely to have low prosocial behaviors at the end of the year.

Classification profiles suggested that the readiness items have utility as universal screeners. Overall classification rates of end-of-year academic and social behavior outcomes ranged from .82 to .87. These classification rates compare favorably with those observed in past research using longer screeners (Distefano & Kamphaus, 2007; Racz et al., 2013). For instance, Distefano and Kamphaus (2007) reported an overall classification rate of .62 for the 23-item preschool version of the BESS in identifying students with behavior disorders at the same time point. Across all outcomes, the readiness items had very favorable specificity and negative predictive values, whereas sensitivity and positive predictive values were more modest. Given the low false positive rates observed in these analyses and the high odds associated with adverse end-of-year outcomes, children who receive poor readiness ratings should receive further evaluation and support planning. On the other hand, many students (40–50%) who went on to have negative end-of-year outcomes were not rated as having poor readiness. This indicates that other indicators are needed to identify these students. For instance, we used teacher ratings of readiness at a single time point early in the academic year. Asking the readiness items again at midyear or even on a monthly basis may yield even higher positive predictive value as more children at risk for negative end-of-year outcomes may come to the teacher's attention as the year progresses. Additionally, in this study we used a single category (poor) on a 5-point scale to indicate lack of readiness. Future research is needed to explore whether expanding the number of rating categories to include more than one indicator of low readiness may increase the sensitivity of the measures.

The academic and readiness items had small but significant relations with internalizing symptoms. Poor readiness on either scale doubled the risk for elevated internalizing symptoms at six-month follow-up. Moreover, when combined into a single category, children with poor readiness in either domain had 3.5 times the risk of experiencing internalizing symptoms. This combined rating also had adequate overall classification profiles of internalizing problems, again with higher specificity than sensitivity. One limitation of this finding concerns the adequacy of teacher ratings of child internalizing symptoms. Prior research suggests that teachers often miss internalizing problems in students, especially when they have not been trained to recognize these symptoms (Achenbach, McConaughy, & Howell, 1987; Auger, 2004; Kamphaus & Frick, 2002; Loeber et al., 1991), so future research is needed to supplement teacher-ratings with student self-report. Still, given the limited number of quality internalizing screeners, this combined readiness scale warrants further investigation as a quick and efficient screen of kindergarten students at risk for these problems.

Problem Solving Framework

Within a problem solving model, school-based teams can use the academic and behavior readiness items as quick screens of entire kindergarten classrooms at the start of each year (Tilly, 2008). Given their negative outcomes observed in this study, students rated as having poor readiness on either item should be assessed for their need for additional academic or behavior supports. This support could come from the teacher or from other professionals in schools. Small groups of students who need more academic support could be formed, and specific skills that are common that they need support with could be targeted. Plans for individuals who need more behavior support could also be developed including those who need help with concentration problems, prosocial behaviors, internalizing symptoms, and/or disruptive behavior problems. It would also be prudent to investigate why teachers rate children as poor in

their readiness to help guide further assessments and intervention.

Additional rating scales for children who are screened as poor could be used to target specific behavior, emotional, and academic areas in need of targeting and monitoring. Poor academic readiness is a broad area and teams and/or teachers can use brief criterion referenced tests or curriculum based measures to try to target more specific academic skills to target for intervention (e.g., AIMSWEB or DIBELS; Howe & Shinn, 2002; Good & Kaminski, 2002). Given the high risk for negative academic outcomes as assessed by both teachers and academic performance (WJ) for children rated as poor for academic readiness in this study, there is a clear need for targeting the specific skills in which children show deficits. It may be that children have deficits in underlying reading (language) and math (number sense) prerequisite skills early in the year that are in need of developing for children to benefit from instruction (Stormont et al., 2012). In future research, the K-ABR could be utilized later in the school year to determine if children identified early in the year continue to have readiness issues and if the readiness items predict readiness for 1st grade. The items could also be explored as potential progress monitoring tools in tiered models of support.

Overall, there are many screeners that could be utilized to determine children who are at risk for struggling during kindergarten and in later years. However, these measures tend to be less feasible for schools to utilize within a problem solving framework with existing resources. Even existing scales that are very brief create feasibility challenges in practice. For example, one potential barrier to the widespread use of the DBR is that most studies have aggregated 5 to 10 days of ratings to produce adequate psychometric outcomes, which adds time and inconvenience to the screening process. Additionally, DBR ratings require teachers to complete a brief training before using the tool. Also, little is known about how well DBR items in kindergarten predict outcomes over the school year or whether DBR items can be used to identify students with internalizing symptoms.

