CASP
CALIFORNIA
ASSOCIATION
OF SCHOOL
PSYCHOLOGISTS

The California School Psychologist, Vol. 12, pp. 31 –45, 2007

The Use of Response to Intervention Practices for Behavior: An Examination of the Validity of a Screening Instrument

Paul Muyskens &
Doug Marston
Minneapolis Public Schools

Amy L. Reschly
University of Georgia

Behavioral difficulties of school-aged students are typically dealt with in a reactive, rather than preventative manner. This article examines a proactive approach, consistent with the Response-to-Intervention model, using a screening measure designed to identify students at risk for behavior difficulties and targeting these students for early intervention. Participants in this study were 22,056 kindergarten through 8th grade students in the Minneapolis Public Schools. Teachers completed a 12-item behavior screener for each student in the fall. The scores on this measure were significantly correlated with suspensions, achievement scores and attendance data over the course of the school year. Potential implications for behavioral planning and interventions are discussed.

KEYWORDS: Behavior, Behavior Screening, Response to Intervention

In a recent report, the National Academy of Sciences Committee on Minority Representation in Special Education noted, "There is substantial evidence with regard to both behavior and achievement that early identification and intervention is more effective than later identification and intervention." (p. 6, Donovan & Cross, 2002). While the logic of preventing or addressing problems before they become severe is both obvious and compelling, historically, schools have been reactive, rather than proactive institutions (Walker et al., 1996). In terms of addressing academic or behavioral concerns, this reactivity is often referred to as a "wait-to-fail" model. Criticism of this approach is well-documented and there are increasing calls in the literature for screening and/or early intervention for academic and behavior problems (e.g., Donovan & Cross, 2002; O'Shaughnessy, Lane, Gresham, & Beebe-Frankenberger, 2003; Walker et al., 1996; Reschly & Ysseldyke, 2002).

Despite the link between student behavior and academic performance (Alexander, Entwisle, & Dauber, 1993) and the association between behavior and long-term social outcomes for students (Kaplan, Peck, & Kaplan, 1997), the knowledge base for systematic screening and tracking of emotional and behavior problems lags behind that of academics, particularly reading (Donovan & Cross, 2002). In addition to interfering with other students' learning, problem behaviors can be violent and/or destructive. For example, in 1999-2000, 29% of principals reported that student bullying occurred daily or weekly, while 19% of principals reported acts of disrespect to teachers with the same frequency (DeVoe et al., 2004). Many educators report difficulty managing student behavior on a day-to-day basis and both parents and the general public are concerned about discipline and safety in our schools. For example, when Americans were asked about the biggest problems facing education, concerns over school discipline/control were tied for 2nd as the most frequent response (behind school funding and tied with overcrowding; Gallup, 2004); and 1/3 of parents worry about their children's physical safety at school, a concern that peaks with the parents of middle school-age children (Gallup, 2002).

Please send correspondence to Paul Muyskens, at Minneapolis Public Schools; 807 NE Broadway; Minneapolis, MN, USA; 55413-2398 or e-mail Paul.Muyskens@mpls.k12.mn.us

Although there is evidence for the efficacy of proactive, school-wide behavior support models (e.g., Sugai et al., 2000), schools typically rely on disciplinary action as the primary means of managing student misbehavior. Zero tolerance policies, the elimination of social promotion, and "get tough" rules are the current trends. Unfortunately, there is little evidence for the effectiveness of these approaches for improving school safety or long-term student outcomes (Skiba & Peterson, 1999). In addition, the result of disciplinary policies based upon clear standards or zero tolerance is typically not instructional, but rather exclusionary. Students and schools can ill afford exclusionary policies that result in lost opportunities for instruction, an outcome that is associated with lower student and school-level achievement (Skiba & Rausch, 2004) and may be related to increased dropout rates, or students dropping out at a younger age (Bock, Tapscott, & Savner, 1998; DeRidder, 1990).

Another issue in managing student behavior primarily through disciplinary approaches is that the application of these practices varies by school characteristics, principal attitudes, and even by teachers within the same school (Rausch & Skiba, 2004; Skiba & Rausch, 2006). Whenever there is uneven or inconsistent application of an educational practice or policy, harmful disparities may result. In the case of discipline, there is evidence that the use of harsh disciplinary practices varies by socioeconomic status and race, negatively affecting students of lower socioeconomic backgrounds and African-American students. As noted by Skiba and Rausch (2006), "Over thirty years of consistent data concerning African American over-representation in suspension and expulsion indicates that disciplinary school exclusion may carry inherent risks for creating or exacerbating racial and socioeconomic disadvantage." (p. 93).

Concerns over fairness and bias in managing and evaluating student behavior spill over into the special education arena. On a national level, African-American students are disproportionately represented in special education programs for Emotional Disturbance and Mental Retardation, and American-Indian students are disproportionately represented in programs for Learning Disabilities, while Asian Americans, Hispanics, and Whites are underrepresented (Donovan & Cross, 2002; MacMillan & Reschly, 1998). While the causes of over- and under- representation are complex and the debate controversial, there is concern about bias in interpretation of students' behavior and the seemingly haphazard nature of the referral procedure. The identification and referral process for emotional and behavioral concerns has been described as, "unsystematic, idiosyncratic, and late in the development of a behavioral problem" (Donovan & Cross, 2002, p.296), while the President's Commission on Excellence in Special Education (U.S. Department of Education, Office of Special Education and Rehabilitative Services, 2002, p. 26) noted that, "Minority children are much more likely to be placed in the emotional disturbance category because of behavioral characteristics associated with the cultural context in which a child is raised. A major factor is the role of teacher referral."

