A Longitudinal Study of Two Teacher-Report Screening Measures for Student Mental Health: Comparing the SWTRS and SAEBRS

PRESENTATION BY ANTHONY J. ROBERSON, PHD
2019 ANNUAL CONFERENCE ON ADVANCING SCHOOL MENTAL HEALTH
NOVEMBER 8, 2019
Introduction
Youth Mental Health

1 in 5 youth with MH disorder\(^1\)

Links with negative outcomes

- Substance use\(^2\)
- Adult depressive episodes\(^3\)
- Panic attacks\(^4\)
- Poorer academic achievement\(^5\)

Behavioral-academic problem interaction\(^6\)

MH among the costliest health care expenses in U.S.\(^7\)
School-based Mental Health Support

A quarter of youth estimated to be identified and provided community services

School becomes “de facto” MH care provider
- Kind and quality of services often far weaker than outpatient

Despite SPED services, ED students often remain at heightened risk for continuing behavioral/academic deterioration

Calls for improving school-based mental health care
- President’s New Freedom Commission on Mental Health

Improved intervention requires improved identification
- Traditional referral method common but problematic
Screening for Risk

Derived from medical context
World Health Organization\textsuperscript{13}:

1. Method of identifying signs and symptoms of distress
2. Efficient
3. Distinguish reasonably well from those who are and are not at risk
4. Not diagnostic for any particular condition
5. Plan must be in place for follow-up assessment and intervention
Universal Mental Health Screening

MH screening of entire school populations on common behavioral criteria gaining traction\textsuperscript{14}

Advantages\textsuperscript{15, 16}:

- Systematic way of identifying at-risk youth
- Helps identify students who would otherwise be missed
- Allows for calculation of local base rates
- Individual- and group-level comparisons possible
- Links with MTSS\textsuperscript{17}
Standards for Universal Mental Health Screening in School

1. Technical Adequacy
   ◦ Psychometric properties
   ◦ Reliability and validity evidence
   ◦ Degree to which evidence and logic support score interpretation and use

2. Usability
   ◦ Practical considerations
   ◦ Defensible to key stakeholders
   ◦ Low time and money cost

3. Appropriateness
   ◦ Match between school goals and screening procedure
   ◦ Current or future disorder identification
   ◦ Conceptualization of mental health
Conceptualizing MH for Screening

Historical focus on pathology

Increased focus on wellbeing

Evidence for value of both distress and wellbeing
  - Related but distinct concepts varying along two dimensions
  - *Dual-factor MH (DFMH)*

DFMH useful in assessing youth risk
  - Greenspoon & Saklofske (2001)
  - Suldo et al. (2008; 2011)
  - Kim, Furlong, Dowdy, & Felix (2014)
  - Dowdy et al. (2014)

Need strong wellbeing instruments to use as part of DFMH screening
Wellbeing-Focused Screening Measures

Some measures available; all limited for certain school-based screening applications with elementary students

1. **Social and Emotional Health Survey (SEHS)**\(^{25}\)
   - Length, not school-specific, self-report, adolescents

2. **Student Subjective Wellbeing Questionnaire (SSWQ)**\(^{26}\)
   - Self-report, adolescents

3. **Positive Experiences at School Scale (PEASS)**\(^{27}\)
   - Self-report, upper elementary

4. **Devereux Student Strengths Assessment-Mini (DESSA-Mini)**\(^{28}\)
   - Unidimensional

Two developing measures of interest

Brief, multi-dimensional, teacher-report, school-specific student wellbeing screeners appropriate for elementary school

1. **Social, Academic, and Emotional Behavior Risk Screener (SAEBRS)**\(^{29}\)
2. **Student Wellbeing Teacher-Report Scale (SWTRS)**\(^{30}\)
SAEBRS

Multiple studies supporting its development and validation with elementary samples\textsuperscript{29, 31, 32}

Developed based on theories linking behavior to school success

Four scores: \textbf{Academic} (6 items), \textbf{Social} (6 items), \textbf{Emotional} (7 items), and \textbf{Total} (19 items) behavior