Single-Item Measures

The idea of using single-item measures of any construct likely challenges the sensibilities of psychometrically minded school psychologists. Traditional psychometric theory emphasizes the value of multiple items; in fact, the classic Spearman-Brown formula includes scale length as a critical determinant of reliability under the assumption that increasing number of items reduces random measurement error. For some constructs and purposes, however, using repetitive and lesser items may introduce other sources of nonrandom error including rater fatigue or frustration (Robins, Hendin, & Trzesniewski, 2001). Moreover, many single-item scales have been developed and validated as measures of global self-esteem, well-being, affect, and life satisfaction (see Robins et al., 2001). For instance, in one study, a single-item measure of stress had practical utility for screening mental health well-being comparable with longer stress scales (Elo, Leppanen, & Jahkola, 2003). When single-item scales are shown to correlate with related constructs in expected ways and to predict meaningful outcomes, as found with the readiness items in this study, it suggests that they provide an acceptable balance between practical and psychometric concerns (i.e., that they are technically adequate). Further inquiry into single-item assessments in educational settings may uncover other useful tools for meeting the real world demands and needs of schools.

Limitations

The study involved participants from a large, urban school district with a high percentage of African American students so it is not known how well findings will generalize to other students in other school settings. The screener items were also contextualized with the stem *compared with other students at this school*, which can be viewed as strength and a limitation. On the one hand, the stem may limit the expectations to the culture of a school; on the other hand, the stem captures naturally occurring variability in school expectations and perceptions, which may be an essential aspect of its predictive utility. Also, only a small number of teachers were included. The readiness items and the majority of the outcome variables were

teacher-rated, and this raises some potential bias issues. However, the academic achievement test and direct observations of disruptive behavior also provided some validity information for the readiness items and were administered by independent research assistants, who were not aware of the teacher ratings. Finally, future research should explore the use of this screener with a more representative sample of children. It is likely that some schools will have more children who are considered at risk, whereas others will have less. Future research should explore this readiness screener in high and low performing schools to determine the percentages of children who are rated as poor in their readiness for kindergarten as compared with other children their age rather than other children at that school.

Conclusion

Children with early academic and behavior difficulties are at risk for a range of negative outcomes including retention, ongoing academic and behavioral difficulties, and peer rejection (Reinke et al., 2008; Stormont et al., 2012). Children who are living in poverty are at higher risk for struggling in their transition to kindergarten and are more likely to have academic and behavior deficits that likely interfere with their success. Boys are also at increased risk for not meeting the social behavior and academic expectations required at school entry. Teachers are prime candidates to accurately identify children at risk through screening and early identification methods. Having brief (less than one minute) screening options that teachers can use with all incoming kindergarten students can be helpful toward early identification of students who would benefit from additional academic or behavioral supports.

References

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

Athanasiou, M. S. (2006). It takes a village: Children's transition to kindergarten. *School Psychology Quarterly, 21,* 468–473. http://dx.doi.org/ 10.1037/h0084133

Auger, R. W. (2004). The accuracy of teacher reports in the identification of middle school students with depressive symptomatology. *Psychology in the Schools, 41,* 379–389. http://dx.doi.org/10.1002/ pits.10164

Baglici, S. P., Codding, R., & Tryon, G. (2010). Extending the research on the tests of early numeracy: Longitudinal analyses over two school years. *Assessment for Effective Intervention, 35,* 89–102. http://dx.doi.org/ 10.1177/1534508409346053

Ball, C. R., Finch, W. H., & Gettinger, M. (2014). Classroom-level effects on the reading and behavior of at-risk kindergarten teachers. *Preventing School Failure: Alternate Education for Children and Youth, 58,* 80–89.

Bowman, B. T., Donovan, M. S., & Burns, M. S. (2000). *Eager to learn: Educating our preschoolers.* Washington, DC: National Academies Press.