MINNEAPOLIS PUBLIC SCHOOLS PROBLEM-SOLVING/RESPONSE TO INTERVENTION MODEL

A Response-to-Intervention (RtI), or problem-solving, approach has been put forth as a potential alternative to the traditional refer-test-place process for special education (e.g., Gresham, 2002; Grimes, 2002; Marston, 2002; Reschly & Tilly, 1999; Shinn, Good, & Parker, 1999). Problem-solving has been described as, "…a self-correcting methodology to identify, analyze, and intervene with difficulties at the individual, group, and system levels" (Reschly, 2005). As described by Fuchs and Fuchs (2006), the first step in an RtI model is to identify students at-risk for not responding to interventions. This activity should take place early in the school year, and involves screening all students on a tool that is tied to

subsequent performance. An RtI or problem-solving model is a means of counteracting the *wait-to-fail* approach because once students are identified as at-risk, their progress is monitored and if progress is not made interventions are provided. Students' responses, or outcomes to interventions, are then used to determine the need for more intensive interventions and services. When personnel and services are organized around intervention, rather than *assessment for placement*, the time and resources for screening and early intervention activities become available.

While most recent attention and controversy about RtI concerns the use of this model in identifying students with learning disabilities (e.g., Bradley, Danielson, & Hallahan, 2002; Naglieri, 2005; Reschly, 2005), large-scale applications (Iowa, Minneapolis Public Schools) use this methodology to address both academic *and* behavioral concerns. In addition, there is some evidence that the utilization of a model like RtI reduces minority overrepresentation in special education (Barbour, 2006), possibly due to the use of direct (e.g., Curriculum-Based Measurement; structured behavioral observation) rather than indirect measures (e.g., IQ) of student performance.

Minneapolis Public Schools has used a Problem-Solving Model (PSM) as an alternative to traditional special education eligibility criteria since 1993. The steps in the PSM include defining the problem and establishing the current level of performance, generating possible solutions to the problem, implementing the best solution, and reviewing the results of the intervention (see Deno, 2002). As implemented in Minneapolis, students in need of intervention cycle through this process first within the general education classroom, then with increasing intensity at the building assistance team level and, if necessary, in the special education evaluation phase. Further information regarding the rationale of the PSM or its implementation in Minneapolis Public Schools may be found in Marston, Muyskens, Lau, and Canter (2003) and Lau, Sieler, Muyskens, Canter, VanKueren and Marston (2006). The expansion of the role of the PSM in the Minneapolis Public Schools was facilitated by a voluntary compliance agreement with the Office of Civil Rights. The purpose of this agreement was to address the disproportionate representation of some minority groups in special education programs by implementing screening and early interventions for students having difficulty with academics and/or school behavior. Implementation of behavioral screening for kindergarten through 8th grades was a part of this agreement.

Although sites across the country vary in their implementation of the PSM, a hallmark of the Minneapolis model is the use of both academic and behavioral screening measures. In this model, students go from screening directly into a series of activities or queries related to the design of an intervention. More specifically, high scores on the screening measures indicate a need for further investigation, and school staff is subsequently directed to complete the *Classroom Intervention Worksheet* for these students. Subsequent completion of the *Classroom Intervention Worksheet* prompts teachers to initiate communications with parents, students, other staff members and a review of records. Further, teachers are asked to describe current levels of performance, identify student strengths, review health information, and specify potential intervention strategies. This information is used to design a classroom-based intervention that is implemented by the teacher. Students who do not demonstrate improved behaviors in response to this intervention may be referred to the problem-solving team. This team implements a problem-solving process which is used to design a more intensive intervention, along with the collection of more frequent progress-monitoring data. Students who do not respond to these interventions may then be referred for special education evaluation (see Marston et al., 2003).

There is a knowledge base for screening students in core academic areas (Donovan & Cross, 2002). Minneapolis Public Schools utilizes students' performance on the District-wide achievement test (the Northwest Achievement Levels Test, or NALT), Curriculum-Based Measurement probes, and informal

reading inventories. In many cases, scores on these measures are tied, or benchmarked, to passing scores on the state of Minnesota high-stakes assessment, the Minnesota Comprehensive Assessment – Series II (MCA-IIs). As mentioned previously, however, the methodology to systematically screen and track students with behavioral concerns lags behind that which is available in academic areas. To address the need for behavioral screening in the PSM, staff in the Minneapolis Public Schools developed a new screening measure, the Behavior Screening Checklist (BSC).

The determination that a new behavioral screening measure was needed was largely done through a process of elimination. There were several school-related variables that were plausible indicators of behavioral problems, such as office referral and suspension data, teacher report card ratings, and observations. However, there were problems with using these variables for screening purposes. For example, office referral and suspension data were typically too reactionary for a screening measure (i.e., data are not collected until after a problem is severe), and the data are often not compiled systematically. Report card ratings were quite global in nature (e.g., satisfactory, satisfactory plus) and the use of observations for screening purposes would have been very time-intensive and expensive. In selecting a measure for screening, it was important that the data not only have evidence for technical adequacy and validity but also amenable to quick and systematic collection from all students. Although rating scale reports can be limited by the context or relationship with the child and/or informant and are generally of little utility for determining etiology or intervention planning (McConaughy 1993), rating scales are efficient, relatively objective and reliable, allow for a collection of information about a broad range of observations and teacher interactions, and indicate risk status (Feil, Severson, & Walker, 2002). Finally, behavioral rating scale data can be used to estimate symptom severity across times, observers, and settings within a problem-solving model (Busse, 2005).