Both adaptive and maladaptive behaviors
  \begin{itemize}
    \item Based on dual-factor evidence
    \item Reverse score maladaptive behaviors
  \end{itemize}

Strong internal consistency evidence for all scales ($\alpha \geq .80$)
SAEBRS

Correlation with other risk-classification systems
- Social Skills Improvement System ($r$ range = 0.36–0.51)
- Behavioral and Emotional Screening System ($r$ range = 0.72–0.94)

At least adequate sensitivity and specificity of each scale based on BESS risk status

Logistic regression analyses suggest significant unique contribution of each scale

Modular

Factor analytic evidence supports...
- $AB + SB = TB$
- $AB + SB + EB = TB$

Measurement invariance evidence for gender and between white and black students
SWTRS

Only pilot study completed to date\textsuperscript{30}
- Elementary school sample

Intended as teacher-report extension of “Feeling good” and “Functioning well” conception of wellbeing
- Incompatible behaviors from SIBS and SEBS
- All positively worded

EFA suggested three robust factors instead: \textbf{Academic} (6 items), \textbf{Social} (6 items), and \textbf{Emotional} (5 items) wellbeing
- Strong factor loadings (> .50)
- Communalities (> .60)
- Internal consistency ($\omega \geq .89$)
Correlations with concurrent variables:
- Percent of time spent on-task \( r \) range = 0.46–0.72
- English Language Arts \( r \) range = 0.19–0.57
- Student Internalizing Behavior Screener \( r \) range = -0.45–-0.63
- Student Externalizing Behavior Screener \( r \) range = -0.54–-0.89

Variable-Centered Dual-factor Analyses
- SWTRS scores collectively superior to SIBS and SEBS scores in explaining variance in...
  - Percent of time on-task
  - Academic achievement in math and ELA
  - Number of absences
SAEBRS has strong evidence while SWTRS only has preliminary evidence

Several conceptual and formal similarities:

- Intended use as teacher-report student mental health screener
- Behavioral domains
- Length
- Frequency-based response scale

Reverse scoring maladaptive SAEBRS items potentially critical difference...
Unknown how well reverse scoring maladaptive SAEBRS items supports integrated dual-factor interpretation

Three points of concern:

1. All previous dual-factor research used separate PTH and WB instruments
   - Novel application of DFMH logic
   - Requires additional study

2. Mixing item valence may bias response compared to uniform directionality\(^{33}\)

3. Reverse scoring may introduce construct-irrelevant variance to scores
   - Requires higher level of inference to justify construct representation\(^{34}\)
The Current Study

SWTRS intended as half of dual-factor screening assessment, not integrated

All items positively worded
  ◦ Do not require reverse scoring

Evidence that WB variables tend to be stronger predictors than problem behavior variables\textsuperscript{23}

**Two broad goals of present study:**

1. Continue development of SWTRS via structural validation and examining score-use validity evidence
2. Compare relative performance of SWTRS and SAEBRS
Research Questions

**Question 1:** What is the best fitting structural representation of the SWTRS items?

**Question 2:** How strongly do the SWTRS scores correlate with their counterpart scores on the SAEBRS and with the criterion variables of interest at Time 1?

**Question 3:** How well do the SWTRS scores function as indicators of risk as indicated by risk on the SDQ-T?

**Question 4:** How well do the cut-scores derived for the SWTRS and SAEBRS based on Time 1 data compare as concurrent and predictive indicators of risk status at Time 1 and Time 2?
Method
Participants: Teachers

Elementary school (Pre K–5th grade) teachers and students

Teachers

- $N = 18$
- 94% female
- Average age = 35.1 years ($SD = 10.7$)
- Average years teaching = 9.3 ($SD = 8.2$)
- 52% held master’s degree
- 56% White, 22% Black/AA, 17% Asian, 7% multiracial identities
Data Collection

Two collection time-points via secure online surveys

**Time 1**
- Researcher led brief in-person training with teachers covering...
- Informed consent information
- Data collection procedure
- Answers to participant questions
- One week to complete for all students in their class

