Using Latent Class Analysis to Identify Academic and Behavioral Risk Status in Elementary Students

Article *in* School Psychology Quarterly · March 2015

DOI: 10.1037/spq0000111 · Source: PubMed

3 authors:

Kathleen King
University of Alabama

Erica Lembke
University of Missouri

Wendy M Reinke
University of Missouri

School Psychology Quarterly

Using Latent Class Analysis to Identify Academic and Behavioral Risk Status in Elementary Students

Kathleen R. King, Erica S. Lembke, and Wendy M. Reinke

Online First Publication, March 9, 2015. http://dx.doi.org/10.1037/spq0000111

CITATION

King, K. R., Lembke, E. S., & Reinke, W. M. (2015, March 9). Using Latent Class Analysis to Identify Academic and Behavioral Risk Status in Elementary Students. *School Psychology Quarterly*. Advance online publication. http://dx.doi.org/10.1037/spq0000111

School Psychology Quarterly

1045-3830/15/$12.00 http://dx.doi.org/10.1037/spq0000111

Using Latent Class Analysis to Identify Academic and Behavioral Risk Status in Elementary Students

Kathleen R. King
East Carolina University

Erica S. Lembke and Wendy M. Reinke
University of Missouri

Identifying classes of children on the basis of academic and behavior risk may have important implications for the allocation of intervention resources within Response to Intervention (RTI) and Multi-Tiered System of Support (MTSS) models. Latent class analysis (LCA) was conducted with a sample of 517 third grade students. Fall screening scores in the areas of reading, mathematics, and behavior were used as indicators of success on an end of year statewide achievement test. Results identified 3 subclasses of children, including a class with minimal academic and behavioral concerns (Tier 1; 32% of the sample), a class at-risk for academic problems and somewhat at-risk for behavior problems (Tier 2; 37% of the sample), and a class with significant academic and behavior problems (Tier 3; 31%). Each class was predictive of end of year performance on the statewide achievement test, with the Tier 1 class performing significantly higher on the test than the Tier 2 class, which in turn scored significantly higher than the Tier 3 class. The results of this study indicated that distinct classes of children can be determined through brief screening measures and are predictive of later academic success. Further implications are discussed for prevention and intervention for students at risk for academic failure and behavior problems.

Keywords: academic screening, behavior problems, latent class analysis

The co-occurrence of behavioral and academic difficulties in students, and the negative outcomes associated with these difficulties, has been documented (see Darney, Reinke, Herman, Stormont, & Ialongo, 2013; Reinke, Herman, Petras, & Ialongo, 2008), including earlier (Rutter & Yule, 1970) and more recent (Bub, McCartney, & Willett, 2007; Lane, Barton-Arwood, Nelson, & Wehby, 2008) studies. Despite research demonstrating that students with co-occurring academic and behavior problems may be at increased risk for long-term negative academic and behavioral outcomes, schools often compartmentalize academics and behavior. Schools using universal screening often screen for one or the other, but less often look at academic and behavioral risk in combination (Cook, Volpe, & Livanis, 2010; Romer & McIntosh, 2005). One reason for this may be that schools are often focused on screening for specific skill deficits (e.g., oral reading fluency) while failing to consider other factors that may be contributing to poor student outcomes (e.g., inattention; Cook et al., 2010). Researchers are now advocating for universal screening at the "whole child" level (Cook et al., 2010; Kalberg, Lane, & Menzies, 2010); however, questions remain about the utility of using a combination of screening measures in reading, mathematics, and behavior to identify students who may be

Editor's Note. Shane Jimerson served as the action editor for this article.

Kathleen R. King, Department of Psychology, East Carolina University; Erica S. Lembke, Department of Special Education, University of Missouri; Wendy M. Reinke, Department of Educational, School and Counseling Psychology, University of Missouri.

The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. The research reported here was supported by the Institute of Educational Sciences, U.S. Department of Education, through Grants R305A100342, R305A130375, and R305A130143 to the third author.

Correspondence concerning this article should be addressed to Kathleen R. King, Department of Psychology, East Carolina University, 104 Rawl Building, Greenville, NC 27858-4353. E-mail: kingkath14@ecu.edu

at risk across the interrelated domains of behavior and academics.

Co-Occurrence of Behavioral and Academic Need

When asked about the needs of their students at risk, teachers often remark that it is difficult to discriminate whether a student has behavioral issues that are impeding academics or if lack of success in academics is prompting behavioral issues (the classic 'chicken or the egg' scenario, if you will). Although most teacher speculation is largely based on observation, there is a growing body of research, dating back to the 1970s, to support the link between behavior and academics (e.g., Rutter & Yule, 1970). When examining comorbidity rates, academic difficulties and disruptive behaviors have been found to co-occur significantly more often than would have happened by chance (Hinshaw, 1992). Several studies, examining reading specifically, have found that struggling readers have significantly higher rates of behavior difficulties than nonstruggling readers (Berger, Yule, & Rutter, 1975; Heiervang, Lund, Stevenson, & Hugdahl, 2001; Rutter & Yule, 1970; Willcutt & Pennington, 2000). Academic deficits have been found to go beyond reading, however. Students with emotional behavior disorders (EBD) have been found to have broad deficits across all academic areas, including vocabulary, listening comprehension, spelling, social studies, mathematics, science, and standardized achievement tests (Nelson, Benner, Lane, & Smith, 2004; Reid, Gonzalez, Nordness, Trout, & Epstein, 2004).

The factors associated with academic and behavioral risk can be identified early in a child's life and are predictive of later skill deficit. Young children who show behavioral difficulties at 2 years of age are more likely to struggle academically in 1st grade (Bub, McCartney, & Willett, 2007). In addition to the early onset of skill deficits, the academic deficits of children with EBD tend to worsen over time, and decrease to a level below that of peers with learning disabilities (LD; Anderson, Kutash, & Duchnowski, 2001). Anderson and colleagues (2001) found that while children with EBD performed higher than children with LD in reading and math in kindergarten, by the sixth grade the children with EBD were performing significantly lower academically than those with LD. The children in the LD group made improvements in academic skills over the course of the 5 years of the study, whereas those with EBD did not (Anderson et al., 2001). Findings over time suggest that the academic deficits of children with behavior difficulties stabilize and may even intensify (Nelson et al., 2004; O'Shaughnessy, Lane, Gresham, & Beebe-Frankenberger, 2003).

Researchers have proposed several hypotheses for the interaction of academic and behavioral difficulties. One explanation of the link is the moderating and mediating effects of co-occurring attention problems associated with Attention Deficit Hyperactivity Disorder (ADHD). Findings that attention problems may account for much of the relationship between academic achievement and delinquency support this theory (Fleming, Harachi, Cortes, Abbott, & Catalano, 2004). Other explanations include the moderating effects of disruptive behavior in which higher disruptive behavior leads to less time academically engaged (Dishion, French, & Patterson, 1995), or the function-based theory of poor behavior as an escape from aversive and challenging academic tasks (Lee, Sugai, & Horner, 1999; McIntosh, Horner, Chard, Dickey, & Braun, 2008). For many students the question of which came first, maladaptive behavior or academic deficits, is not clear.

