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
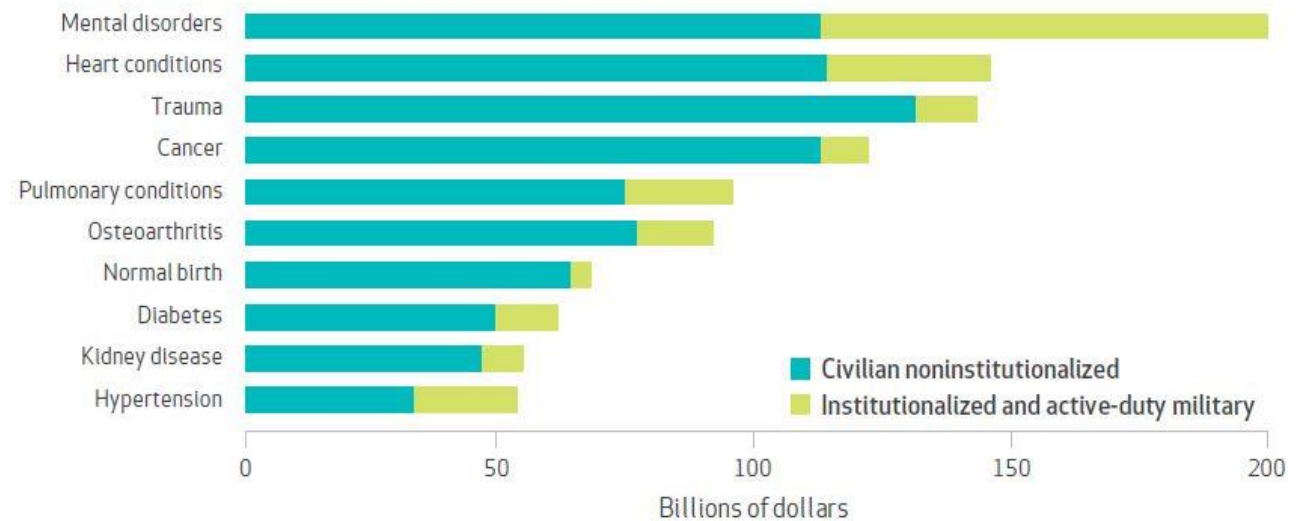
1 in 5 youth with MH disorder[1]

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]

Ten medical conditions with the highest estimated spending in 2013

Mental disorders
Heart conditions
Trauma
Cancer
Pulmonary conditions
Osteoarthritis
Normal birth
Diabetes
Kidney disease
Hypertension

■ Civilian noninstitutionalized
■ Institutionalized and active-duty military

0   50   100   150   200
Billions of dollars

SOURCE Author's analysis of study data. NOTES Institutionalized populations include nursing home residents, long-term patients in psychiatric hospitals, and prisoners. Trauma is fractures and wounds. Pulmonary conditions include chronic obstructive pulmonary disease, asthma, and other pulmonary diseases.

School-based Mental Health Support

A quarter of youth estimated to be identified and provided community services[8]

School becomes "de facto" MH care provider[9]
- 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[10]

Calls for improving school-based mental health care
- President's New Freedom Commission on Mental Health[11]

Improved intervention requires improved identification
- Traditional referral method common but problematic[12]