**Time 2**
- Same procedure as Time 1 without in-person training component
## Participants: Students

<table>
<thead>
<tr>
<th></th>
<th>Time 1</th>
<th>Time 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>371</td>
<td>332</td>
</tr>
<tr>
<td>Median Students per Class</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>Mean Students per Class</td>
<td>20.6 (( SD = 7.8 ))</td>
<td>22.1 (( SD = 5.1 ))</td>
</tr>
<tr>
<td>Active IEP</td>
<td>4%</td>
<td>4%</td>
</tr>
<tr>
<td>504 Plan</td>
<td>4%</td>
<td>5%</td>
</tr>
<tr>
<td>Non-English First Language</td>
<td>16%</td>
<td>16%</td>
</tr>
</tbody>
</table>

52% Male  
Mean age = 7.8 (\( SD = 1.7 \))  
Majority Black/AA (53%) with 9 other racial identities included
Measures: SWTRS

Teacher-report screener for student wellbeing

Three school-specific wellbeing domains plus a total score:

- **Academic**—e.g., “Comfortable working independently,” “Engaged in learning”
- **Social**—e.g., “Treats classmates kindly,” “Listens to teachers”
- **Emotional**—e.g., “Seems happy in class,” “Smiles at school”

Teachers asked to rate frequency of each student behavior over the past month
Measures: SAEBRS

Teacher-report behavioral risk screener

Reverse code negatively worded items

Four domains of student functioning:
- **Academic**—e.g., “Difficulty working independently,” “Production of acceptable work”
- **Social**—e.g., “Impulsiveness,” “Cooperation with peers”
- **Emotional**—e.g., “Sadness,” “Adaptable to change”

Teachers asked to rate frequency of each student behavior over the past month
Outcome Measures

**Strengths and Difficulties Questionnaire—Teacher Version (SDQ)**

- Widely used 25 item measure of youth functioning
- Content related to **internalizing** and **externalizing** problems, and **prosocial** behavior
- Temporal frame altered from past six months to past month
  - CFA confirmed structure
  - Risk established for scores at or above a 16 for Total difficulties composite (SDQ-T)

**Time On-Task**

- “In the past month, what percent of time was the student on-task during class?”
- 11-point response scale: 0% = *Never* to 100% = *Always*
- Risk established for scores at or below 20th percentile: 60–70%
Outcome Measures

**Academic Achievement**
- “In the past month, how well has the student performed in English Language Arts/Math?”
- Five-point response scale: 1 = *Far below grade level* to 5 = *Far above grade level*
- 1 and 1 or 2 both used as separate indicators of risk

**Intervention Involvement**
- “Is the student currently receiving additional intervention in any of the following areas? [check all that apply]...(a)math, (b) ELA, (c) behavior/mental health”
- Any level of involvement with (a) academic or (b) behavioral intervention dummy coded 1
Data Analyses: Latent Factor Structure

Three measurement models were compared via Confirmatory Factor Analysis

1. **Unidimensional model**
   - One global wellbeing latent factor

2. **Correlated factors model**
   - Three correlated first-order latent factors
     - AWB, SWB, and EWB

3. **Bifactor model**
   - Four orthogonal first-order latent factors
     - Three domain-specific factors—AWB, SWB, and EWB
     - One domain-general wellbeing factor

**Model fit**: $\chi^2 p > .05$, $CFI \geq .90$, $RMSEA \leq .08$, $SRMR \leq .08$

**Internal reliability**: *Latent*—$H \geq .70$; *Observed*—Hierarchical $\omega \geq .70$
Data Analyses: Classification Accuracy

Cut-scores were established for **BOTH** the SWTRS and SAEBRS scales based on local data (rather than pulling from previously established cuts) to reduce bias in favor of the SWTRS.