After a review of the literature and existing instruments, it was determined that existing comprehensive rating systems such as the Behavior Assessment System for Children (BASC) – Second Edition (Reynolds & Kamphaus, 2004) or Child Behavior Checklist system (Achenbach, 2002) were too long for our purpose of screening all students. For example, the BASC – Second Edition Teacher Rating Scale takes about 10 to 15 minutes to complete. It is recommended that teachers complete it on one occasion with minimal distractions (BASC; Reynolds and Kamphaus, 2004). This time commitment does not include scoring, and thus, would be impractical for use as a class- or school-wide screening instrument.

The BSC was developed as a general purpose measure in consultation with district teachers and behavioral experts. A key in the development of our screener was weighing the need for brevity while still tapping a range of important classroom behaviors with adequate psychometric properties. The primary purpose of the BSC is to help teachers identify students early in the year who have behavioral issues who may benefit from general education interventions within the PSM. In this study, we sought to examine the technical adequacy of this measure in terms of, a) the inter-rater and internal consistency reliability and b) the predictive validity of the BSC with subsequent behavioral and academic difficulties. Measures of achievement are included in our examination of validity because of the interconnected nature of student behavior and academic performance, and because of the paucity of objective measures of behavior.

METHOD

Participants

The subjects of this study included 22,056 kindergarten through grade 8 students in the Minneapolis Public Schools (MPS), a large urban district located in Minnesota. Of these students 49.9% were

female and 51.1% were male. The racial/ethnic distribution of the sample was: 40% African American, 28% White American, 16.6% Hispanic American, 11.8% Asian American, and 3.6% Native American. Approximately 68% of the sample received free or reduced lunch services.

Measures

Behavior Screening Checklist. When first implemented in the District during the 1998-99 academic year, the BSC was comprised of 10 items; items regarding property destruction and out of place behavior were added during the first year based on teacher and staff feedback. The 12 items included in the BSC were grouped into 3 categories: *Classroom Behaviors*, *Externalizing Behaviors*, and *Socialization* (See Appendix A). The *Classroom Behaviors* section included four items: evaluating student attention, following directions, completing work, and class involvement. The *Externalizing Behaviors* section included: physical behavior toward others, verbal behavior toward others, physical behavior toward property, and out of place. The last section of the BSC, *Socialization*, also had four items: adult interactions, peer interactions, coping with change, and projected self-image. On each item staff was asked to rate the student on a scale of 1 to 5, with 1 representing the most appropriate behavior related to that area and a 5 associated with significant concerns. When introducing the BSC to teachers it was emphasized that this tool is a screening measure, and should be thought of as a prompt to think of all students using the same questions. If questions arose about a particular rating, teachers were encouraged to go with their first impression.

Thirty-six was used as the BSC cut-score for those students in need of additional classroom interventions. This score was initially chosen as the cut score because it was the point in the distribution at which approximately 5% of the student population fell at or above this score. This is similar to the level of students considered to be in need of tertiary prevention or intense individualized interventions in some systems of Positive Behavior Support (Sugai & Horner, 2002) or Response-to-Intervention Models (Batsche, 2005). It was also thought that this targeted a manageable number of students for additional intervention. For these students teachers were directed to begin interventions related to the problem of concern within their classrooms, and to determine their responses to these interventions.

Northwest Achievement Levels Test. The Northwest Achievement Levels Test[1], or NALT, is a standardized multiple-choice test that was administered to all students in the district in grades 2 through 7, and 9 in the areas of reading and mathematics. From the item bank, district personnel chose items that corresponded to the district's curriculum standards, allowing for measurement of students' knowledge and skills with respect to these standards. The NALT is a *levels-test* with national growth norms and, therefore, student performance may be linked from level-to-level and grade-to-grade. Further, it is possible to measure growth over time at the student, classroom, school, and district levels. The norms are based on a sample of over 1 million students. A student's performance is based on two factors: the number of correct answers and the difficulty of the questions. Each student's normal curve equivalent (NCE) score was used in the analysis. NCE scores for reading ranged from 1 to 99 with a mean of 48.7 and a standard deviation of 21.8. For math the NCE scores ranged from 1 to 99 with a mean of 54.6 and a standard deviation of 22.0.

Student behavior: Suspensions and attendance. The total number of suspensions from school during the entire academic year was used as one criterion measure of student behavior. Suspension scores for students in our sample ranged from 0 to 21 with a mean of 0.35 and a standard deviation of 1.24. Student attendance was calculated by dividing the number of days absent from school by the total number of

1 Further information may be found at http://www.nwea.org/

days enrolled (days present and absent). The percentage of days absent in the sample ranged from 0% to 77.4% with a mean of 4.9% and a standard deviation of 5.8%.

PROCEDURES

BSC data were collected in the fall. Teachers completed the BSC for all students in their classrooms by October 1st of the academic year. Data were collected district-wide as part of the PSM process described in the introduction. Teachers were asked to complete screeners for all students, typically all at one time. While they were encouraged to go with their first impression, expanded scoring definitions were available. Other criterion measures were collected from District records.

