Using Composite Scores to Determine Behavioral Risk with Direct Behavior Rating

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

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

Introduction

In an educational climate which emphasizes the use of multi-tiered intervention and prevention frameworks for the provision of behavioral services to children, it is crucial to identify efficient screening tools that provide reliable and valid results. Direct Behavior Rating (DBR) is one such suggested tool, wherein an observer makes an estimate of the percentage of time a student was engaged in one or more target behaviors during a pre-specified observation period. While a practitioner is able to measure multiple behaviors simultaneously using DBR, it is unclear how best to combine screening information across behaviors, and what the implications of various methodologies for creating such composite scores are when constructing and analyzing Receiver Operating Characteristic (ROC) curves. This poster presentation describes results from a study of students sampled across grades and geographic regions, along with the diagnostic accuracy of DBR ratings for these students as determined utilizing different methods for composite score creation.

Method

Participants. Approximately 1900 public-school students enrolled in a total of 192 1st, 2nd, 4th, 5th, 7th, and 8th grade classrooms across three states (Missouri, New York, and Connecticut) were enrolled in this study. Prior to the Fall time point, ten students were randomly selected from each participating teacher's roster for ratings. As identified in the Fall, 52.2% of student participants were male. The racial identity of a majority of participants was White (82.5%), with 13.0% of the participants identified as African-American and 1.7% as Asian. The ethnicity of most participants was non-Hispanic (92.6%). 13% of students received special education supports as part of a formal special education identification.

Procedures. Teachers completed the BASC-BESS (Kamphaus & Reynolds, 2007) and DBR Single Item Scales (DBR-SIS) for students during the Fall, Winter, and Spring of the 2011-12 academic year. Using DBR-SIS, three behaviors (Academically Engaged, Disruptive, and Respectful) were measured on an 11-point scale, with observations structured such that a first group of five students was rated twice-daily for five days, followed by a second group of five students.

Composite Scores. A mean DBR rating was calculated if a given student was rated at least six of a possible ten times within a time point. These means were then aggregated into a simple sum composite score, with Disruptive behavior reverse-coded such that a higher score indicated less-disruptive behavior.

Analyses. Analyses were conducted in R 2.15.3 with the pROC, plyr, and ggplot2 packages. ROC curves were constructed for all combinations of grade group and time point, with a combined-gender BESS *t*-score above 60 used as the criterion for risk. Bootstrapped sensitivity (SN), specificity (SP), and AUC statistics were generated using 10000 replications, with SN and SP calculated for thresholds in increments of 0.10. Optimal thresholds were determined using a researcher-created algorithm with rules specified in Table 1 using median statistic values.

Preparation of this poster was supported by a grant from the Institute for Education Sciences (IES), U.S. Department of Education (R324A110017). Please direct all correspondence to Dr. Sandra Chafouleas, University of Connecticut (sandra.chafouleas@uconn.edu).

Results

Figure 1. *Values and 95% Confidence Intervals for Sensitivity and Specificity Statistics for Composite Scores by Grade Group and Time Point.*

