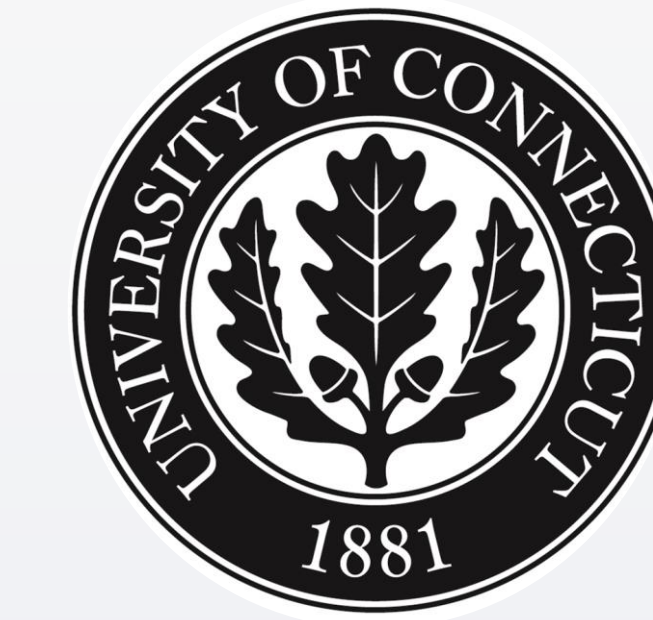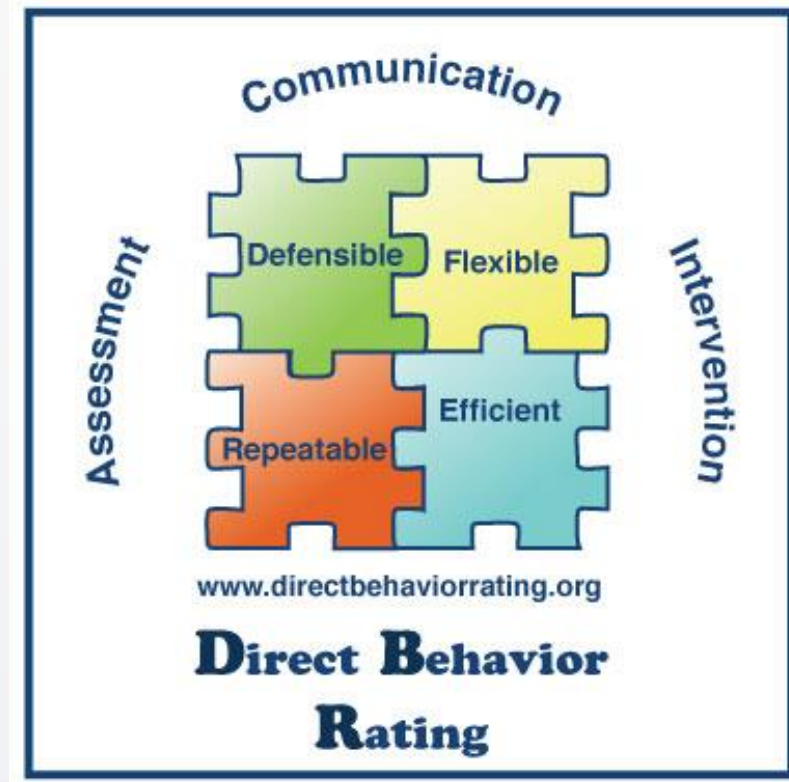# An Evaluation of Universal Screening Methods to Identify Behavioral Risk

Faith G. Miller[1], Megan E. Welsh[1], Austin H. Johnson[1], Sandra M. Chafouleas[1],
T. Chris Riley-Tillman[2], & Gregory A. Fabiano[3]

University of Connecticut[1], University of Missouri[2], University at Buffalo[3]

Introduction

Background

With the advent of multi-tiered systems of support, the use of reliable and valid screening procedures to identify students at-risk for school failure has become increasingly important. A wide variety of commercially-available programs have helped create parameters around screening within academic domains. However, behavioral screening has received considerably less attention. Given relevance of behavior toward overall student success, it is critical that systematic research is carried out to develop and evaluate assessment strategies to meet key needs within systemic problem-solving models of service delivery. Specifically, there is a pressing need for assessment systems which can be used to accurately identify students at-risk for behavioral difficulties.

Objective

The purpose of the current study was to examine the relationship between universal behavioral screening methods and differences in classification accuracy between methods. In particular, four prominent screening methods were of interest in the present investigation: (a) Direct Behavior Ratings – Single Item Scales (DBR-SIS: www.directbehaviorrating.org), (b) Social Skills Improvement System (SSIS) Performance Screening Guide (Elliott & Gresham, 2007), (c) Behavior Assessment System for Children – 2, Behavioral and Emotional Screening System – Teacher Form (BESS: Kamphaus & Reynolds, 2007), and (d) office discipline referrals (ODRs).