The relationship between academic and behavior problems is complex and has yet to be fully explained. Reinke and colleagues (2008) argue that a more useful examination of the relationship between behavior and academics is by identifying subclasses of children based on the presentation of symptoms in both areas. This allows for interventions that target both academic and social behavioral deficits when needed. Using a person-centered, latent variable approach, the authors classified children into optimal grouping categories based on broad academic and behavioral indicators. Results of the class analysis indicated that there were four latent classes within the male sample and three within the female sample. Although the majority of the males in the sample had no behavior or academic problems (59%), a total of 14% of the sample were in a class with both academic and behavioral problems. A smaller class contained male students with only academic problems (11%) and a larger class contained boys with only behavioral problems (16%). The female sample was generally consistent with the male sample; however, there was no class for behavior problems only. More important than the discovery of these classes is the prediction these classes made for academic success several years later. The boys with behavior only and academic and behavior problems were several times more likely to be referred to special education, have poor or failing grades, be suspended from school, and be identified by 6th grade teachers as having high levels of conduct problems. The results for girls were similar, except that those with academic problems only were also more likely to have negative outcomes in 6th grade (Reinke et al., 2008). Further, the co-occurring groups for both boys and girls persisted in having the most negative outcomes when followed into Grade 12, including being at increased risk for dropping out of school (Darney et al., 2013). These studies identified a subgroup of children in first grade for whom there was later and lasting lower academic achievement, based on both academic and behavior indicators. Although these studies demonstrated that subgroups of students with co-occurring academic and behavior problems exist, the measures used were research driven and not readily available to schools for use in screening students. The current study investigates whether commercially available academic and behavioral screening tools used by schools can identify subgroups of students at the start of the school year.

Benefit of Using Screening Measures to Detect Early Difficulties

Identifying children at-risk for behavioral and academic difficulties is a growing area of interest for practitioners and researchers alike. Since the inception of No Child Left Behind (NCLB), and tiered-intervention models such as MTSS and Response to Intervention (RTI), identification of students at risk and early intervention have become the primary focus of assessment efforts (Cook et al., 2010; Glover & Albers, 2007). Universal screening within tiered models allows school personnel to quickly and accurately determine student need, and is central to the early implementation of intervention and prevention efforts necessary for student success (Benner, Kutash, Nelson, & Fisher, 2013; Gresham, 2007; Oakes, Lane, Cox, & Messenger, 2014).

Despite growing evidence of the relationship between behavior and academics, school practices typically include screening, and subsequent intervention, within problem/deficit areas only (Cook et al., 2010). Tiered-models of service delivery, like the one proposed by Lane and colleagues (2014), integrate screening and intervention to collaboratively address multiple areas at once (Collier-Meek, Fallon, Sanetti, & Maggin, 2013). Incorporating components of RTI and MTSS, this approach begins with systematic screening across domains of student functioning, including behavior, academics, and social skills (Lane et al., 2014).

Traditional identification methods of at-risk students, like teacher referral, failing grades, and office discipline referrals, are problematic methods of assessing student risk status (Eklund et al., 2009; Lane, Gresham, & O'Shaughnessy, 2003), and are inconsistent with tiered-models of intervention and prevention. These methods are considered reactionary, termed "wait-to-fail," because they rely on some level of disruptive behavior or failing school grades before identification and subsequent intervention can occur. Thus, developing methods to quickly and accurately identify at-risk children is now of substantial concern to many school districts (Glover & Albers, 2007).

Measures to Detect Early Academic Problems

While teachers are encouraged to adopt measures that are technically adequate for use in detecting early academic problems, they also need measures that are efficient to administer and relatively easy to score. Curriculum Based Measurement (CBM) tools fit these needs because they are efficient to administer, short in duration, technically adequate, and the measures serve as indicators of academic proficiency (Fuchs, 2004). The term "indicators" is used to signify that the short duration (usually 1 to 3 minute) measures are strongly related to other measures of academic proficiency in that content area. CBM measures essentially serve as proxies for academic proficiency. Although a common measure of CBM in reading is the number of words read correctly in one minute (oral reading fluency or RCBM), this measure serves as a broader indicator of academic proficiency in reading. CBM measures are available in a variety of academic areas including reading, writing, spelling, mathematics, and content-area learning. Several web-based systems for utilizing CBM measures for schoolwide use have been developed. One such system is AIMSweb. Created by researchers with roots in CBM, AIMSweb software is used broadly across the United States. The AIMSweb system provides measures in all academic areas, at all grade levels, as well as behavioral screening measures.

Measures to Detect Early Behavior Problems

A recent national survey of school district administrators found that behavior screening was used in only about 12% of schools (Bruhn, Woods-Groves, & Huddle, 2014). As Kilgus and colleagues (2013) point out, there are several barriers to the implementation of behavior screening measures, including lack of school resources and a negative perception of the effectiveness of behavior screening. Indeed, some behavior screening measures can be costly in terms of dollars and time, and may require specialized training in scoring and interpretation of results. There are, however, a number of validated behavior screeners that can be used universally in much the same way as academic screeners.

The Behavioral and Emotional Screening System (BESS; Kamphaus & Reynolds, 2007) is one such universal screening instrument that was developed from the full-length BASC-2 (Reynolds & Kamphaus, 2004). The BESS has demonstrated adequate reliability and validity (see Kamphaus et al., 2007; King & Reschly, 2014; King, Reschly, & Appleton, 2012). Recently, the BESS was added as the behavioral screening measure used within the AIMSweb system (Howe & Shinn, 2002), and can be purchased as an add-on assessment for schools who currently use AIMSweb or as a stand-alone measure from the publisher (www.pearsonassessments.com).

Other screening measures, like the Behavior Screening Checklist (BSC; Muyskens, Marston, & Reschly, 2007) and the Student Risk Screening Scale (SRSS; Drummond, 1994), are universal behavior screeners that are free to use and efficient to administer. In addition, each has been found to have predictive validity of generalized outcomes (Ennis, Lane, & Oakes, 2012; King & Reschly, 2014). Overall, a number of well-validated behavior screeners are available; however, they vary in efficiency, application to RTI/MTSS, and cost-effectiveness. Furthermore, despite the availability of universal behavior screeners and development of multitiered student support structures, schools have been slow to implement behavior screening (Cook et al., 2010; Romer & McIntosh, 2005).

Benefits of Using a Combined Set of Measures for Early Detection

Although once daunting, the process of screening all students in a building or district is becoming streamlined and efficient with the help of manualized products specifically designed for screening and benchmarking, such as AIMSweb (Howe & Shinn, 2002). The AIMSweb system is the first widely implemented, manualized system available for assessing student risk in academics and behavior. This system allows schools to quickly assess entire student populations in the areas of reading and math, with brief CBM measures, and in behavior, with the addition of the BESS (Kamphaus & Reynolds, 2007). Adding the behavioral component to the academic screening system comes at an extra cost to schools; however, it also comes with added convenience, as it uses the same system of inputting scores and producing graphical results.