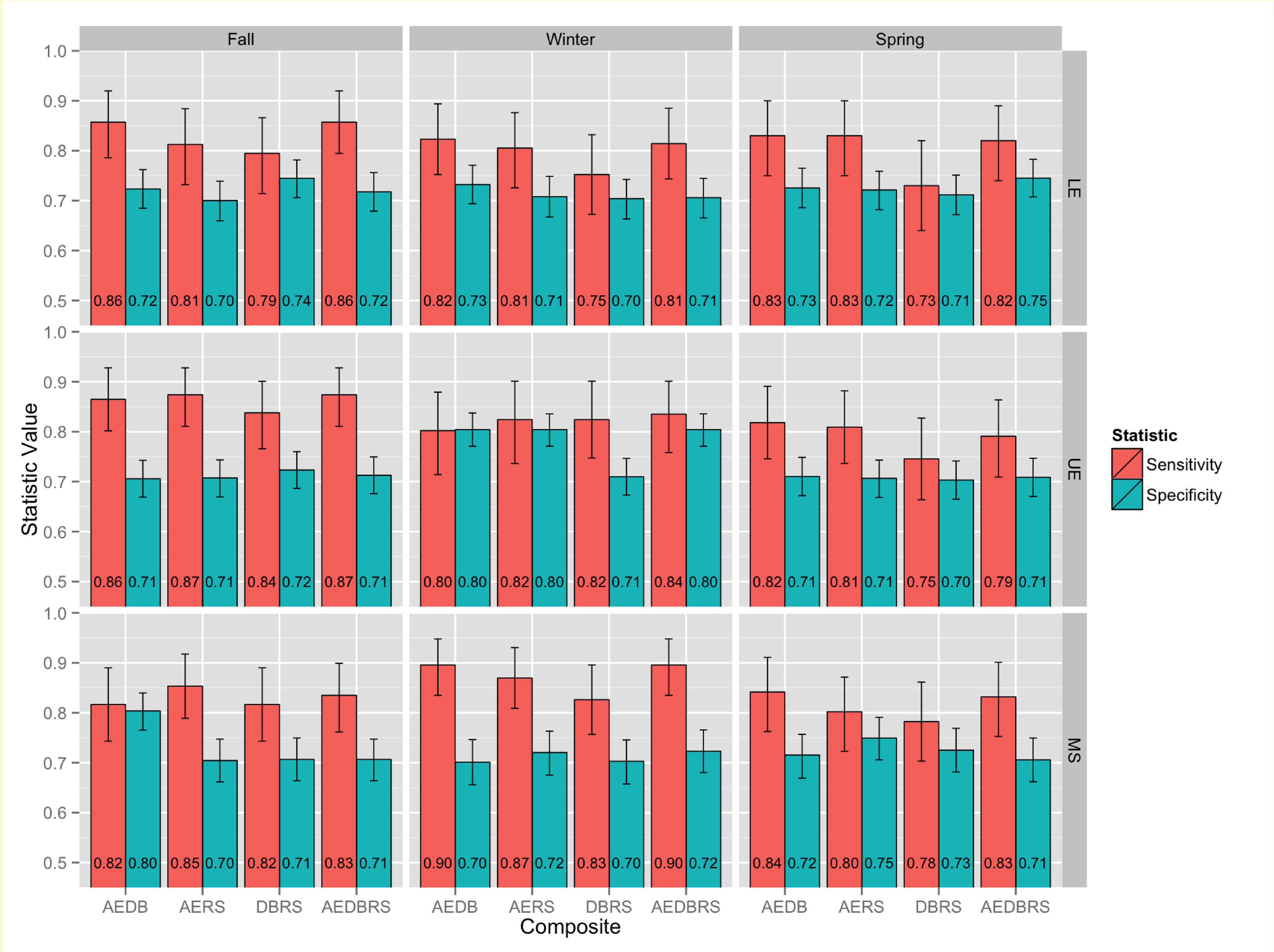

Figure 2. *Values and 95% Confidence Intervals for Area Under the Curve (AUC) Statistics for Composite Scores by Grade Group and Time Point.*