Screening for Risk

Derived from medical context
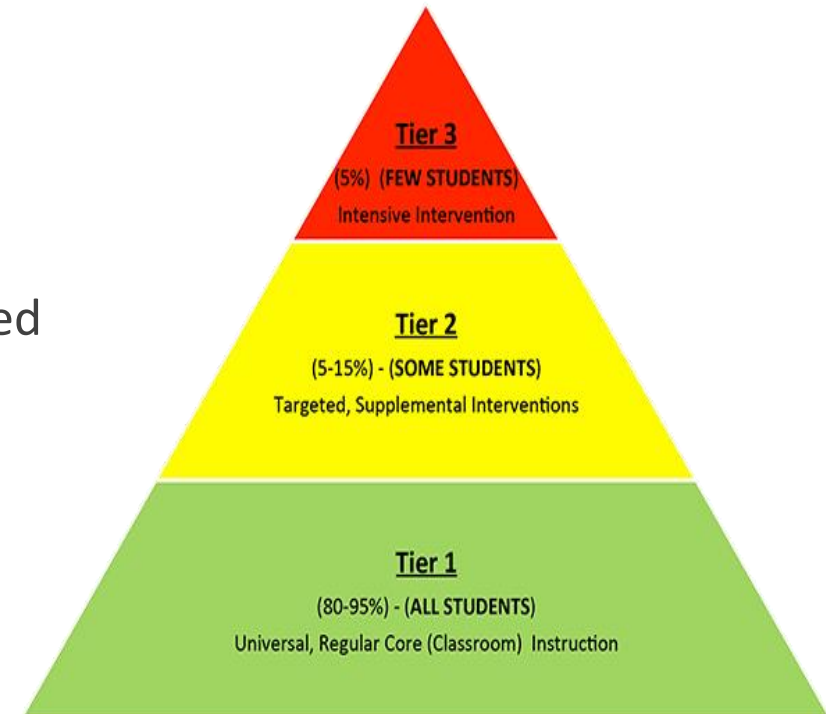
World Health Organization[13]:

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[14]

Advantages[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[17]

Tier 3
(5%) (FEW STUDENTS)
Intensive Intervention

Tier 2
(5-15%) - (SOME STUDENTS)
Targeted, Supplemental Interventions

Tier 1
(80-95%) - (ALL STUDENTS)
Universal, Regular Core (Classroom) Instruction

Standards for Universal Mental Health Screening in School[18]

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[19]

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)[20]
◦ Suldo et al. (2008; 2011)[21, 22]
◦ Kim, Furlong, Dowdy, & Felix (2014)[23]
◦ Dowdy et al. (2014)[24]

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[29, 31, 32]

Developed based on theories linking behavior to school success

Four scores: **Academic** (6 items), **Social** (6 items), **Emotional** (7 items), and **Total** (19 items) behavior

Both adaptive and maladaptive behaviors
◦ Based on dual-factor evidence
◦ Reverse score maladaptive behaviors

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[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: **Academic** (6 items), **Social** (6 items), and **Emotional** (5 items) wellbeing

- Strong factor loadings (> .50)
- Communalities (> .60)
- Internal consistency ($\omega \geq .89$)

SWTRS

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

The Current Study

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…

The Current Study

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[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

	Time 1	Time 2
N	371	332
Median Students per Class	24	24
Mean Students per Class	20.6 (*SD* = 7.8)	22.1 (*SD* = 5.1)
Active IEP	4%	4%
504 Plan	4%	5%
Non-English First Language	16%	16%

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)[35]

- 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: χ2 p > .05 , CFI ≥ .90, RMSEA ≤ .08, SRMR ≤ .08

Internal reliability: *Latent*—*H* ≥ .70; *Observed*—Hierarchical ω ≥ .70
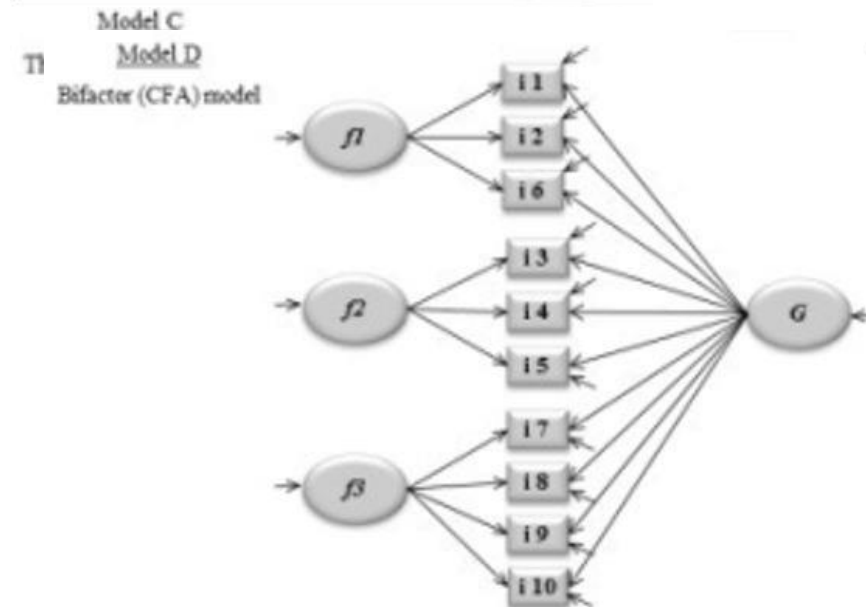
Data Analyses: Classification Accuracy