Given the increased use of screening in schools, particularly those implementing RTI/MTSS, investigating the utility of these screening measures to identify subgroups of students with co-occurring academic and behavior risk can be useful toward developing targeted interventions to prevent these students from school failure, including academic underachievement. Therefore, the purpose of this study was to determine whether subgroups of students with co-occurring academic and behavior problems can be identified among a third grade sample of students who were screened using AIMSweb in the fall of the school year. Third grade is often considered a critical academic year, as it is typically the first year that students participate in high-stakes state assessment, with results used to determine overall school quality, teacher quality, student performance, and even grade retention. In many school districts, students must earn passing scores on these high-stakes tests in 3rd grade to receive promotion to 4th grade. We anticipated that similar to Reinke and colleagues (2008), three to four subgroups of students will be identified, including one group with minimal academic and behavioral risk, one group with both academic and behavioral risk, and possibly a group with academic only risk and a group with behavior only risk. Further, to validate the identified student subgroups, we evaluated whether subgroups identified in the fall of third grade would be predictive of success on the spring administered 3rd grade state assessment subtests. We hypothesized that students identified with the greatest academic and behavioral risk in the fall would perform worse on the state assessment in comparison to other subgroups of students.

Method

Participants

Institutional Review Board (IRB) approval was obtained from the school district to access de-identified data from the district database. These data were part of a larger, multiyear, district-wide screening initiative among Kindergarten through 12th grade students in an urban school district in the midwest during the 2009 through 2012 school years. This school district was chosen for inclusion in the current study based on convenience sampling. All students participated in the screening as part of district expectations, so measures were collected as part of typical classroom instruction. District demographics provide a diverse sample, with an enrollment of 17,882 students, 72% of whom are African American, 57% who are eligible for free or reduced lunch, and 13% of whom receive special education services.

The current study includes 517 3rd grade students who were administered academic and behavior screening assessments during the 2011–2012 school year. Several schools did not conduct behavior screening and therefore these students were excluded from the study. The sample of students included in the current study was 55% boys with 65% receiving free or reduced lunch; 87% receiving general education services; and 4% receiving ESL services The ethnic diversity of the sample was as follows: 62% African American, 1% American Indian, 1% Asian, 3% Hispanic, and 33% Caucasian. See Table 1 for sample demographics.

Measures

Academic and behavioral screening indicators collected in fall of 3rd grade. Academic and behavior measures gathered in September of third grade were used as indicators of class membership. The following describes the AIMSweb reading, math, and behavioral screening measures that were used in this study. Each measure was made categorical to fit within a three-tiered model of determining risk. For instance, category 1 included all students who scored as falling into Tier 1, or meeting benchmark and minimal risk for each measure, category 2 included students who scored in the Tier 2 range, or slightly below benchmark and at-risk for each measure, and category 3 included students who scored in Tier 3, or not meeting benchmark and high risk for each

Table 1
Sample Demographics

Variable	*n*	% of sample
Gender		
Female	231	44.7
Male	286	55.3
Ethnicity		
African American	319	61.7
American Indian	3	0.6
Asian	7	1.4
Hispanic	17	3.3
Caucasian	171	33.1
Meal status		
Free	287	55.5
Reduced	44	8.5
Full priced	182	35.2
ESL		
Receives services	23	4.4
Does not receive	494	95.6
Special education		
Receives services	65	12.6
Does not receive	452	87.4
Disability category		
Intellectual disability	2	0.4
Emotional disturbance	7	1.4
Specific learning disability	9	1.7
Other health impairment	10	1.9
Autism spectrum	3	0.6
Language impairment	12	2.3
Speech impairment	22	4.3

measure. Tier decisions were made according to AIMSweb cut score criteria, which are presented in Table 2.

Reading skills. The *AIMSweb Reading Curriculum-Based Measurement (R-CBM)* is a 1-minute, individually administered, reading measure strongly related to grade-level standards. Students' scores are based on the number of words read aloud correctly during one minute. Oral reading measures, like those used in AIMSweb R-CBM, have been well-validated for use as indicators of overall reading competence (Good & Jefferson, 1998; Reschly, Busch, Betts, Deno, & Long, 2009). Reliability studies of R-CBM probes have found the alternate-probe reliability of Grade 3 level probes to be 0.85 for single probes and 0.94 for the projected mean of 3 probes (based on Spearman-Brown single-probe reliabilities; Howe & Shinn, 2002). Test–retest reliability coefficients, across 4 months, for Grade 3 R-CBM probes was found to be 0.94 from fall to winter and 0.95 from winter to spring (Christ & Silberglitt, 2007).

R-CBM data used in the current study were benchmark scores. Student scores are compared with national normative data to determine level of risk. For the 3rd graders in our sample, students correctly reading more than 76 words per minute were considered Tier 1, or meeting benchmark, students reading between 42 and 76 were considered Tier 2, or slightly below benchmark, and students reading less than 42 were considered Tier 3, well below benchmark. This tiered system of categorization aligns well with an RTI tiered system. For the purpose of these analyses, students meeting Tier 1 criteria were coded 1, Tier 2 were coded as 2, and Tier 3 were coded as 3.

In addition to the R-CBM, the *AIMSweb MAZE* was used as a measure of student reading abilities. The MAZE measure (www.aimsweb.com) is used as an indicator of overall reading proficiency through a silent reading, context-based task. In a MAZE passage every seventh word is replaced with a choice of three words (the correct word and two distractors). Students are directed to read the passage and select, by circling, the missing words from the choices. Students' scores on the MAZE are calculated by totaling the number of words circled correctly during the 3-min administration. MAZE instruments have been found to be valid and reliable measures of students' reading skills. MAZE has strong criterion-related validity with CBM-R, with coefficients ranging from .77 to .86 (Espin, Deno, Maruyama, & Cohen, 1989). The concurrent validity of MAZE has been established with other group-administered tests of reading achievement (Jenkins & Jewell, 1993). AIMSweb has established MAZE cut scores to determine level of risk using procedures identical to those described for R-CBM above. For our 3rd grade sample, students scoring greater than 10 were classified as Tier 1 or meeting benchmark, students scoring between 6 and 10 were considered Tier 2 or slightly below benchmark, and students scoring less than 6 on the MAZE were classified as Tier 3 or below benchmark. For the purpose of these analyses students meeting Tier 1 criteria were coded as 1, Tier 2 were coded as 2, and Tier 3 were coded as 3.