RESULTS

Reliability Analyses

Inter-rater reliability. Inter-rater reliability and agreement were used to examine this form of reliability. At the time the BSC was first implemented in the District, a collaborative teaching project was in place in which two teachers were assigned to every third grade classroom. These teachers shared instructional responsibilities for their classes, which presented an optimal situation for checking inter-rater agreement of the behavior screener. Research staff recruited six pairs of collaborative teachers to complete a behavior screener with entire classes of students (2 raters for 143 students). These analyses were run with the 10-item version of the BSC. An inter-rater reliability coefficient was calculated for each pair as well as the percent agreement for those below and above the district cut-score[1] identifying students in need of additional classroom interventions.

Inter-rater reliability coefficients were calculated for six pairs of 3rd grade teachers who taught 143 students in six classrooms. Sample sizes for those ratings ranged from 17 to 28 students per classroom. All coefficients were significant at or below .01. Coefficients ranged from .659 to .965, with a mean rating of .825 (see Table 1). Inter-rater agreement between teachers for those students below and at or above the cut-score was 91%.

TABLE 1. *Inter-rater reliability coefficients for six pairs of teachers*

Teacher Group	Number of Students	Reliability Coefficient
Pair 1	25	.865
Pair 2	25	.965
Pair 3	27	.847
Pair 4	17	.659
Pair 5	28	.819
Pair 6	21	.796
Average Correlation		.825

Internal Consistency Reliability

Chronbach's alpha was used to examine the internal consistency reliability of the BSC. Chronbach's alpha is reported for the original 10-item version of the BSC (N = 143) and the current 12-item BSC.

1 Thirty-six is the cut score used with the 12-item BSC. A cut score for the 10-item BSC that corresponds with the same portion of the distribution was 31.

The distribution of the sample used in analysis of the 12-item BSC by grade may be found in Table 2. The alpha coefficient for the original 10-item BSC was .93. On the more recent 12-item version alpha coefficients ranged from .92 to .95 for grades kindergarten to grade 8 (kindergarten: .93; 1st grade: .92; 2nd grade: .93; 3rd grade; .93, 4th grade; .93, 5th grade: .93; 6th grade: .94; 7th grade: .95; and 8th grade: .94).

Descriptive Analyses

Descriptive data for the entire sample of students are presented in Tables 2 and 3. Scores on the BSC ranged from 12 (which represents a rating of "1" on each item) to 60 (a rating of "5" on each item). The mean total score on the Behavior Screener Checklist was 20.37 with a standard deviation of 8.40. The distribution of scores was positively skewed, with most students earning scores in the low-end of the scale, indicating few behavioral concerns (See Table 2). Approximately 74% of the students scored at or below 24 on the BSC, while 5.2% scored above 36, the cut-off score for initiation of classroom interventions.

TABLE 2. *BSC Descriptive Statistics by Grade*

Grade	Mean	N	Std. Deviation	Minimum	Maximum
Kdg	19.68	2607	7.69	12	60
1	19.29	2623	7.55	12	56
2	19.84	2710	7.83	12	57
3	19.53	2431	8.01	12	59
4	20.17	2443	8.01	12	59
5	19.43	2521	7.62	12	57
6	21.87	2192	9.01	12	56
7	22.33	2197	9.95	12	60
8	21.82	2332	9.37	12	60
Total	20.37	22056	8.40	12	60

Note: Distribution Characteristics of the total Sample Skewness = 1.297 (SE = .016); Kurtosis = 1.711 (SE = .033)

The means, standard deviations and sample sizes at each grade level are presented in Table 2. It is apparent that there is a gradual increase in scores on the BSC as students grow older, with the highest scores recorded for students in the middle school years. Behavior screening scores also vary by gender and ethnicity (Table 3). Male students scored higher than female students. BSC scores are somewhat higher for African-American and Native-American students. Relatively lower scores were obtained for Asian-American and White-American students.

Predictive validity analyses

Predictive criterion validity of the BSC was studied by using separate correlation analyses for students in K to 5 (N = 14,335) and grades 6 to 8 (N = 6,721), corresponding with the change in school level from elementary to middle school. It was anticipated that BSC scores would vary as a function of school level, with more reported difficulties for older students. Fall BSC data were correlated with criterion measures (NALT, suspensions, attendance) collected in the spring of the same academic year.

TABLE 3. *BSC Descriptive Statistics by Grade*

Variable	Mean	N	SD	Minimum	Maximum
Gender					
Female	18.71	10995	7.32	12	60
Male	22.02	11061	9.06	12	60
Total	20.37	22056	8.40	12	60
Ethnicity					
Native American	22.58	783	8.82	12	59
African American	23.45	8828	9.29	12	60
Asian American	17.37	2609	6.29	12	60
Hispanic	19.06	3664	7.15	12	58
White	17.73	6172	6.77	12	59
Total	20.37	22056	8.40	12	60

TABLE 4. *Correlations between BCS and criterion variables (K-5 sample)*

	Number of Suspensions	Number of Absences	NALT Reading NCE	NALT Math NCE
Behavior Screener	0.28*	0.19*	-0.39*	-0.42*
N	15335	15335	6135	6117
Number of Suspensions		0.18*	-0.20*	-0.19*
N		15335	6135	6117
Number of Absences			-0.13*	-0.18*
N			6135	6117
NALT Reading NCE				0.80*
N				6092