**ROC curve analysis**

Conditional probability statistics:
- Area under the ROC curve (AUC)
- Sensitivity (SENS)
- Specificity (SPEC)
- Positive Predictive Value (PPV)
- Negative Predictive Value (NPV)

**SENS:** ≥ .90 = **optimal**, ≥ .80 = **acceptable**, and ≥ .70 = **borderline**

**SPEC:** ≥ .80 = **optimal**, ≥ .70 = **acceptable**, and ≥ .60 = **borderline**
Data Analyses: Classification Accuracy

Multistep cut-score search procedure

1. **Optimal** sensitivity and specificity
2. **Acceptable** sensitivity/optimal specificity
3. **Optimal** sensitivity/acceptable specificity
4. **Acceptable** sensitivity and specificity
5. **Borderline** sensitivity/acceptable specificity
6. **Acceptable** sensitivity/borderline specificity

Procedure terminates when a score meets criteria

If multiple scores meet criteria at same step, select score that minimizes discrepancy
Data Analyses: Comparative Concurrent and Predictive Validity

Risk on the SWTRS and SAEBRS scales at Time 1 predicting risk classification on all outcomes at Time 1 and Time 2

Generalized Linear Mixed Modeling (GLMM)

Hox model building procedure

Model comparison statistics: AIC, BIC, Likelihood ratio tests, ICC

Predictor comparisons: Odds Ratios
Data Analyses: Comparative Concurrent and Predictive Validity

For each outcome, four models tested...

1. \( y = TWB + (\text{Teacher Error}) + (\text{Rand. Error}) \)
2. \( y = TB + (\text{Teacher Error}) + (\text{Rand. Error}) \)
3. \( y = AWB + SWB + EWB + (\text{Teacher Error}) + (\text{Rand. Error}) \)
4. \( y = AB + SB + EB + (\text{Teacher Error}) + (\text{Rand. Error}) \)

**Like terms were compared across SWTRS and SAEBRS**
Q1: Structural Validity
SWTRS 18-Item Unidimensional Model

χ² (135) = 1095.93
CFI = .943
RMSEA[90% CI] = .139 [.132, .147]
SRMR = .089
SWTRS 18-Item Correlated Factors Model

χ² (132) = 692.76
CFI = .967
RMSEA[90% CI] = .107[.100, .115]
SRMR = .060
SWTRS 18-Item Bifactor Model

\[ \chi^2 (117) = 382.91 \]
\[ \text{CFI} = .984 \]
\[ \text{RMSEA}[90\% \text{ CI}] = .079 [.070, .087] \]
\[ \text{SRMR} = .039 \]
SWTRS 12-Item Bifactor Model

\[ \chi^2 (42) = 89.39 \]

CFI = .995

RMSEA[90% CI] = .055 [.039, .071]

SRMR = .025
$\omega = .98, .98$

$\omega_H = .87, .88$

$H = .95, .95$
Q2: Convergent & Discriminant Validity
### SWTRS Correlations ($r$)

<table>
<thead>
<tr>
<th></th>
<th>AWB</th>
<th>SWB</th>
<th>EWB</th>
<th>TWB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. SWTRS AWB</td>
<td><strong>.83</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. SWTRS SWB</td>
<td>.64</td>
<td><strong>.83</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. SWTRS EWB</td>
<td>.64</td>
<td>.63</td>
<td><strong>.72</strong></td>
<td></td>
</tr>
<tr>
<td>4. SWTRS TWB</td>
<td>.89</td>
<td>.87</td>
<td>.85</td>
<td><strong>.83</strong></td>
</tr>
<tr>
<td>5. SAEBRS AB</td>
<td>(.87)</td>
<td>.66</td>
<td>.57</td>
<td>.82</td>
</tr>
<tr>
<td>6. SAEBRS SB</td>
<td>.52</td>
<td>(.89)</td>
<td>.50</td>
<td>.73</td>
</tr>
<tr>
<td>7. SAEBRS EB</td>
<td>.60</td>
<td>.63</td>
<td>(.74)</td>
<td>.75</td>
</tr>
<tr>
<td>8. SAEBRS TB</td>
<td>.79</td>
<td>.86</td>
<td>.71</td>
<td>(.91)</td>
</tr>
<tr>
<td>9. TOT</td>
<td>(.78)</td>
<td>.62</td>
<td>.54</td>
<td>.76</td>
</tr>
<tr>
<td>10. Math Ach</td>
<td>.62</td>
<td>.28</td>
<td>.33</td>
<td>.49</td>
</tr>
<tr>
<td>11. Reading Ach</td>
<td>.58</td>
<td>.25</td>
<td>.25</td>
<td>.43</td>
</tr>
<tr>
<td>12. SDQ Tot</td>
<td>-.67</td>
<td>-.83</td>
<td>-.68</td>
<td>(-.83)</td>
</tr>
<tr>
<td>13. SDQ Int</td>
<td>-.48</td>
<td>-.49</td>
<td>(-.69)</td>
<td>-.62</td>
</tr>
<tr>
<td>14. SDQ Ext</td>
<td>-.62</td>
<td>(-.85)</td>
<td>-.49</td>
<td>-.76</td>
</tr>
<tr>
<td>15. SDQ PS</td>
<td>.53</td>
<td>(.71)</td>
<td>.59</td>
<td>.69</td>
</tr>
</tbody>
</table>