Math skills. Two *AIMSweb Mathematics* measures were examined: *Concepts and Applications* and *Computation*. *Concepts and Applications (M-CAP)* is a measure of general mathematics problem-solving skills. The 3rd Grade M-CAP is an 8-min, group-administered measure. The measure includes problem-solving and applied mathematics tasks that are aligned with standards for each grade level. It can be administered up to three times per year as a benchmarking tool to identify children at risk for mathematics difficulties. As such, student scores on this measure can assist in differentiating student level of risk. Third grade students scoring greater than 4 are considered Tier 1 or meeting benchmark, students scoring between 2 and 4 were classified as Tier 2 or slightly below benchmark, and students scoring less than 2 were considered Tier 3 or below benchmark. A normative study of 3rd Grade M-CAP probes, conducted by AIMSweb, found a split-half reliability of .83, and an alpha of .80 (AIMSweb, 2009). For the purpose of these analyses stu-

Table 2
Third Grade Cut-Scores for Determining Tier Level

AIMSweb	Tier 1	Tier 2	Tier 3
MAZE	>10	6–10	<6
RCBM	>76	42–76	<42
MCAP	>4	2–4	<2
MCOMP	>19	10–19	<10
BESS-T	<59	60–69	>69

dents meeting Tier 1 criteria were coded as 1, Tier 2 were coded as 2, and Tier 3 were coded as 3.

In addition to the M-CAP, the *AIMSweb Mathematics Computation (M-COMP*) was used in this study as a measure of math skills. The M-COMP is an 8-min, paper-based test that measures general math computation. It can be administered up to 3 times per year as a benchmarking tool of math computation. Problem types include mixed addition, subtraction, multiplication, and division problems that are differentiated and aligned by grade level standards. Student scores are compared to the national normative sample to determine level of risk. Third grade students scoring greater than 19 are considered Tier 1 or meeting benchmark, those scoring between 10 and 19 are classified as Tier 2 or slightly below benchmark, and those with scores of less than 10 are considered to be Tier 3 or below benchmark. The technical adequacy of M-COMP was evaluated with a total of 971 3rd grade students. Internal consistency, measured with split-half reliability, was found to be .90, and Cronbach's alpha was .89 (AIMSweb, 2010). For the purpose of these analyses students meeting Tier 1 criteria were coded as 1, Tier 2 were coded as 2, and Tier 3 were coded as 3.

Measure of behavior risk. *The Behavioral and Emotional Screening System* (BESS; Kamphaus & Reynolds, 2007) is a universal behavior screening measure that can either be purchased as a stand-alone measure, or administered as part of the AIMSweb system. The BESS was developed through the use of principal components analysis of the full-length Behavioral Assessment System for Children, Second Edition (BASC-2; Reynolds & Kamphaus, 2004). The result is a screening instrument designed to quickly and efficiently determine children's emotional and behavioral strengths and weaknesses. Like the BASC-2, the BESS has three report forms: Parent, Teacher, and Student. The data included in the current study are from the BESS Teacher Form. The BESS has recently been adopted by AIMSweb, as the behavioral component, to the well-established screening and progress monitoring system.

The BESS Teacher Form consists of 27 items, scored on a 4-point Likert-type scale, and reportedly takes 5 to 10 minutes to complete, per child. BESS forms result in an overall *t* score that uses combined or gender specific normative data to classify a child as having normal, elevated, or extremely elevated levels of risk. Students with *t* scores less than 60 are considered normal (Tier 1), those scoring between 60 and 69 are elevated (Tier 2), and those scoring greater than 69 are considered extremely elevated (Tier 3). Combined gender norms were used for analyses in the current study. Like the other AIMSweb tools, the BESS can be administered up to three times per year as a universal screening instrument. The BESS Teacher Form has been well validated by both the authors and independent researchers. The BESS manual reports internal-consistency by age group. The median (across age group) split-half reliability coefficient was .96 (Kamphaus & Reynolds, 2007). Test–retest reliability is reported to be .91, and interrater reliability between teachers was found to be .70. Data included in this study were collected during the September administration, and were completed online by general and special education teachers for each student in their classroom. For the purpose of these analyses students with a *t* score less than 60 were coded as 1, students with a *t* score between 60 and 69 were coded as 2, and students with a *t* score of 70 or higher were coded as 3.

Outcome measure. The statewide assessment measure administered in the spring of the school year was used as the outcome measure for this study.

Missouri Assessment Program (MAP). The MAP is a standardized, statewide assessment administered to students in Grade 3 through 8 in the spring of every school year. This criterion-referenced test was designed to measure student achievement toward state-level standards. Data included in the current study are from the 3rd Grade Mathematics and Communication Arts subtests of the MAP. Two types of scores are reported for the MAP tests in each area, a scale score and its associated level of achievement: Below Basic, Basic, Proficient, and Advanced. For the 2012 3rd Grade Communication Arts MAP test, a score of 591 or lower was considered Below Basic, 592 through 647 was Basic, 648 through 672 was Proficient, and scores at or above 673 were Advanced. On the Mathematics test, scores of 567 and lower were Below Basic, scores 568 through 627 were Basic, 628 through 666 were Proficient, and scores 667 and higher were Advanced. For this school district, a total of 43% of students across all participating grades scored as Proficient or Advanced in the area of Communication Arts, and 40% scored Proficient or Advanced in Mathematics on the 2012 tests. For comparison, the state levels of Proficient or Advanced on these tests were 56% in both areas. An independent contractor evaluated the technical adequacy of the MAP tests. Data published in the Technical Report of the 2012 version of the MAP test reported a Cronbach's alpha of .91for the 3rd Grade Communication Arts test, and .91 for the Mathematics test. Moreover, the 3rd Grade Communication Arts and Mathematics tests were correlated at .69 (Missouri Department of Elementary and Secondary Education, 2012).

Statistical Methods

Latent class analysis (LCA) was used to examine the underlying structure of the academic and behavior indicators in the fall of third grade. LCA is a person-centered approach that allows researchers to identify latent classes of children with similar patterns of behavior, and form groups based on these patterns (McCutcheon, 1987). The goal of LCA is to determine the smallest number of classes that describe the association between the selected indicators. Determining the appropriate number of classes is based on several factors, including model fit indices and theoretical relevance (Muthén & Muthén, 2010). Classes are formed based on similar patterns of behavior on the indicators. LCA uses categorical observed variables as indicator variables and is a person-centered approach that assumes an underlying latent variable that determines an individual's class membership (McCutcheon, 1987). Thus, the likelihood of a child being in each latent class is reported, and meaningful classes (based on probability) are formed.

After determining the appropriate number of classes, the classes were used to predict distal achievement outcomes. The Mplus Auxillary function (Muthén & Muthén, 2010) was used to evaluate differences on mean MAP subtest scores between classes. This evaluates the association between the classes identified with LCA and the mean score of students within each class on the high stakes state assessment.