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
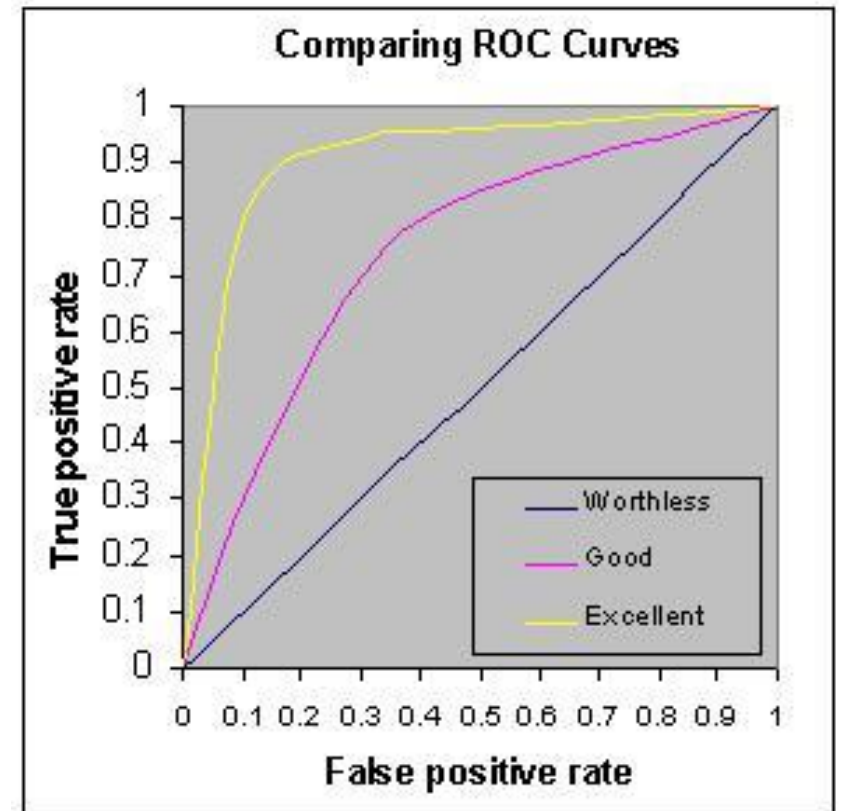
Data Analyses: Classification Accuracy

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 + (Teacher Error) + (Rand. Error)

2. y = TB + (Teacher Error) + (Rand. Error)

3. y = AWB + SWB + EWB + (Teacher Error) + (Rand. Error)