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology, 53,* 371–399. http://dx.doi .org/10.1146/annurev.psych.53.100901.135233

Burke, M. D., Hagan-Burke, S., Kwok, O., & Parker, R. (2009). Predictive validity of early literacy indicators from the middle of kindergarten to second grade. *The Journal of Special Education, 42,* 209–226. http://dx.doi.org/10.1177/0022466907313347

Chafouleas, S. M., Kilgus, S. P., Jaffery, R., Riley-Tillman, T. C., Welsh, M., & Christ, T. J. (2013). Direct behavior rating as a school-based behavior screener for elementary and middle grades. *Journal of School Psychology, 51,* 367–385. http://dx .doi.org/10.1016/j.jsp.2013.04.002

Chafouleas, S. M., Riley-Tillman, T. C., & Christ, T. J. (2009). Direct Behavior Rating (DBR): An emerging method for assessing social behavior within a tiered intervention system. *Assessment for Effective Intervention, 34,* 195–200. http://dx.doi .org/10.1177/1534508409340391

Child Trends. (2010). *Early school readiness.* Retrieved from http://www.childtrendsdatabank.org/ ?q=node/291

Clarke, B., Baker, S., Smolkowski, K., & Chard, D. J. (2008). An analysis of early numeracy curriculum-based measurement examining the role of growth in student outcomes. *Remedial and Special Education, 29,* 46–57. http://dx.doi.org/10.1177/ 0741932507309694

Clarke, B., & Shinn, M. R. (2004). A preliminary investigation into the identification and development of early mathematics curriculum-based measurement. *School Psychology Review, 33,* 234–248.

Coie, J. D., & Jacobs, M. R. (1993). The role of social context in the prevention of conduct disorder. *Development and Psychopathology, 5,* 263–275. http://dx .doi.org/10.1017/S0954579400004387

Conduct Problems Prevention Research Group (CPPRG). (1995). *Teacher – Social Competence Scale*. Available from the Fast Track Project Website http://www.fasttrackproject.org

Curriculum Associates. (2010). *Brigance Comprehensive Inventory of Basic Skills II: Standardized*. North Billerica, MA: Author.

Darney, D., Reinke, W. M., Herman, K. C., Stormont, M., & Ialongo, N. S. (2013). Children with co-occurring academic and behavior problems in first grade: Distal outcomes in twelfth grade. *Journal of School Psychology, 51,* 117–128. http://dx.doi.org/10.1016/j.jsp.2012.09.005

Dauber, S. L., Alexander, K. L., & Entwisle, D. R. (1993). Characteristics of retainees and early precursors or retention grade: Who is held back. *Merrill-Palmer Quarterly, 39,* 326–343.

Distefano, C. A., & Kamphaus, R. W. (2007). Development and validation of a behavioral screener for preschool-age children. *Journal of Emotional and Behavioral Disorders, 15,* 93–102. http://dx.doi.org/10.1177/10634266070150020401

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

Elo, A. L., Leppänen, A., & Jahkola, A. (2003). Validity of a single-item measure of stress symptoms. *Scandinavian Journal of Work, Environment & Health, 29,* 444–451. http://dx.doi.org/10.5271/sjweh.752

Goffreda, C. T., Diperna, J. C., & Pedersen, J. A. (2009). Preventive screening for early readers: Predictive validity of the dynamic indicators of basic early literacy skills (DIBELS). *Psychology in the Schools, 46,* 539–552. http://dx.doi.org/10.1002/pits.20396

Good, R., & Kaminski, R. (2002). *DIBELS oral reading fluency passages for first through third grades* (Tech. Rep. No. 10). Eugene, OR: University of Oregon.

Gouley, K. K., Brotman, L. M., Huang, K. Y., & Shrout, P. E. (2008). Construct validation of the social competence scale in preschool-age children. *Social Development, 17,* 380–398. http://dx.doi.org/10.1111/j.1467-9507.2007.00430.x

Greenwood, C. R., Kratochwill, T. R., & Clements, M. (Eds.). (2008). *Schoolwide prevention models: Lessons learned in elementary schools*. New York, NY: Guilford Press.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes.