*$p < .001$

TABLE 5. *Correlation between BSC and Criterion Variables (6-8 sample)*

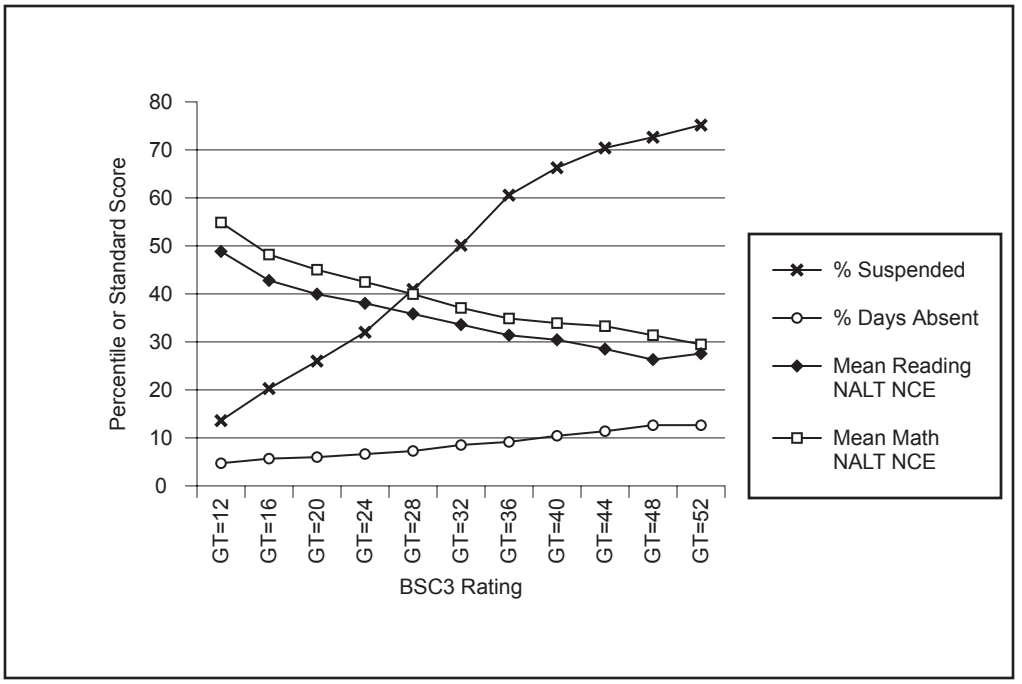
	Number of Suspensions	Number of Absences	NALT Reading NCE	NALT Math NCE
Behavior Screener	0.51*	0.36*	-0.45*	-0.48*
N	6721	6721	3657	3357
Number of Suspensions		0.46*	-0.30*	-0.34*
N		6721	3657	3357
Number of Absences			-0.20*	-0.26*
N			3657	3357
NALT Reading NCE				0.84*
N				3320

*$p < .001$

Because of the skewed distribution of BSC scores and the ordinal nature of the scale, a Spearman's rho correlation was used for the measure of predictive validity for both the K-5 and 6-8 samples. The total score on the BSC administered in the fall was correlated with the end of the year number of suspensions, percentage of days absent from school, and district achievement scores in reading and math from spring test administration. For both groups the correlations between all of these variables were significant at the .01 level (Tables 4 and 5). Correlations between the BSC and criterion variables were higher in all cases for the grade 6-8 group. There were moderate correlations between fall BSC scores and the number of suspensions and reading and math achievement in the spring. The correlation between the BSC and suspensions was higher for the grade 6-8 group (.28 and .51, respectively, K-5 and grades 6-8). Correlations between the achievement measure and the BSC were also better for the grade 6-8 group (an average correlation of -.40 for K to grade 5; -.46 for the upper grade level group).

A visual analysis of the predictive nature of the BSC was also conducted. A cross-tabs procedure was used to display scores on the BSC with the number of suspensions, achievement scores in reading and math, and total days absent from school. This analysis, presented in Figure 1, was conducted because we were interested in examining how higher scores on the fall BSC interacted with the related

Figure 1. Percentage of students suspended, median reading and math NALT NCE scores, and average percentage of days absent for groups of students scoring at or above selected BSC total scores.*



* Values on Y-axis represent three different scales associated with suspensions, achievement, and attendance.

variables. Note that the vertical axis in Figure 1, which ranges from 0 to 80, represents three scales: percentage of students suspended; mean NCE scores in reading and math for each BSC group, and average percentage of days absent from each BSC group. There appears to be a linear relationship between fall BSC scores and the likelihood of being suspended during the year. While only 13.6% of the total sample was suspended, 60.3% of those students scoring 36 or above on the fall behavior screener were suspended.

The relationship between the BSC scores and achievement scores appears to be greater at the lower end of the scale, but continues in an inverse manner as the behavior scores rise. The student group as a whole achieved over 6 NCE's higher in reading and math than those scoring 16 or above on the BSC, and 17 to 19 NCE's higher in reading and math than those scoring 36 or above on the BSC. This difference can be largely accounted for by the average NALT NCE score for reading being above 60 for those students scoring 12 on the BSC (mean score = 61.0 for reading and 67.6 for math (N=2,007 and 1,949, respectively). The mean score for those students scoring 16 on the BSC was 51.1 for reading (N=449) and 57.7 for math (N=485). The relationship between the fall BSC scores and percentage of days absent from school was not strong. There is a gradual increase in percentage of absences as the BSC scores rise, before a plateau around a score of 48.