4. y = AB + SB + EB + (Teacher Error) + (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
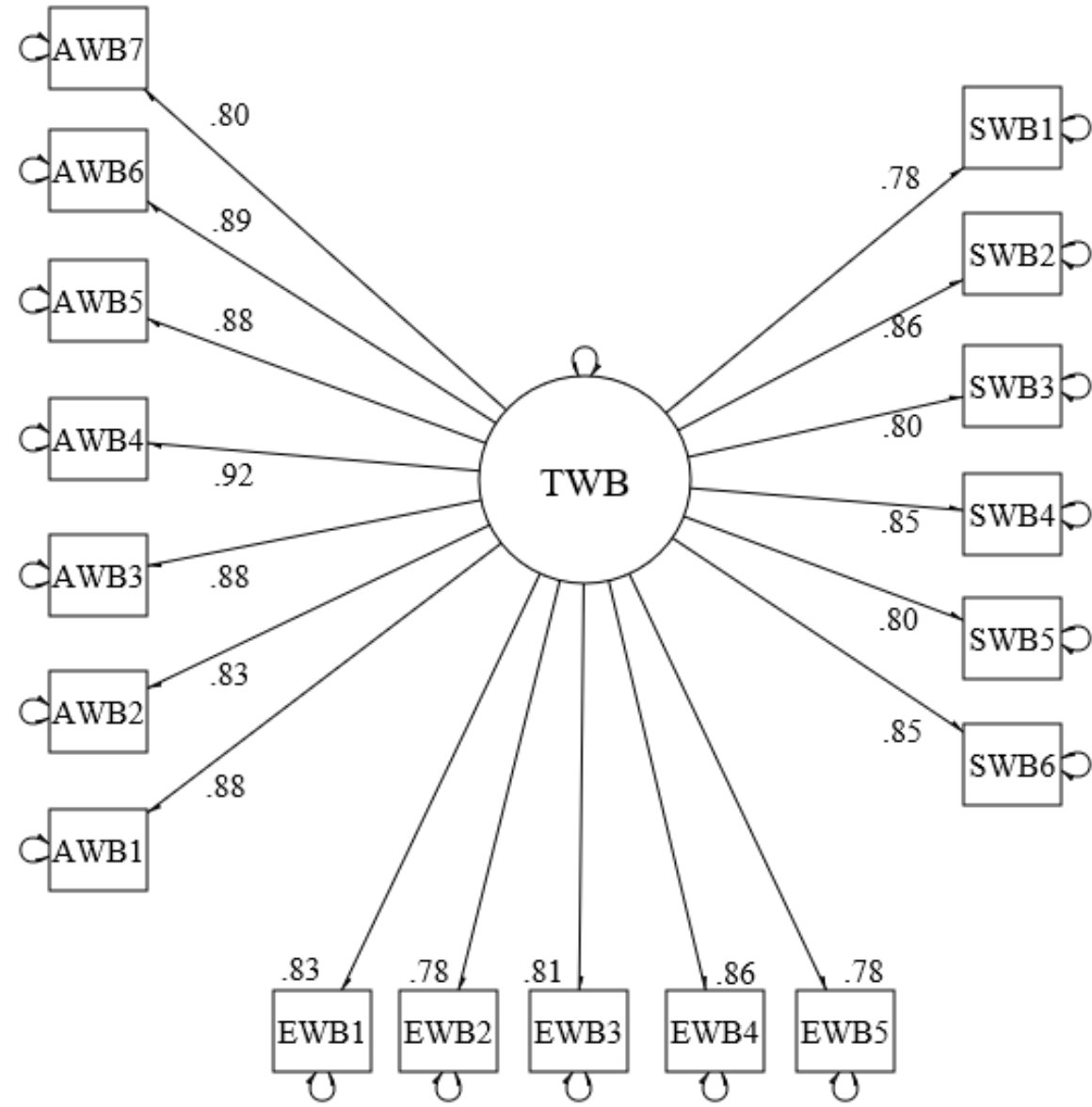
SWTRS 18-Item Correlated Factors Model

χ2 (132) = 692.76

CFI = .967

RMSEA[90% CI] = .107[.100, .115]