Herman, K. C., Lambert, S. F., Ialongo, N. S., & Ostrander, R. (2007). Academic pathways between attention problems and depressive symptoms among urban African American children. *Journal of Abnormal Child Psychology, 35,* 265–274. http://dx.doi.org/10.1007/s10802-006-9083-2

Herman, K. C., Lambert, S. F., Reinke, W. M., & Ialongo, N. S. (2008). Low academic competence in first grade as a risk factor for depressive cognitions and symptoms in middle school. *Journal of Counseling Psychology, 55,* 400–410. http://dx.doi.org/10.1037/a0012654

Herman, K. C., Reinke, W., Parkin, J., Traylor, K., & Agarwal, G. (2009). Childhood depression: Rethinking the role of the school. *Psychology in the Schools, 46,* 433–446. http://dx.doi.org/10.1002/pits.20388

Hinshaw, S. P. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms. *Psychological Bulletin, 111,* 127–155. http://dx.doi.org/10.1037/0033-2909.111.1.127

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

Kamphaus, R. W., & Frick, P. J. (2002). *Clinical assessment of child and adolescent personality and behavior* (2nd ed.). Boston, MA: Allyn & Bacon.

Kamphaus, R. W., & Reynolds, C. R. (2007). *Behavioral & Emotional Screening System*. Bloomington, MN: NCS Pearson.

Kellam, S. G., & Rebok, G. W. (1992). Building developmental and etiological theory through epidemiologically based preventive intervention trials. In J. McCord & R. Ernest (Eds.), *Preventing antisocial behavior: Interventions from birth through adolescence* (pp. 162–195). New York, NY: Guilford Press.

Kilgus, S. P., Riley-Tillman, T. C., Chafouleas, S. M., Christ, T. J., & Welsh, M. E. (2014). Direct behavior rating as a school-based behavior universal screener: Replication across sites. *Journal of School Psychology, 52,* 63–82. http://dx.doi.org/10.1016/j.jsp.2013.11.002

Koth, C. W., Bradshaw, C. P., & Leaf, P. J. (2009). Teacher Observation of Classroom Adaptation—Checklist: Development and factor structure. *Measurement and Evaluation in Counseling and Development, 42,* 15–30. http://dx.doi.org/10.1177/0748175609333560

Lilles, E., Furlong, M. J., Quirk, M., Felix, E., Dominguez, K., & Anderson, M. (2009). Preliminary development of the Kindergarten Student Entrance Profile. *California School Psychologist, 14,* 71–80. http://dx.doi.org/10.1007/BF03340952

Lilles, E., Quirk, M., & Furlong, M. (February, 2011). *Kindergarten Student Entrance Profile*. Paper presented at the annual meeting of the National Association of School Psychologists, San Francisco, CA.

Loeber, R., Green, S. M., Lahey, B. B., & Stouthamer-Loeber, M. (1991). Differences and similarities between children, mothers, and teachers as informants on disruptive child behavior. *Journal of Abnormal Child Psychology, 19,* 75–95. http://dx.doi.org/10.1007/BF00910566

Mardell-Czudnowski, C., & Goldenberg, D. S. (1998). *Developmental indicators for the assessment of learning.* Monterey, CA: American Guidance Service.

Martinez, R. S., Missall, K. N., Graney, S. B., Aricak, O. T., & Clarke, B. (2009). Technical adequacy of early numeracy curriculum-based measurement in kindergarten. *Assessment for Effective Intervention, 34,* 116–125. http://dx.doi.org/10.1177/1534508408326204

McCoy, A. R., & Reynolds, A. J. (1999). Grade retention and school performance: An extended investigation. *Journal of School Psychology, 37,* 273–298. http://dx.doi.org/10.1016/S0022-4405(99)00012-6

McIntosh, K., Horner, R. H., Chard, D. J., Boland, J. B., & Good, R. H. (2006). The use of reading and behavior screening measures to predict nonresponse to school-wide positive behavior support: A longitudinal analysis. *School Psychology Review, 35,* 275.

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

McWayne, C. M., Hahs-Vaughn, D. L., Cheung, K., & Wright, L. E. G. (2012). National profiles of school readiness skills for Head Start children: An investigation of stability and change. *Early Childhood Research Quarterly, 27,* 668–683. http://dx.doi.org/10.1016/j.ecresq.2011.10.002

Meisels, S. J., Marsden, D. B., Wiske, M. S., & Henderson, L. W. (1997). *Early Screening Inventory – Revised. Examiner's manual.* Ann Arbor, MI: Rebus.

National Reading Panel (U.S.), National Institute of Child Health, & Human Development (U.S.). (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups.* Bethesda, MD: National Institute of Child Health and Human Development, National Institutes of Health.