DISCUSSION

Student behavior figures prominently in our schools, with links to academic performance and long-term social outcomes for students. The most frequent or common approach to managing student behavior is reactionary, rather than proactive, in nature. The use of harsh disciplinary practices – zero tolerance policies – has increased, despite the absence of evidence for the effectiveness of these approaches for improving student behavior or school climate. Indeed, these practices may negatively affect student achievement and school completion, and there is evidence of socioeconomic and racial disparities in the disciplinary management of students in our nations' schools. The logic of preventing or intervening early with problems, whether academic or behavioral, before these problems become severe and sometimes intractable, is obvious; however, this requires a major shift in thinking and practice. The knowledge base for screening and tracking emotional and behavioral problems is not well-established. The focus of this study was to examine a classroom behavior screener, the BSC, developed for use in the PSM utilized by the Minneapolis Public Schools. The purpose of the BSC was to identify students with behavior needs to inform general education interventions as part of the PSM.

One of the first questions in the development and use of a new measure relates to the technical adequacy. The inter-rater reliability coefficients examined in this study were technically adequate, as was the inter-rater agreement for teacher ratings of those below and at or above the cut score signaling the need for additional intervention. It should also be noted these coefficients tended to be higher than those often obtained for cross-informant ratings of student behavior. In a meta-analysis of 119 studies, Achenbach, McConaughy, and Howell reported that the average correlation between respondents of similar roles, such as 2 teachers, was .60, while the correlation among respondents of different roles was considerably lower (reported in McConaughy, 1993b). The internal consistency coefficients were extremely high and also documented technical adequacy in the area of reliability.

There are concerns over bias and disparity in both disciplinary actions and special education placements. The pattern of scores on the BSC is similar to that found in other research: African American, Native American, middle school, and male students received higher scores on the BSC. Screening with an instrument like the BSC will not address concerns over misinterpretation of student behavior due to

cultural or racial differences; however, screening with a focus on early intervention is a step in the right direction and may help attenuate some of the poor outcomes associated with harsh disciplinary policies (e.g., loss of instructional time, dropout, special education placements). Donovan and Cross (2002) argued the overall impact of early intervention models will be to reduce the numbers of students who fail, which in turn meets the educational needs of students of color.

BSC predictive validity was evidenced by moderate correlations between fall ratings and behavioral performance over the course of the school year and academic assessments administered in the spring. In addition, the results showed a clear and consistent relationship between scores on the BSC and the likelihood of suspension. Those students scoring 36 or above in the fall have a 50% chance of being suspended over the course of the school year.

This study provided some evidence to support the use of the BSC as an efficient, reliable, and valid screening tool to aid teachers and other school staff in identifying students likely to encounter significant behavioral difficulties, particularly those at-risk for suspensions. As noted by Gersten and Dimino (2006), one of the failures of the prereferral model of the late 1980's was that, "Student Study Teams typically made intervention recommendations on the basis of the classroom teacher's description of the student's academic or behavioral performance. Many of the descriptions were anecdotal; few were data based." (Gersten & Dimino, 2006, p. 101). The ability to use the BSC to identify students in need of interventions may help teachers to examine their students' behaviors earlier in the school year, and provides a measure which prompts teachers to rate all students' behaviors on objective items. This should also lessen the bias introduced by teacher nomination, and will hopefully provide for intervention before a cycle of frustration or failure is created.

In practical terms, the provision of teachers with a tool which in less than a minute can help them identify students likely to present significant behavioral or academic concerns over the course of the school year seems worthy. On average those students scoring at or above 36 on the BSC were over 4 times as likely to be suspended as the total sample, were likely to be absent almost twice as much, and scored around one-half standard deviation lower on measures of achievement. This knowledge can be used to target students, classrooms or programs for intervention programs before these problems arise.

There is a tension between maintaining scientific rigor while conducting research in real-world settings. This study is no exception. The strengths of this study – the design and implementation of a screening measure for use in a PSM and subsequent research with a large, diverse school population – are also weaknesses. There was little time between development and implementation of the BSC as a district-wide K-8 screener; analyses of its adequacy were post-hoc. Future study should address refining the BSC, including examination of individual items, the need for and utility of age-specific versions, and the establishment of norms and cut-scores for screening purposes. In addition, validity studies relating the BSC to other forms of behavioral assessment are needed as further evidence of its validity. Furthermore, although behavior rating scales in general are ideal for screening purposes, due to their relative efficiency, objectivity, and reliability (Feil, Severson, & Walker, 2002), these scales do not provide information regarding etiology or intervention design (McCanaughy, 1993). This is true of the BSC as well. The purpose of this screener was to identify students *in need of early intervention*; it does not provide information to teachers on how to intervene nor is it a sensitive measure of how a specific behavior changes due to the intervention. Future research must address the critical link between screening and intervention planning and implementation within the PSM.

There are a range of demands on the educational system, including increased accountability, *No Child Left Behind*, budgeting concerns, and discipline. In the midst of these demands, however, it is

clear that the traditional wait-to-fail model used for academic and behavioral difficulties does not meet the needs of our students or the system as a whole. While wait-to-fail is the past and present, screening and early interventions are the future. In their call for a "paradigm shift," Reschly and Ysseldyke (2002) argue that data-based problem solving is crucial to system reform. A response-to-intervention model promotes the early identification of students' academic and behavioral difficulties and provides interventions and monitoring with increasing intensity until effective interventions and environments are found for students. Although further research is needed, the behavior screener described is a model for addressing significant behavioral issues in a proactive way that is linked to problem solving and better outcomes for at risk children and youth.