Method

This study was conducted as a part of a larger federally-funded project designed to provide unified validation of school-based behavior assessments for screening and progress monitoring purposes.

Participants and Setting

- The analytic sample included 1974 students.

- Students were enrolled in a total of 20 different schools, including rural, suburban, and urban districts.

- Public school settings were geographically located in Connecticut, New York, and Missouri.

Table 1: *Student Demographic Characteristics*

Characteristic	n	%
Gender		
Male	1029	52
Female	945	48
Race		
White	1611	81
African American	238	12
Asian	31	1
Other	94	47
Ethnicity		
Hispanic	145	7
Non-Hispanic	1829	93
Grade		
Lower elementary (1-2)	658	33
Upper elementary (4-5)	725	37
Secondary (7-8)	591	30

Method

Measures

- *Direct Behavior Rating – Single Item Scale* (DBR-SIS)
  - DBR-SIS reflects the teacher's perception of the proportion of time a student is observed engaged in a target behavior (academic engagement, respectful, disruptive) from 0 (never) to 10 (always). Composite scores were created for each student, ranging from 0 (indicating substantial risk) – 30 (indicating little risk).

- *Social Skills Improvement System - Performance Screening Guide* (SSiS; Gresham & Elliott, 1990)
  - The SSiS Performance Screening Guide can be used to screen social and academic behaviors of all students in a class. This screener uses a scale of 1 (Substantial Difficulty), 2 or 3 (Moderate Difficulty), and 4 or 5 (Average). Student functioning is rated across four areas: Motivation to Learn, Prosocial Behavior, Math Skills, and Reading Skills.

- *Behavioral and Emotional Screening System* (BESS Teacher Form; Kamphaus & Reynolds, 2007)
  - The BESS is a brief rating scale that can be useful in screening for behavioral and emotional strengths and weaknesses in children and adolescents. The scoring of the BESS yields an overall T score; a student is considered to be at-risk if his or her score is ≥ 61.

- *Office discipline referrals (ODR)*
  - At the end of the school year, ODR data were collected for each student by month of referral. Total ODRs were then summed by time point. Previous research has suggested that 2 or more ODRs are indicative of risk (McIntosh, Campbell, Carter, & Zumbo, 2009).

Procedures

- Participants completed behavioral screenings during two-week data collection periods in the fall, winter, and spring of the 2011-12 school year. At each time point, teachers completed the three behavior assessments on a random sample of approximately 10 students in their classroom.

Results

Data Analysis

- The following classification accuracy statistics were calculated using the BESS as a criterion: sensitivity, specificity, positive predictive value, negative predictive value, and hit rate (see Table 2).

- Receiver Operating Characteristic curves were calculated using IBM SPSS Statistics (Version 21).

  - Area under the curve (AUC) statistics and corresponding 95% confidence intervals were calculated.

Results

- Sensitivity values for DBR, SSiS, Prosocial, and SSiS Motivation exceeded .80 across time points.
  - Sensitivity values for ODR data did not exceed .36 across time points
- Specificity values for ODR data exceeded .90 across time points
  - Adequate levels of sensitivity were exhibited by DBR and SSiS scales, all exceeding .70.
- AUC values for DBR and SSiS scales suggest these scales perform significantly better than chance when identifying students at-risk for behavioral concerns.
  - ODR data did not perform significantly better than chance.

Table 2: *Classification Accuracy Statistics*

Screening measure	SN	SP	PPV	NPV	Hit Rate	AUC	95% CI
Fall							
DBR Composite	.86	.71	.39	.96	.74	.86*	.84-.88
SSiS Prosocial	.86	.78	.47	.96	.79	.88*	.86-.90
SSiS Motivation	.91	.77	.48	.97	.80	.90*	.88-.91
ODR	.21	.97	.58	.84	.82	.50	.47-.54
Winter							
DBR Composite	.85	.75	.43	.96	.77	.87*	.85-.89
SSiS Prosocial	.84	.79	.48	.96	.80	.87*	.85-.89
SSiS Motivation	.90	.77	.47	.97	.79	.90*	.88-.92
ODR	.32	.94	.55	.86	.82	.49	.46-.52
Spring							
DBR Composite	.81	.72	.38	.95	.74	.84*	.82-.86
SSiS Prosocial	.83	.80	.47	.96	.80	.87*	.85-.89
SSiS Motivation	.85	.77	.44	.96	.78	.88*	.86-.90
ODR	.36	.92	.50	.87	.82	.50	.46-.53

*$p < .05$

Summary and Conclusions

- Classification accuracy statistics were similar across DBR and SSiS scales
  - DBR and SSiS scales identified a greater proportion of students at-risk than the BESS.
  - ODR data resulted in substantial under identification of students at-risk.
- Further research should examine classification accuracy statistics with alternative criterion.
  - Examination of classification accuracy statistics by grade group is also needed.
- In selecting a screening measure, high sensitivity levels are desirable such that students can be identified and interventions developed early in the process.