Determining model fit. All analyses were conducted using MPlus 6.0 (Muthén & Muthén, 2010). There are several statistical indicators of model fit used in LCA, and the best model is chosen based on these indicators, as well and substantive theory. To determine the relative fit of the models, models with differing numbers of classes were compared using the Akaike information criterion (AIC; Akaike, 1987), the Bayesian information criterion (BIC; Schwarz, 1978), and the sample-size adjusted Bayesian information criterion (aBIC; Sclove, 1987). Typically, the smaller the information criteria, the better the model fit to the data. In addition, we evaluated the classification precision as indicated by estimated posterior class probabilities, summarized by the entropy measure (Ramaswamy, DeSarbo, Reibstein, & Robinson, 1993). Entropy values close to 1.0 indicate higher classification precision (Muthén, 2004). The Bayesian Information Criterion is given more weight when determining model fit as it has been shown to be the most reliable indicator of true model fit (Nylund, Asparouhov, & Muthén, 2007). Further, the Vuong-Lo-Mendell-Rubin Likelihood Ratio Test (VLMR LRT) is another indicator of model fit that provides a *p* value of the model strength over a model with one fewer class. Significant *p* values on the VLMR LRT indicate that the current model is a significantly better fit for the data than a model with one fewer class (Nylund et al., 2007).

Treatment of missing data. The Mplus program assumes that data are missing at random, and uses full information maximum likelihood estimation (MAR; Arbuckle, Marcoulides, & Schumacker, 1996; Little, 1995), which is a widely accepted way of handling missing data (Muthén & Shedden, 1999; Schafer & Graham, 2002). The covariance coverage for all variables ranged from 0.98 to 1.0, well above minimum thresholds for establishing adequate coverage (e.g., .10; Muthén & Muthén, 2010).

Results

Latent Class Analysis of Academic and Behavior Problems

LCA were conducted to determine the optimal number of classes and the academic and behavioral characteristics of third grade students associated with each class. We considered the AIC, BIC, and aBIC indices, with the smaller value indicating a better-fit model. In addition, entropy was considered in the determination (see Table 3). LCA model fit indices

Table 3
Model Fit Indices for 1–4 Class Solutions of Academic and Behavior Problems

Model	AIC	BIC	Adj BIC	VLMR LRT	Entropy
1-class solution	4953.180	4995.661	4963.919	—	—
2-class solution	4485.624	4574.832	4508.174	0.00	0.80
3-class solution	**4407.276**	**4543.213**	**4441.639**	**0.002**	**0.73**
4-class solution	4379.382	4562.048	4425.558	0.02	0.75

Note. LC = Latent class; AIC = Akaike information criterion; BIC = Baysian information criterion; aBIC = adjusted Baysian information criterion. Bold indicates best fit: The three-class solution had the lowest BIC and the VLMR LRT and the Bootstrap LRT indicated the 3-class solution provided a better fit than the 4-class solution. All entropy ratings indicate acceptable fit. Entropy values close to 1.0 indicate higher classification precision.

for class solutions are summarized in Table 3. The three-class solution emerged as the optimal fit for the data as evidenced by lowest BIC value for this solution. In addition, the VLMR indicated that the three-class solution was a better fitting model than the two-class solution and the entropy was adequate. We then included class prevalence and interpretability (the extent to which an additional class provided unique information) as additional criteria while selecting the best-fitting models.

Figure 1 summarizes the percentage of students within each class that met Tier 1 (minimal risk), Tier 2 (some risk), and Tier 3 (high risk) across indicators. Class labels were assigned based on the overall pattern and presentation of academic and behavior problems for each class. Class 1 was characterized as a subclass of students falling into Tier 1, indicating minimal risk for academic or behavior problems ($n = 167$; 32%). Class 2 was characterized by a subclass of students who had a higher probability of having academic and behavior problems falling into Tier 2, indicating some risk for academic and behavior problems ($n = 189$; 37%). Finally, Class 3 was characterized by a subclass of students with a higher probability of demonstrating significant academic and behavior concerns at the Tier 3 level ($n = 161$; 31%).

Outcomes on State Academic Assessment With Each Class

To validate the subclasses and determine whether students in the subclass with academic and behavior indicators of risk were more likely to perform poorly than other classes, the mean score of students within each class on the high stakes state assessment were evaluated (see Table 4). Statistically significant differences were observed among the classes in terms of the Communications Arts and Mathematic subtests of the spring administered state academic assessment. The Tier 1 class had the highest mean score for Communication Arts ($M = 658.05$; $SD = 2.29$) scoring significantly higher than Tier 2 class ($\chi^2 = 61.33$, $p < .001$) and Tier 3 class ($\chi^2 = 160.58$, $p < .001$). The Tier 2 class exhibited the second highest mean score for Communication Arts on the state assessment ($M = 640.81$; $SD = 2.19$), scoring significantly higher than the Tier 3 class ($\chi^2 = 27.83$, $p < .001$). Students in the Tier 3 scored the lowest on the Communication Arts subtest ($M = 611.91$; $SD = 2.85$), with the mean level score falling into the Basic level, which is not considered a passing score on the assessment.

With regard to the Mathematics subtest on the state academic assessment, a similar pattern emerged. The Tier 1 class had the highest mean score for Mathematics ($M = 644.86$; $SD = 2.53$), scoring significantly higher than Tier 2 class ($\chi^2 = 16.79$, $p < .001$) and Tier 3 class ($\chi^2 = 72.39$, $p < .001$). The Tier 2 class exhibited the second highest mean score for Mathematics on the state assessment ($M = 620.78$; $SD = 2.17$), scoring significantly higher than the Tier 3 class ($\chi^2 = 47.78$, $p < .001$). Students in the Tier 3 class scored the lowest on the Mathematics subtest ($M = 598.85$; $SD = 4.79$), with the mean level score falling into the Basic level, which is considered failing by state standards. Results are summarized in Table 4.

Discussion

The purpose of this study was to evaluate whether students with co-occurring academic

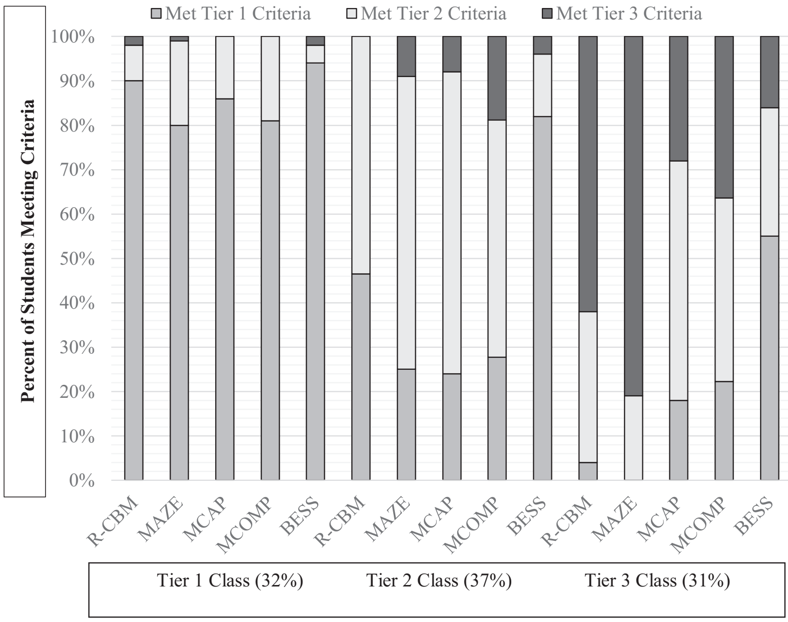

Figure 1. Percentage of students within each class meeting Tier 1, Tier 2, and Tier 3 criteria across indicators for each class of students in fall of third grade.