Quirk, M., Nylund-Gibson, K., & Furlong, M. J. (2013). Exploring patterns of Latino/a children's school readiness at kindergarten entry and their relations with grade 2 achievement. *Early Childhood Research Quarterly, 28,* 437–449. http://dx.doi.org/10.1016/j.ecresq.2012.11.002

Quirk, M., Rebelez, J., & Furlong, M. J. (2014). Exploring the dimensionality of a brief school readiness screener for use with Latino/a children. *Journal of Psychoeducational Assessment, 32,* 259–264. http://dx.doi.org/10.1177/0734282913505994

Racz, S. J., King, K. M., Wu, J., Witkiewitz, K., & McMahon, R. J. (2013). The predictive utility of a brief kindergarten screening measure of child behavior problems. *Journal of Consulting and Clinical Psychology, 81,* 588–599. http://dx.doi.org/10.1037/a0032366

Reinke, W. M., Herman, K. C., Petras, H., & Ialongo, N. S. (2008). Empirically derived subtypes of child academic and behavior problems: Co-occurrence and distal outcomes. *Journal of Abnormal Child Psychology, 36,* 759–770. http://dx.doi.org/10.1007/s10802-007-9208-2

Reinke, W. M., & Newcomer, L. (2010). *Student Teacher Classroom Interaction Observation (ST-CIO).* Columbia, MO: University of Missouri.

Reinke, W. M., Stormont, M., Herman, K. C., Wang, Z., Newcomer, L., & King, K. (in press). Use of coaching and behavior support planning for students with disruptive behavior within a universal classroom management program. *Journal of Emotional and Behavioral Disorders.*

Rimm-Kaufman, S. E., Pianta, R. C., & Cox, M. J. (2000). Teachers' judgments of problems in the transition to kindergarten. *Early Childhood Research Quarterly, 15,* 147–166. http://dx.doi.org/10.1016/S0885-2006(00)00049-1

Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin, 27,* 151–161. http://dx.doi.org/10.1177/0146167201272002

Stormont, M., Beckner, R., Mitchell, B., & Richter, M. (2005). Supporting successful transition to kindergarten: General challenges and specific implications for students with problem behavior. *Psychology in the Schools, 42,* 765–778. http://dx.doi.org/10.1002/pits.20111

Stormont, M., Reinke, W. M., & Herman, K. C. (2011). *The Kindergarten Academic Behavior Readiness Screener (K-ABR).* Columbia, MO: University of Missouri.

Stormont, M., Reinke, W., Herman, K., & Lembke, E. (2012). *Tier two interventions: Academic and behavior supports for children at risk for failure.* New York, NY: Guilford Press.

Tapp, J., Wehby, J., & Ellis, D. (1995). A multiple option observation system for experimental studies: MOOSES. *Behavior Research Methods, Instruments & Computers, 27,* 25–31. http://dx.doi.org/10.3758/BF03203616

Tilly, W. D. (2008). The evolution of school psychology to science-based practice: Problem solving and the three-tiered model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 17–36). Washington, DC: National Association of School Psychologists.

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

Wang, Z., Rohrer, D., Chuang, C., Fujiki, M., Reinke, W. M., & Herman, K. C. (in press). Five methods to score TOCA-C and to examine group differences. *Journal of Experimental Education*.

Werthamer-Larsson, L., Kellam, S., & Ovesen-McGregor, K. (1990). Teacher interview: Teacher Observation of Classroom Adaptation—Revised (TOCA-R). In *Johns Hopkins Prevention Center training manual*. Baltimore, MD: Johns Hopkins University.

Wesley, P. W., & Buysse, V. (2003). Making meaning of school readiness in schools and communities. *Early Childhood Research Quarterly, 18,* 351–375. http://dx.doi.org/10.1016/S0885-2006(03)00044-9

Wight, V., Thampi, K., & Briggs, J. (2010). *Who are America's poor children? Examining food insecurity among children in the United States*. New York, NY: Columbia University, National Center for Children in Poverty Mailman School of Public Health.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2007). *Woodcock-Johnson III Tests of Achievement*. Rolling Meadows, IL: Riverside Publishing.

Yates, D., Moore, D., & McCabe, G. (1999). *The practice of statistics* (1st ed.). New York, NY: W. H. Freeman.

Zill, N. (1999). Promoting educational equity and excellence in kindergarten. In R. C. Pianta & M. J. Cox (Eds.), *The transition to kindergarten* (pp. 67–105). Baltimore, MD: Brookes.

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at http://notify.apa.org/ and you will be notified by e-mail when issues of interest to you become available!