**Note:** All $p < .05$ after Holm-Bonferroni correction.
Q3: SDQ-T Risk Detection
SWTRS TWB ROC Curve
Area Under the ROC Curve (AUC)

- Total: SWTRS 0.95, SAEBRS 0.92
- Social: SWTRS 0.91, SAEBRS 0.92
- Emotional: SWTRS 0.87, SAEBRS 0.86
- Academic: SWTRS 0.87, SAEBRS 0.81

SWTRS and SAEBRS
AUC
<table>
<thead>
<tr>
<th></th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>0.9</td>
<td>0.82</td>
</tr>
<tr>
<td>Social</td>
<td>0.86</td>
<td>0.81</td>
</tr>
<tr>
<td>Emotional</td>
<td>0.86</td>
<td>0.84</td>
</tr>
<tr>
<td>Academic</td>
<td>0.73</td>
<td>0.67</td>
</tr>
</tbody>
</table>

**SENS:** ≥ .90 = optimal, ≥ .80 = acceptable, and ≥ .70 = borderline

**SPEC:** ≥ .80 = optimal, ≥ .70 = acceptable, and ≥ .60 = borderline
Q4: Criterion Validity
SWTRS and SAEBRS T1 Risk → T1 SDQ-T Risk

**Odds Ratio**

<table>
<thead>
<tr>
<th>Domain</th>
<th>SAEBRS</th>
<th>SWTRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total†</td>
<td>56.26</td>
<td>19.11</td>
</tr>
<tr>
<td>Academic</td>
<td>5.58</td>
<td>4.31</td>
</tr>
<tr>
<td>Social</td>
<td>12.68</td>
<td>8.85</td>
</tr>
<tr>
<td>Emotional</td>
<td>5.53</td>
<td>4.81</td>
</tr>
</tbody>
</table>

* p < .05
** p < .01
*** p < .001
† Significantly different
SWTRS and SAEBRS T1 Risk  →  T2 SDQ-T Risk

Odds Ratio

<table>
<thead>
<tr>
<th>Category</th>
<th>Total</th>
<th>Academic</th>
<th>Social</th>
<th>Emotional</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAEBRS</td>
<td>31.5</td>
<td>9.87</td>
<td>12.81</td>
<td>2.44</td>
</tr>
<tr>
<td>SWTRS</td>
<td>26.05</td>
<td>5.81</td>
<td>12.43</td>
<td>2.86</td>
</tr>
</tbody>
</table>

*  p < .05
** p < .01
*** p < .001
†Significantly different
TIME ON-TASK (TOT)
SWTRS and SAEBRS T1 Risk → T1 TOT Risk