and behavior problems can be readily identified at the start of the school year using screening measures common in school practices, particularly among schools implementing tiered models of support such as RTI and MTSS. Further, we examined the link between early indicators of behavior and academics and later risk status in academics, as measured by a state high-stakes assessment. Although many studies have examined behavioral risk and academic risk of elementary age students, few have examined both concurrently. In addition, this study extended the literature on the use of short-duration, efficient-to-administer screening instruments by combining information from the two sources of data to create integrated classes of student risk. Identifying potential classes of student academic and behavioral risk is important to schools and teachers as they work to better define and integrate the treatment that might result in the best outcomes for students at risk. In addition, for schools that are using RTI/MTSS models, identification of classes of students with both behavior and academic risk helps to better align resources, targeting tiered levels of intervention based on the child's complete profile.

Results of this study indicated three, fairly similar sized classes of students: a class with minimal academic and behavioral risk (Tier 1),

Table 4
Class Counts, Proportions, and Equality Tests Across Classes on Screening Measures

MAP subtest	Class 1 $n = 167$, 32%	Class 2 $n = 189$, 37%	Class 3 $n = 161$, 31%	Overall test of significance	Significant class comparisons
Comm. Arts	658.01 Proficient	640.74 Basic	611.52 Basic	166.94	Class 1 vs. 3; 1 vs. 2; & 2 vs. 3***
Mathematics	644.91 Proficient	620.73 Basic	602.55 Basic	117.49	Class 1 vs. 3; 1 vs. 2; & 2 vs. 3***

Note. Chi-square *p* values.
*** $p < .001$.

a class with some academic, yet little behavioral risk (Tier 2), and a class with a high degree of academic and some behavioral risk (Tier 3). These findings were similar to prior studies (Reinke et al., 2008). Interestingly, there was not a class of students who was clearly at-risk in only behavior, but not at-risk in academics. In fact, although most of the children with behavior risk (Tier 3 on BESS) were grouped into one class (the Tier 3 class), a large portion of students in this class did not exhibit behavioral risk. In addition, there were no clear delineations between classes for reading and mathematics performance on the academic screeners, as can be seen in Figure 1. The current study expanded on the results of prior studies (e.g., Reinke et al., 2008) by demonstrating that similar classes are found when using CBM measures as indicators of class membership. This is an encouraging finding given that CBM data are becoming more readily available in schools, more pertinent to decision making within MTSS models, and is central in the development of cohesive and integrated universal screening models (Cook et al., 2010).

Past research has demonstrated a positive association between CBM measures and state test outcomes (Good, Simmons, & Kame'enui, 2001; McGlinchey & Hixson, 2004). However, little if any work has been completed to examine the use of a combination of short duration academic and behavior measures to predict state outcomes. We evaluated whether subgroups of students with co-occurring academic and behavior problems could be identified in the fall of 3rd grade and determine their performance on spring administered 3rd grade state assessment tests. We hypothesized that students identified with the co-occurring academic and behavioral risk in the fall would perform poorly on the state assessment in comparison to other subgroups of students. This hypothesis was confirmed. Class means on both subtests of the statewide achievement test were significantly different between each class in the expected directions. However, results indicated that both the Tier 2 and Tier 3 class performed in the Basic level of proficiency on both subject tests, which is not considered passing by state standards. The Tier 1 class, on average, was the only class to pass the statewide test in either subject area.

Limitations

Using brief academic and behavior screening instruments has been found to provide predictive and concurrent validity; however, these scores are only one estimate of academic and behavioral functioning gathered at one time point. This study, like many others, has not yet answered the question of onset of academic and behavior problems. Thus, the course of academic problems influencing behaviors, and vice versa, is yet to be determined.

This study was limited to one school district that, although diverse, cannot generalize to other populations or geographic locations. Replication of this study is necessary. Limitations of using an existing de-identified dataset were also present in this study, including lack of school-level and teacher demographic data. Additionally, the outcome measure of the current study was a statewide standardized achievement test. This testing has powerful implications for students and school districts; however, this type of test is not the only measure of success for students. Other outcome variables of interest may include referral to special education, graduation, grade level retention, social skill and adjustment, and other measures of academic achievement.

Future Studies

An interesting finding for the current study is that there was not a single class where a majority of the children experienced behavioral difficulties. The Tier 3 class contained the majority of the children with significant scores on the BESS; however, this class was still primarily composed of children without behavioral risk. Thus, our ability to detect differences in academic risk among children with behavioral risk is limited. Future studies, using latent class procedures with samples of children identified as at-risk behaviorally (i.e., clinical populations), might reveal differing levels of academic risk even among children already at risk behaviorally. In other words, it would be of practical importance to identify classes of children, at known behavioral risk, with varying levels or combination of academic risk, if those classes are predictive of achievement outcomes, as these classes might inform intervention services.

The current study is useful and significant in that predictive classes of children were identified. However, these classes were formed with data drawn from one time point. It is impossible within this research design to determine which risk factors, behavior or academic, developed first. Using a latent transition approach, researchers would be able to study the movement of children between groups across several time points. This approach could indicate subclasses of children who first developed behavioral problems, then academic, and vice versa.

Implications for Practice

Students are not their reading scores or behaviors scores; rather, they are individuals with varying levels of skills across areas, each influencing the other. Latent class analysis identifies groups of children with similar patterns of skills, and examines the predictive ability of those groups. By assessing risk in both academics and behavior using brief screening instruments, school personnel can choose interventions targeted to both academics and behavior, as determined by student need. The process of identifying children with concurrent behavioral and academic risk, and developing applicable interventions based on risk status, is impeded when separate "behavior" and "academic" data teams are established. Integrated data teams, like those proposed by Lane and colleagues (2014), allow for examination of student risk in all areas of school functioning, including academic, behavior, and social. These teams, comprising school personnel with comprehensive areas of expertise, work to efficiently assess and adequately intervene at the "whole child" level (Lane et al., 2014).