SRMR = .060
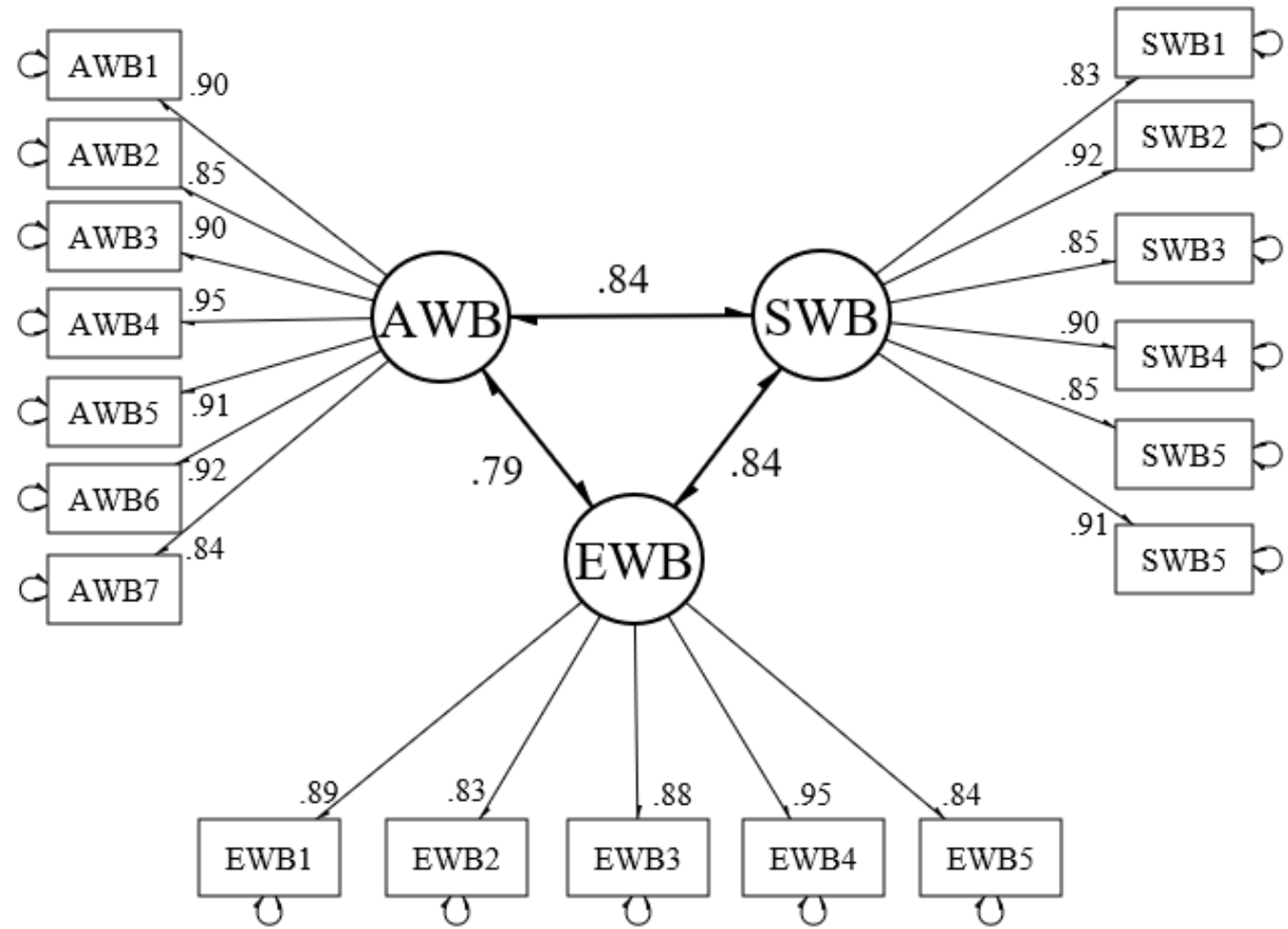
SWTRS 18-Item Bifactor Model

χ2 (117) = 382.91

CFI = .984

RMSEA[90% CI] = .079[.070, .087]