Appendix A

BEHAVIOR SCREENING CHECKLIST III

Student Name:_____ Student ID#:_____ Date:_____

Rate the student on the following continuum:

I. AREA: CLASSROOM BEHAVIORS

Attention:

1	2	3	4	5
Consistently attends to classroom activities		Sometimes follows along with classroom activities		Rarely follows along with classroom activities

Follows Directions:

1	2	3	4	5
Consistently follows rules		Sometimes follows rules		Rarely follows rules

Completing Work:

1	2	3	4	5
Consistently completes work independently		Sometimes completes work independently		Rarely completes work independently

Class Involvement:

1	2	3	4	5
Participates Well		Sometimes participates		Rarely participates

II. AREA: EXTERNALIZING BEHAVIORS

Physical Behavior Toward Others:

1	2	3	4	5
Physically appropriate		Occasionally physically appropriate		Rarely physically appropriate

Verbal Behavior:

1	2	3	4	5
Uses appropriate verbal behavior		Sometimes uses appropriate verbal behavior		Rarely uses appropriate verbal behavior

Physical Behavior Toward Materials or Property:

1	2	3	4	5
Is consistently respectful of Materials or Property		Is sometimes respectful of Materials or Property		Is rarely respectful of Materials or Property

Out of Place:

1	2	3	4	5
Remains in assigned area		Sometimes remains in assigned area		Rarely is in assigned area

III. AREA: SOCIALIZATION

Coping with Change:

1	2	3	4	5
Handles change appropriately		Occasionally handles change appropriately		Rarely handles change appropriately

Adult Interactions:

1	2	3	4	5
Seeks positive relationships		Sometimes seeks positive relationships		Rarely seeks positive relationships

Peer Interactions:

1	2	3	4	5
Seeks positive relationships		Sometimes seeks positive relationships		Rarely seeks positive relationships

Scores: Classroom Behaviors____ Externalizing Behaviors____ Socialization____ Total_____

REFERENCES

Achenbach, T. M. (2002). *Achenbach System of Empirically Based Assessment.* Retrieved June 13, 2005 from ASEBA website, *www.aseba.org/index/html.*

Alexander, K. L., Entwisle, D. R., & Dauber, S. L. (1993). First-grade classroom behavior: Its short-and long-term consequences for school performance. *Child Development, 64*(3), 801-813.

Barbour, C. B. (2006). Engaging RTI: Notes from the field. Presentation at the University of Cincinnati, May 12, 2006.

Batsche, G. M. (2005, October 4). *Problem-Solving and Response to Intervention: Maximizing Opportunities for Student Progression*. Paper presented at the LaGrange Area Department of Special Education. Retrieved May 30, 2006, from http://www.ladse.org/prof%20dev.htm

Bock, S. J., Tapscott, K. E., & Savner, J. L. (1998). Suspension and expulsion: Effective management for students? *Intervention in School and Clinic, 34*(1), 50-52.

Bradley, R., Danielson, L., & Hallahan, D. P. (2002). (Eds). *Identification of learning disabilities: Research to practice*. Mahwah, New Jersey: Lawrence Erlbaum

Busse, R. T. (2005). Rating Scale Applications within the Problem-Solving Model. In R. Brown-Chidsey (Ed.) *Assessment for Intervention: A Problem-Solving Approach* (pp. 200-218). New York: Guilford.

Casat, C. D., Norton, & Whitsel (1999). Identification of elementary school children at risk for disruptive behavioral disturbance: Validation of a combined screening method. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*, 1246-1253.

Deno, S. L. (2002). Problem solving as "Best Practice." In A. Thomas & J. Grimes (Eds.), *Best Practices in School Psychology IV* (pp.37-55). Bethesda, MD: NASP Publications.

DeRidder, L. M. (1990). The impact of school suspensions and expulsions on dropping out *Educational Horizons, 68*, 153-157.

DeVoe, J. F., Peter, K., Kaufman, P., Miller, A., Noonan, M., Snyder, T. D., & Baum, K. (2004). *Indicators of School Crime and Safety: 2004* (NCES 2005-002/NCJ 205290). U.S. Department of Education and Justice. Washington, DC: U.S. Government Printing Office.

Donovan, M. S., & Cross, C. T. (Eds.). (2002). *Minority students in special and gifted education.* Washington DC: National Academy Press.

Feil, E. G., Severson, H. H., & Walker, H. M. (2002). Early screening and intervention to prevent the development of aggressive, destructive behavior patterns among at-risk children. In M. Shinn, H. Walker, & G. Stoner (Eds.), *Interventions for Academic and Behavior Problems II: Preventive and Remedial Approaches* (pp 143-166. Bethesda, MD: NASP Publications.

Fuchs, D., & Fuchs, L. S. (2006). Introduction to Response to Intervention: What, why, and how valid is it? *Reading Research Quarterly 41*(1), 93-99.

Gallup Poll News Service (September 17, 2002). *Parents concerned about school safety*. Available on-line: *www.gallup.com*

Gallup Poll News Service (October 26, 2004). *Americans on local schools: "Show us the money*". Available on-line: *www.gallup.com*

Gersten, R., & Dimino, J. A. (2006). RTI (Response to Intervention): Rethinking special education for students with reading difficulties (yet again). *Reading Research Quarterly 41*(1), 99-108.