This study adds to the literature on the utility of CBM and behavior screening measures. Screening scores from the beginning of the school year were used to predict, based on class membership, performance on the end of year state test. Given the quick and easy nature of these screening assessments, it is encouraging that the scores can inform intervention to such a degree. Determining risk status using cut-scores of academic and behavior screening measures administered in the fall, schools can identify subgroups of children at most risk of performing poorly on the state assessment, and implement target interventions early in the school year. The results of this study indicated that schools should focus intervention efforts on the subgroup of children with concurrent behavioral and academic difficulty, as this group of children is at greatest risk of poor academic outcomes on the spring statewide achievement test. Data teams are encouraged to carefully evaluate fall benchmarking data to determine overlap of academic and behavioral deficits, and implement intervention early in the school year for these students.

One would expect that students who display both academic and behavioral deficits need targeted, dual interventions. Otherwise, simply focusing on academics may not optimize improved outcomes for these students. Knowing more about combined classes of behavior and academics can help data teams provide more specifically designed academic and social instruction. One example would be teaching students self-regulation strategies to implement as they are working on academic tasks. An example of this is the use of Self-Regulated Strategy Development (SRSD; Graham et al., 2005) in writing with students with behavioral difficulties. Writing can be frustrating for students and writing tasks can sometimes serve as a trigger for behavioral events. SRSD incorporates writing strategies with self-regulatory techniques such as goal setting and self-monitoring, and has demonstrated effectiveness with students with behavioral difficulties (Lane et al., 2010).

The design and results of this study align with a comprehensive RTI/MTSS model, incorporating both behavior and academics. Within RTI/MTSS models, resources are allocated to match the level of student need. Thus, students at most severe risk receive the most intensive interventions. This study provides useful information regarding the identification of students at risk, and indicates that screeners can be used within a RTI/MTSS model to allocate intervention resources. Identifying classes predictive of academic failure could inform an allocation of resources to those students with the greatest need or likelihood of improvement.

References

AIMSweb. (2009). *Math concepts and applications: Administration and technical manual*. San Antonio, TX: Pearson.

AIMSweb. (2010). *Mathematics computation: Administration and technical manual*. Bloomington, MN: Pearson.

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

Anderson, J. A., Kutash, K., & Duchnowski, A. J. (2001). A comparison of the academic progress of students with EBD and students with LD. *Journal of Emotional and Behavioral Disorders, 9,* 106–115. http://dx.doi.org/10.1177/106342660100900205

Arbuckle, J. L., Marcoulides, G. A., & Schumacker, R. E. (1996). Full information estimation in the presence of incomplete data. *Advanced structural equation modeling: Issues and techniques, 24*3–277.

Benner, G. J., Kutash, K., Nelson, J., & Fisher, M. B. (2013). Closing the achievement gap of youth with emotional and behavioral disorders through multi-tiered systems of support. *Education & Treatment of Children, 36,* 15–29. http://dx.doi.org/10.1353/etc.2013.0018

Berger, M., Yule, W., & Rutter, M. (1975). Attainment and adjustment in two geographical areas. II—The prevalence of specific reading retardation. *The British Journal of Psychiatry, 126,* 510–519. http://dx.doi.org/10.1192/bjp.126.6.510

Bruhn, A. L., Woods-Groves, S., & Huddle, S. (2014). A preliminary investigation of emotional and behavioral screening practices in k-12 schools. *Education & Treatment of Children, 37,* 611–634. http://dx.doi.org/10.1353/etc.2014.0039

Bub, K. L., McCartney, K., & Willett, J. B. (2007). Behavior problem trajectories and first-grade cognitive ability and achievement skills: A latent growth curve analysis. *Journal of Educational Psychology, 99,* 653–670. http://dx.doi.org/10.1037/0022-0663.99.3.653

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

Collier-Meek, M. A., Fallon, L. M., Sanetti, L. M., & Maggin, D. M. (2013). Focus on implementation: Assessing and promoting treatment fidelity. *Teaching Exceptional Children, 45,* 52–59.

Cook, C. R., Volpe, R. J., & Livanis, A. (2010). Constructing a roadmap for future universal screening research beyond academics. *Assessment for Effective Intervention, 35,* 197–205. http://dx.doi.org/10.1177/1534508410379842

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

Dishion, T. J., French, D. C., & Patterson, G. R. (1995). *The development and ecology of antisocial behavior*. Hoboken, NJ: Wiley.

Drummond, T. (1994). *The student risk screening scale (SRSS*). Grants Pass, OR: Josephine County Mental Health Program.

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

Ennis, R. P., Lane, K. L., & Oakes, W. P. (2012). Score reliability and validity of the Student Risk Screening Scale: A psychometrically sound, feasible tool for use in urban elementary schools. *Journal of Emotional and Behavioral Disorders, 20,* 241–259. http://dx.doi.org/10.1177/1063426611400082

Espin, C. A., Deno, S. L., Maruyama, G., & Cohen, C. (1989). *The Basic Academic Skills Samples (BASS): An instrument for screening and identifying children at risk for failure in the regular education classroom*. In American Educational Research Association Meeting, San Francisco, CA.

Fleming, C. B., Harachi, T. W., Cortes, R. C., Abbott, R. D., & Catalano, R. F. (2004). Level and change in reading scores and attention problems during elementary school as predictors of problem behavior in middle school. *Journal of Emotional and Behavioral Disorders, 12,* 130–144. http://dx.doi.org/10.1177/10634266040120030101

Fuchs, L. S. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review, 33,* 188–192.

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

Good, R. H., III, Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading, 5,* 257–288. http://dx.doi.org/10.1207/S1532799XSSR0503_4

Good, R., & Jefferson, G. (1998). Contemporary perspectives on curriculum-based measurement validity. *Advanced applications of curriculum-based measurement, 6*1–88.

Graham, S., Harris, K. R., & Mason, L. (2005). Improving the writing performance, knowledge, and self-efficacy of struggling young writers: The effects of self-regulated strategy development. *Contemporary Educational Psychology, 30,* 207–241. http://dx.doi.org/10.1016/j.cedpsych.2004.08.001

Gresham, F. M. (2007). Response to intervention and emotional and behavioral disorders: Best practices in assessment for intervention. *Assessment for Ef-*

fective Intervention, 32(4), 214–222. http://dx.doi.org/10.1177/15345084070320040301

Heiervang, E., Stevenson, J., Lund, A., & Hugdahl, K. (2001). Behaviour problems in children with dyslexia. *Nordic Journal of Psychiatry, 55,* 251–256. http://dx.doi.org/10.1080/080394801681019101

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.

Jenkins, J. R., & Jewell, M. (1993). Examining the validity of two measures for formative teaching: Reading aloud and maze. *Exceptional Children, 59,* 1367–1925.

Kalberg, J. R., Lane, K. L., & Menzies, H. M. (2010). Using systematic screening procedures to identify students who are nonresponsive to primary prevention efforts: Integrating academic and behavioral measures. *Education & Treatment of Children, 33,* 561–584. http://dx.doi.org/10.1353/etc.2010.0007

Kamphaus, R. W., & Reynolds, C. R. (2007). *Behavioral & emotional screening system*. San Antonio, TX: NCS Pearson.