<table>
<thead>
<tr>
<th></th>
<th>Odds Ratio</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>18.92 ***</td>
<td>* &lt; .01</td>
</tr>
<tr>
<td></td>
<td>13.87 ***</td>
<td></td>
</tr>
<tr>
<td>Academic</td>
<td>22.2 ***</td>
<td>** &lt; .01</td>
</tr>
<tr>
<td></td>
<td>30.27 ***</td>
<td>*** &lt; .001</td>
</tr>
<tr>
<td>Social</td>
<td>3.6 **</td>
<td>&lt; .05</td>
</tr>
<tr>
<td></td>
<td>1.88</td>
<td></td>
</tr>
<tr>
<td>Emotional</td>
<td>3.29 ***</td>
<td>** &lt; .01</td>
</tr>
<tr>
<td></td>
<td>2.29</td>
<td>*</td>
</tr>
</tbody>
</table>

*Significantly different
SWTRS and SAEBRS T1 Risk → T2 TOT Risk
READING ACHIEVEMENT
SWTRS and SAEBRS T1 Risk → T1 Below Grade-Level Reading Risk

- **Total**: 5.00 (SWTRS), 4.14 (SAEBRS)
- **Academic**: 16.12 (SWTRS), 21.33 (SAEBRS)
- **Social**: 0.56 (SWTRS), 0.84 (SAEBRS)
- **Emotional**: 1.55 (SWTRS), 0.82 (SAEBRS)

*Significance levels:
* \( p < .05 \)
** \( p < .01 \)
*** \( p < .001 \)
†Significantly different
SWTRS and SAEBRS T1 Risk $\rightarrow$ T2 Below Grade-Level Reading Risk

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Academic</th>
<th>Social</th>
<th>Emotional</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWTRS</td>
<td>4.71</td>
<td>19.49</td>
<td>0.54</td>
<td>1.43</td>
</tr>
<tr>
<td>SAEBRS</td>
<td>6.30</td>
<td>16.78</td>
<td>0.76</td>
<td>1.39</td>
</tr>
</tbody>
</table>

* $p < .05$  
** $p < .01$  
*** $p < .001$  
†Significantly different
MATH ACHIEVEMENT
**SWTRS and SAEBRS T1 Risk → T1 Below Grade-Level Math Risk**

<table>
<thead>
<tr>
<th></th>
<th>Odds Ratio</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAEBRS</td>
<td>5.21</td>
<td></td>
</tr>
<tr>
<td>SWTRS</td>
<td>5.81</td>
<td></td>
</tr>
<tr>
<td><strong>Academic</strong></td>
<td>14.30</td>
<td><strong>p &lt; .01</strong></td>
</tr>
<tr>
<td>SAEBRS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWTRS</td>
<td></td>
<td>*<strong>p &lt; .001</strong></td>
</tr>
<tr>
<td><strong>Social</strong></td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>SAEBRS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWTRS</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td><strong>Emotional</strong></td>
<td>2.23</td>
<td><em>p &lt; .05</em>*</td>
</tr>
<tr>
<td>SAEBRS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWTRS</td>
<td>1.17</td>
<td></td>
</tr>
</tbody>
</table>

*Significantly different

* * *

Significance levels:

* $p < .05$
** $p < .01$
*** $p < .001$

Diagram:

- Bar chart showing odds ratios for Total, Academic, Social, and Emotional domains.
- Odds ratios for SWTRS and SAEBRS.
- Significance levels indicated with stars.

---

**Note:** The image contains a bar chart illustrating the odds ratios for SWTRS and SAEBRS in different domains, with significance levels indicated by stars.
<table>
<thead>
<tr>
<th></th>
<th>SWTRS</th>
<th>SAEBRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>4.62</td>
<td>8.85</td>
</tr>
<tr>
<td>Academic</td>
<td>9.39</td>
<td>12.30</td>
</tr>
<tr>
<td>Social</td>
<td>0.68</td>
<td>0.79</td>
</tr>
<tr>
<td>Emotional</td>
<td>2.39</td>
<td>2.51</td>
</tr>
</tbody>
</table>

**Odds Ratio**

- * $p < .05$
- ** $p < .01$
- *** $p < .001$

†Significantly different
BEHAVIORAL INTERVENTION
SWTRS and SAEBRS T1 Risk → T1 Behavioral Intervention