Ikeda, M. J., Grimes, J., Tilly, W. D., Allison, R., Kurns, S., & Stumme, J. (2002). Implementing an intervention-based approach to service delivery: A case example. In M. Shinn, H. Walker, & G. Stoner (Eds.), *Interventions for Academic and Behavior Problems II: Preventive and Remedial Approaches* (pp 53-69. Bethesda, MD: NASP Publications.

Kaplan, D. S., Peck, B. M., & Kaplan, H. (1997). Decomposing the academic failure-dropout relationship: A longitudinal analysis. *Journal of Educational Research, 90*(9), 331-343.

Lau, M. Y., Sieler, J. D., Muyskens, P., Canter, A., VanKeuren, B., & Marston, D. (2006). Perspectives on the use of the problem-solving model from the viewpoint of a school psychologist, administrator, and teacher from a large Midwest urban school district. *Psychology in the Schools, 43*(1), 117-127.

Marston, D., Muyskens, P, Lau, M, & Canter, A. (2003). Problem-solving model for decision making with high-incidence disabilities: The Minneapolis experience. *Learning Disabilities Research and Practice, 18*(3), 187-200.

Marston, D. (2002). A functional and intervention-based approach to establishing discrepancy for students with learning disabilities. In R. Bradley, D. P. Hallahan, & L. Danielson (Eds.), Identification of learning disabilities: Research to practice. NJ: Lawrence Erlbaum Associates, Inc..

McClelland, M. M., Morrison, F. J., & Holmes, D. (2000). Children at risk for early academic problems: The role of learning-related social skills. *Early Childhood Research Quarterly, 15*(3), 307-329.

McConoughy, S. H. (1993a). Evaluating behavioral and emotional disorders with the CBLC, TRF, and YSR cross-informant scales. *Journal of Emotional and Behavioral Disorders, 1*, 40-52.

McConaughy, S. H. (1993b). Advances in empirically based assessment of children's behavioral and emotional problems. *School Psychology Review*, *22*(2), 285-307.

McEvoy, A., & Welker, R, (2000). Antisocial behavior, academic failure, and school climate: A critical review. *Journal of Emotional and Behavioral Disorders, 8*, 130-140.

Naglieri, J. (2005). "Current perspectives on behavioral interventions" lacks balance. *NASP Communique, 33*(5).

O'Neil, R., Welsh, M., Parke, R. D., Wang, S., & Strand, C. (1997). A longitude assessment of the academic correlates of early peer acceptance and rejection. *Journal of Clinical Child Psychology, 26*, 290-303.

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*(1), 27-35.

Rausch, M. K. & Skiba, R. (2004). Unplanned outcomes: Suspensions and expulsions in Indiana. Education Policy Briefs, 2(2). Available On-line: *http://ceep.indiana.edu/ChildrenLeftBehind/*

Reschly, D. J. (2005). Viewpoint. Child outcomes and school psychology service delivery: Response to Naglieri. *NASP Communique*, 33(5).

Reschly, D. J., & Tilly, W. D. (1999). Reform trends and system design alternatives. In D. J. Reschly, W. D. Tilly, & J. P. Grimes (Eds). *Special Education in Transition: Functional assessment and noncategorial programming*. Longmont, CO: Sopris West.

Reschly, D. & Ysseldyke, J. E. (2002). Paradigm shift: The past is not the future. In A. Thomas & J. Grimes (Eds.), *Best practices in School psychology IV* (pp 3-20). Bethesda, MD: National Association of School Psychologists.

Reynolds, C. R., & Kamphaus, R. W. (2004). *Manual: Behavior Assessment System for Children*. Circle Pines, MN: American Guidance Service.

Skiba, R. & Peterson, R. (1999). The dark side of zero tolerance: Can punishment lead to safe schools? *Phi Delta Kappan, 80*(5), 372-382.

Skiba, R., & Rausch, M. K. (2006). School disciplinary systems: Alternatives to suspension and expulsion. In G. Bear & K. Minke (Eds). *Children's Needs III: Understanding and addressing the developmental needs of children* (pp. 87-102). National Association of School Psychologists.

Sugai, G, & Horner, R. H. (2002). Introduction to the special series on Positive Behavior Support in Schools. *Journal of Emotional and Behavioral Disorders*, 10, 130-135.

Sugai, G., Horner, R. H., Dunlap, G., Hieneman, M., Lewis, T. J., Nelson, C. M., Scott, T., Liaupsin, C., Sailor, W., Turnbull, A. P., Turnbull, H. R. III, Wickham, D., Reuf, M., & Wilcox, B. (2000). Applying positive behavioral support and functional behavioral assessment in the schools. *Journal of positive behavioral interventions*, *2*, 131-143.

U.S. Department of Education, Office of Special Education and Rehabilitative Services (2002). *A new era: Revitalizing special education for children and their families.* Washington DC, U.S. Department of Education.

Walker, H. M., Horner, R. B., Sugai, G., Bullis, M., Sprague, J.R., Bricker, D., & Kaufman, M. J. (1996). Integrated approaches to preventing antisocial behavior patterns among school-age children and youth. *Journal of Emotional & Behavioral Disorders*, *4*(4), 194- 210.

Woodward, L. J., & Fergusson, D. M. (2000). Childhood peer relationship problems and later risks of educational under-achievement and unemployment. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 41*(2), 191-201.