Kamphaus, R. W., Thorpe, J. S., Winsor, A. P., Kroncke, A. P., Dowdy, E. T., & VanDeventer, M. C. (2007). Development and predictive validity of a teacher screener for child behavioral and emotional problems at school. *Educational and Psychological Measurement, 67,* 342–356. http://dx.doi.org/10.1177/00131644070670021001

Kilgus, S. P., Chafouleas, S. M., & Riley-Tillman, T. C. (2013). Development and initial validation of the Social and Academic Behavior Risk Screener for elementary grades. *School Psychology Quarterly, 28,* 210–226. http://dx.doi.org/10.1037/spq0000024

King, K. R., & Reschly, A. L. (2014). A comparison of screening instruments: Predictive validity of the BESS and BSC. *Journal of Psychoeducational Assessment, 32,* 687–698. http://dx.doi.org/10.1177/0734282914531714

King, K. R., Reschly, A. L., & Appleton, J. J. (2012). An examination of the validity of the Behavioral and Emotional Screening System in a rural elementary school: Validity of the BESS. *Journal of Psychoeducational Assessment, 30,* 527–538. http://dx.doi.org/10.1177/0734282912440673

Lane, K. L., Barton-Arwood, S. M., Nelson, J. R., & Wehby, J. (2008). Academic performance of students with emotional and behavioral disorders served in a self-contained setting. *Journal of Behavioral Education, 17,* 43–62. http://dx.doi.org/10.1007/s10864-007-9050-1

Lane, K. L., Graham, S., Harris, K. R., Little, M. A., Sandmel, K., & Brindle, M. (2010). Story writing: The effects of self-regulated strategy development for second-grade students with writing and behavioral difficulties. *The Journal of Special Education, 44,* 107–128. http://dx.doi.org/10.1177/0022466908331044

Lane, K. L., Gresham, F. M., & O'Shaughnessy, T. E. (2003). Review of interventions for children with or at risk for emotional and behavioral disorders. *Behavioral Disorders, 28,* 190–191.

Lane, K. L., Oakes, W. P., & Menzies, H. M. (2014). Comprehensive, Integrated, Three-Tiered (CI3T) Models of Prevention: Why does my school-and district-need an integrated approach to meet students' academic, behavioral, and social needs? *Preventing School Failure: Alternative Education for Children and Youth, 58,* 121–128. http://dx.doi.org/10.1080/1045988X.2014.893977

Lee, Y.-Y., Sugai, G., & Horner, R. H. (1999). Using an instructional intervention to reduce problem and off-task behaviors. *Journal of Positive Behavior Interventions, 1,* 195–204. http://dx.doi.org/10.1177/109830079900100402

Little, R. J. (1995). Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association, 90,* 1112–1121. http://dx.doi.org/10.1080/01621459.1995.10476615

McCutcheon, A. L. (Ed.). (1987). *Latent class analysis* (No. 64). Thousand Oaks, CA: Sage.

McGlinchey, M. T., & Hixson, M. D. (2004). Using curriculum-based measurement to predict performance on state assessments in reading. *School Psychology Review, 33,* 193–203.

McIntosh, K., Horner, R. H., Chard, D. J., Dickey, C. R., & Braun, D. H. (2008). Reading skills and function of problem behavior in typical school settings. *The Journal of Special Education, 42,* 131–147. http://dx.doi.org/10.1177/0022466907313253

Missouri Department of Elementary and Secondary Education. (2012). *Missouri assessment program grade-level assessment* (Tech. Rep. No. 2012). Monterey, CA: Author.

Muthén, B. (2004). Latent variable analysis. *In The Sage handbook of quantitative methodology for the social sciences* (pp. 345–368). Thousand Oaks, CA: Sage Publications.

Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics, 55,* 463–469. http://dx.doi.org/10.1111/j.0006-341X.1999.00463.x

Muthén, L., & Muthén, B. (2010). *Mplus 6.0*. Los Angeles, CA: Author.

Muyskens, P., Marston, D., & Reschly, A. L. (2007). The use of response to intervention practices for behavior: An examination of the validity of a screening instrument. *California School Psychologist, 12,* 31–45. http://dx.doi.org/10.1007/BF03340930

Nelson, J. R., Benner, G. J., Lane, K., & Smith, B. W. (2004). Academic achievement of K-12 students with emotional and behavioral disorders. *Exceptional Children, 71,* 59–73. http://dx.doi.org/10.1177/001440290407100104

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14,* 535–569. http://dx.doi.org/10.1080/10705510701575396

Oakes, W. P., Lane, K. L., Cox, M. L., & Messenger, M. (2014). Logistics of behavior screenings: How and why do we conduct behavior screenings at our school? *Preventing School Failure, 58,* 159–170. http://dx.doi.org/10.1080/1045988X.2014.895572

O'Shaughnessy, T. E., Lane, K. L., Gresham, F. M., & Beebe-Frankenberger, M. E. (2003). Children placed at risk for learning and behavioral difficulties: Implementing a school-wide system of early identification and intervention. *Remedial and Special Education, 24,* 27–35. http://dx.doi.org/10.1177/074193250302400103

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

Reid, R., Gonzalez, J. E., Nordness, P. D., Trout, A., & Epstein, M. H. (2004). A meta-analysis of the academic status of students with emotional/behavioral disturbance. *The Journal of Special Education, 38,* 130–143. http://dx.doi.org/10.1177/00224669040380030101

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

Reschly, A. L., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology, 47,* 427–469. http://dx.doi.org/10.1016/j.jsp.2009.07.001

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior Assessment System for Children* (2nd ed.). San Antonio, TX: Pearson.

Romer, D., & McIntosh, M. (2005). The roles and perspectives of school mental health professionals in promoting adolescent mental health. In D. L. Evans, E. B. Foa, R. E. Gur, H. Hendin, C. P. O'Brien, M. P. Seligman, & T. Walsh (Eds.), *Treating and preventing adolescent mental health disorders: What we know and what we don't know: A research agenda for improving the mental health of our youth* (pp. 597–615). New York, NY: Oxford University Press.

Rutter, M., & Yule, W. (1970). Reading retardation and antisocial behaviour: The nature of the association. In M. Rutter, J. Tizard, & K. Whitmore (Eds.), *Education, health, and behaviour* (pp. 240–255). London, UK: Longmans.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7,* 147–177. http://dx.doi.org/10.1037/1082-989X.7.2.147

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6,* 461–464. http://dx.doi.org/10.1214/aos/1176344136

Sclove, S. L. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika, 52,* 333–343. http://dx.doi.org/10.1007/BF02294360

Willcutt, E. G., & Pennington, B. F. (2000). Psychiatric comorbidity in children and adolescents with reading disability. *Journal of Child Psychology and Psychiatry, 41,* 1039–1048. http://dx.doi.org/10.1111/1469-7610.00691

Received September 15, 2014
Revision received December 9, 2014
Accepted December 17, 2014 ■