SRMR = .039
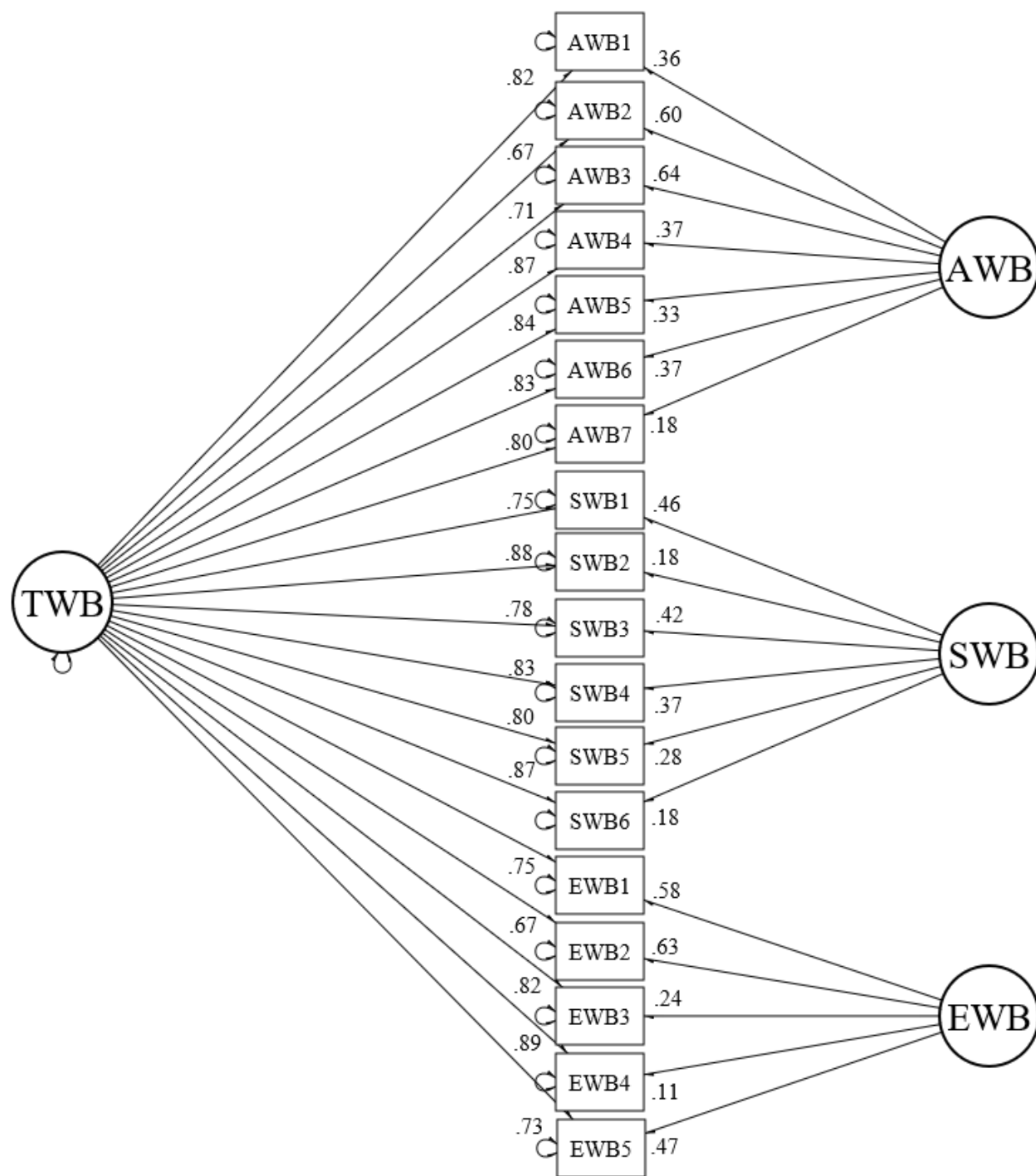
SWTRS 12-Item Bifactor Model