* p < .05
** p < .01
*** p < .001
†Significantly different
SWTRS and SAEBRS T1 Risk → T2 Behavioral Intervention

### Odds Ratio

<table>
<thead>
<tr>
<th></th>
<th>SWTRS</th>
<th>SAEBRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>21.76</td>
<td>14.59</td>
</tr>
<tr>
<td>Academic</td>
<td>1.14</td>
<td>3.86</td>
</tr>
<tr>
<td>Social</td>
<td>8.67</td>
<td>8.85</td>
</tr>
<tr>
<td>Emotional</td>
<td>3.82</td>
<td>4.81</td>
</tr>
</tbody>
</table>

* * p < .05
** ** p < .01
*** *** p < .001
† Significantly different
ACADEMIC INTERVENTION
### SWTRS and SAEBRS T1 Risk → T1 Academic Intervention

<table>
<thead>
<tr>
<th>Category</th>
<th>SWTRS</th>
<th>SAEBRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>3.82</td>
<td>7.10</td>
</tr>
<tr>
<td>Academic</td>
<td>9.78</td>
<td></td>
</tr>
<tr>
<td>Social</td>
<td>0.54</td>
<td>0.88</td>
</tr>
<tr>
<td>Emotional</td>
<td>2.01</td>
<td>1.13</td>
</tr>
</tbody>
</table>

#### Notes:
- * $p < .05$
- ** $p < .01$
- *** $p < .001$
- †Significantly different
SWTRS and SAEBRS T1 Risk → T2 Academic Intervention

Odd Ratio

**SAEBRS**

**SWTRS**

* * * p < .05
** ** p < .01
*** *** p < .001

†Significantly different
Summary & Conclusions
Q1. Structural Validity

➢ 12-Item Bifactor Structure preferred

➢ Four derived scores
  ➢ TWB, AWB, SWB, EWB

➢ Adequate internal consistency independently

➢ TWB strongest latent and observed reliability
Summary

Q2. Convergent & Discriminant Validity

➢ TWB strongest with TB and SDQ-T

➢ AWB strongest with AB and academic metrics (TOT, Math/Reading achievement)

➢ SWB strongest with SB and SDQ Ext

➢ EWB strongest with EB and SDQ Int
Q3. SDQ-T Risk Identification

➢ Adequate cut-points for all SWTRS scores
➢ TWB strongest
➢ AWB weakest. Overidentified (45%)
Summary

Q4. Criterion Validity

➢ TWB most consistently strong among SWTRS scores
➢ TWB > TB for T1 Academic Intervention but < for T1 SDQ-T
➢ Among SWTRS subscales, at least 1 sig. term per model
➢ SWTRS subscale block > SAEBRS for T1/T2 BGL Math and T1/T2 Academic Intervention
Implications and Recommendations

➢ Evidence for interpretation and use for both SWTRS and SAEBRS
➢ Concurrent and short-run future

➢ Instrument selection should be driven by goals of screening...

➢ Broad student risk? → SAEBRS
➢ ≥ risk detection than SWTRS

➢ Target positive functioning specifically? → SWTRS
➢ SAEBRS does not differentiate

➢ High-stakes decision making → Total scores (TWB, TB)

➢ Subscale use better for general domain target for additional assessment
## Limitations

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overfitting SWTRS</td>
<td>• Replication samples / “Training” and “Test” data sets needed</td>
</tr>
<tr>
<td>Staggered T1/T2 Data Collection</td>
<td>• Only reflects 1(^{st}) and 2(^{nd}) instances of screening</td>
</tr>
</tbody>
</table>
| Monomethod Bias                  | • Behavioral and Academic Intervention outcomes included to counteract other more subjective ratings  
                                 | • …some issues still noted                                                                                                               |
| SDQ Proxy for “Gold Standard” of Risk | • Not much longer than SAEBRS  
                                 | • A couple reverse coded items as well                                                                                                  |
Future Directions

- Compare combined and co-use DFMH approach
- Potential for progress monitoring
- Treatment utility
- Control for temporal effects
- More objective criterion measures