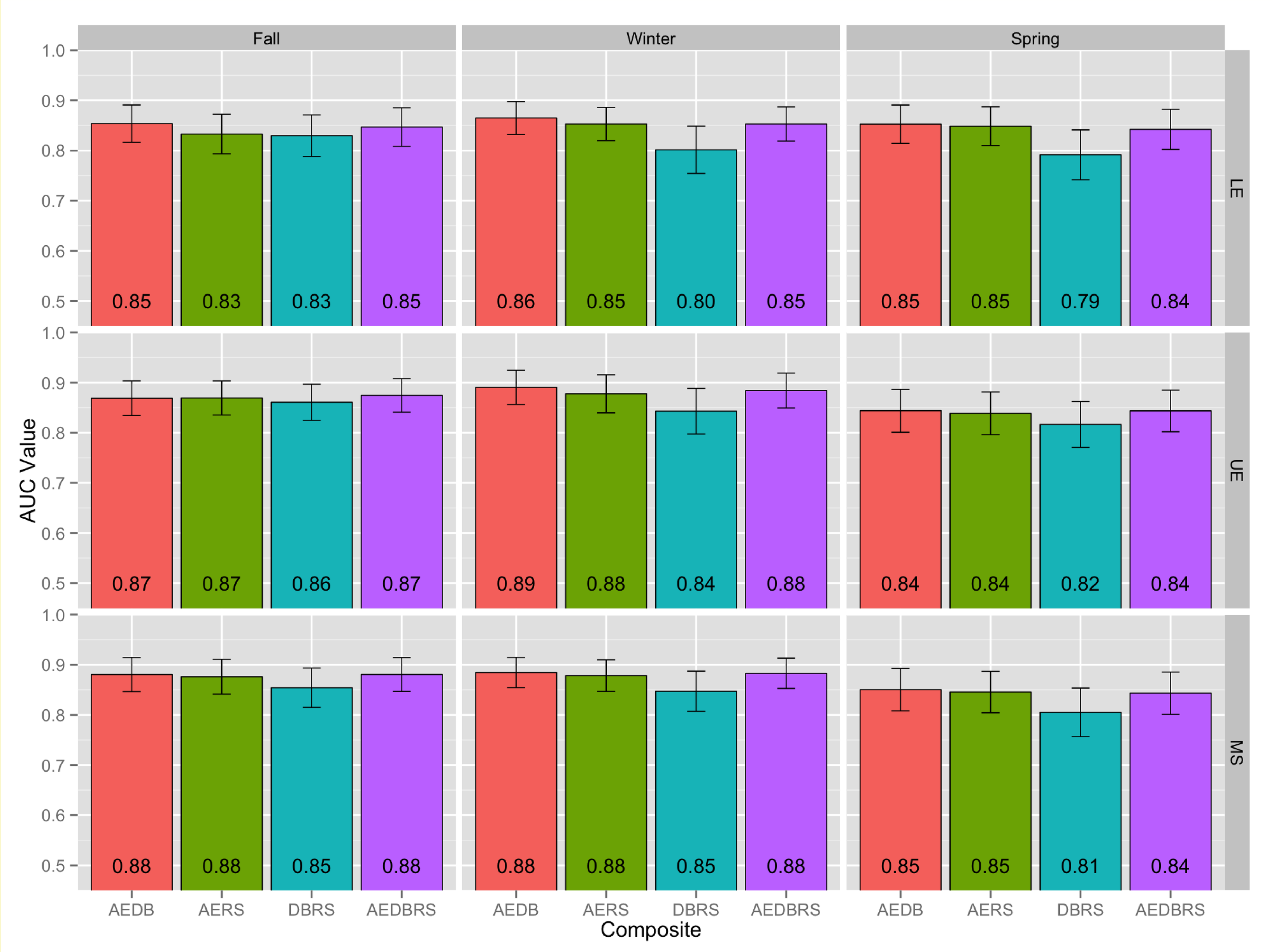

Table 1. *Rules utilized for determining optimal threshold for each grade level and time point.*

	Sensitivity (SN)	Specificity (SP)
Best	0.9	0.9
	0.9	0.8
	0.9	0.7
	0.8	0.8
	0.8	0.7
	0.7	0.7
Worst	Smallest SN/SP discrepancy	

Table 2. *Correspondence between Type and risk change when adding RS to AEDB composite.*

	Composite	
Type	**AEDB**	**AEDBRS**
A	Risk	Risk
B	Risk	No risk
C	No risk	Risk
D	No risk	No risk

Table 3. *Number of students categorized as at-risk or not-at-risk according to AEDB and AEDBRS composites, using Types outlined in Table 2.*

At-risk on BESS

	Grade Group		
Type	**LE**	**UE**	**MS**
Fall			
A	96	95	89
B	0	1	0
C	**0**	**2**	**2**
D	Z	13	18
Winter			
A	90	71	101
B	3	2	2
C	**2**	**5**	**2**
D	18	13	10
Spring			
A	80	87	83
B	3	3	2
C	**2**	**0**	**1**
D	15	20	15

Not at-risk on BESS

	Grade Group		
Type	**LE**	**UE**	**MS**
Fall			
A	130	148	87
B	**13**	**20**	**0**
C	16	16	43
D	358	387	313
Winter			
A	126	97	105
B	**6**	**15**	**13**
C	19	15	2
D	342	445	277
Spring			
A	121	148	106
B	**18**	**11**	**11**
C	8	12	15
D	359	378	279