$\chi 2$ (42) = 89.39

CFI = .995

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

SRMR = .025
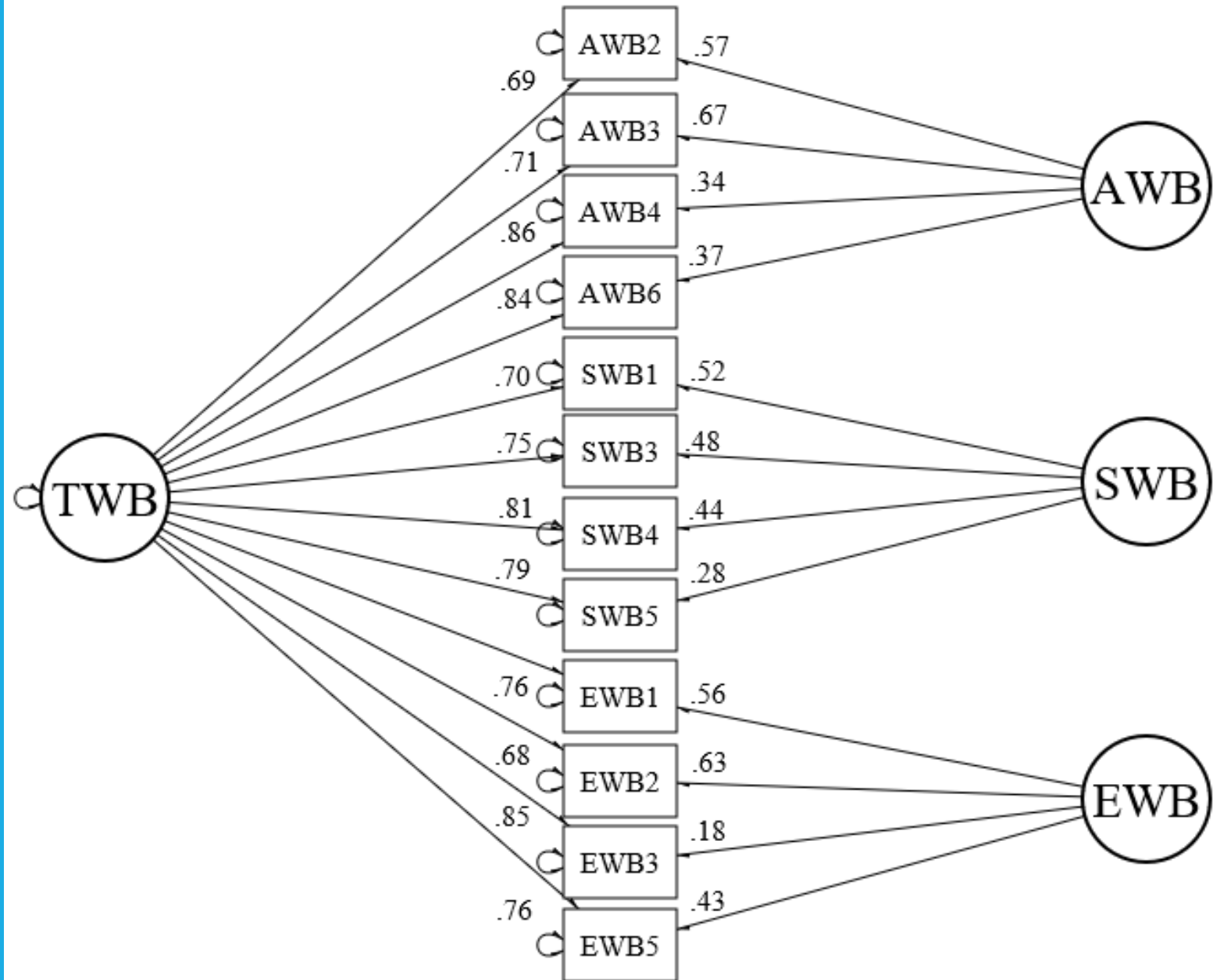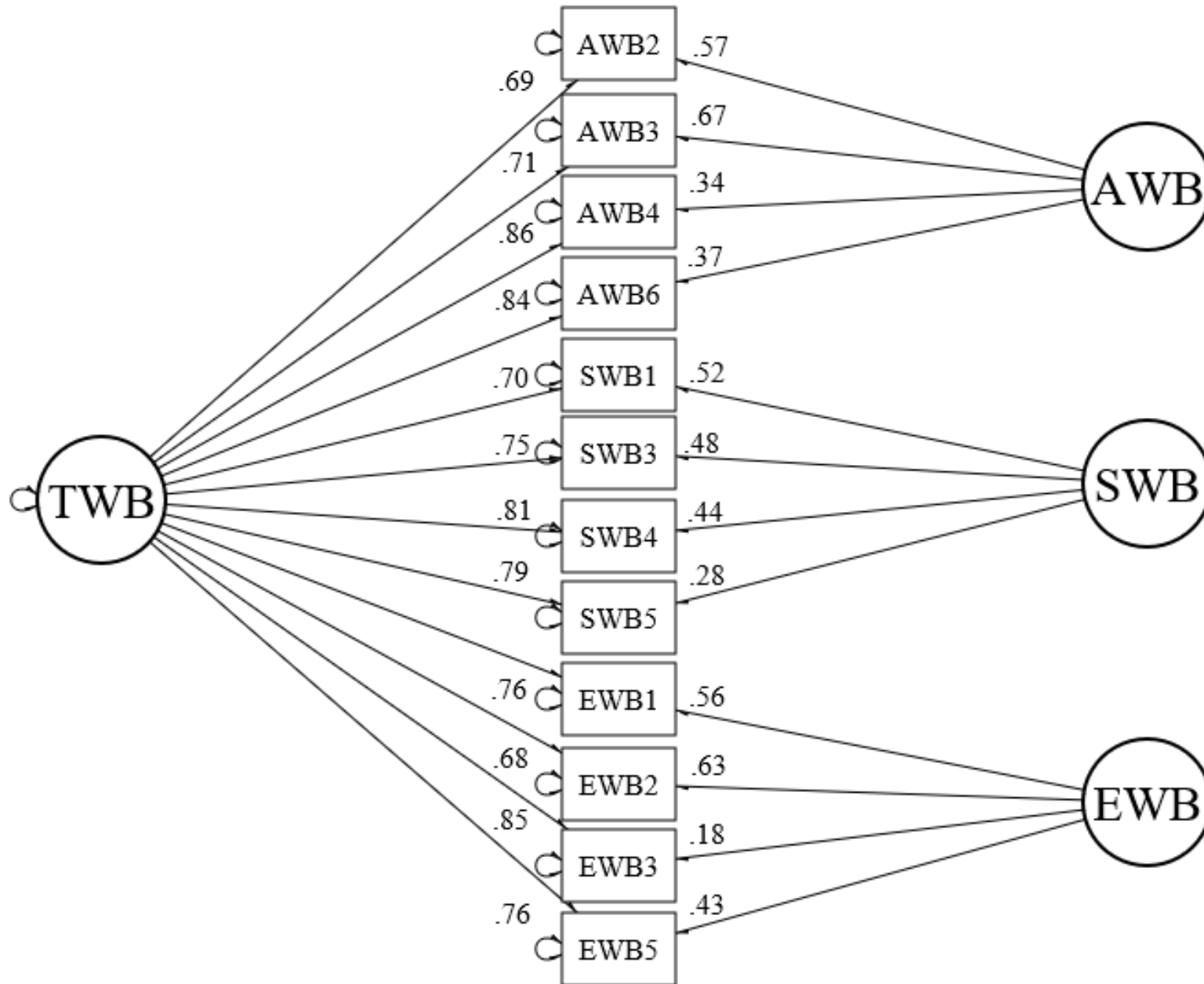
Q2: Convergent & Discriminant Validity

SWTRS Correlations (*r*)

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

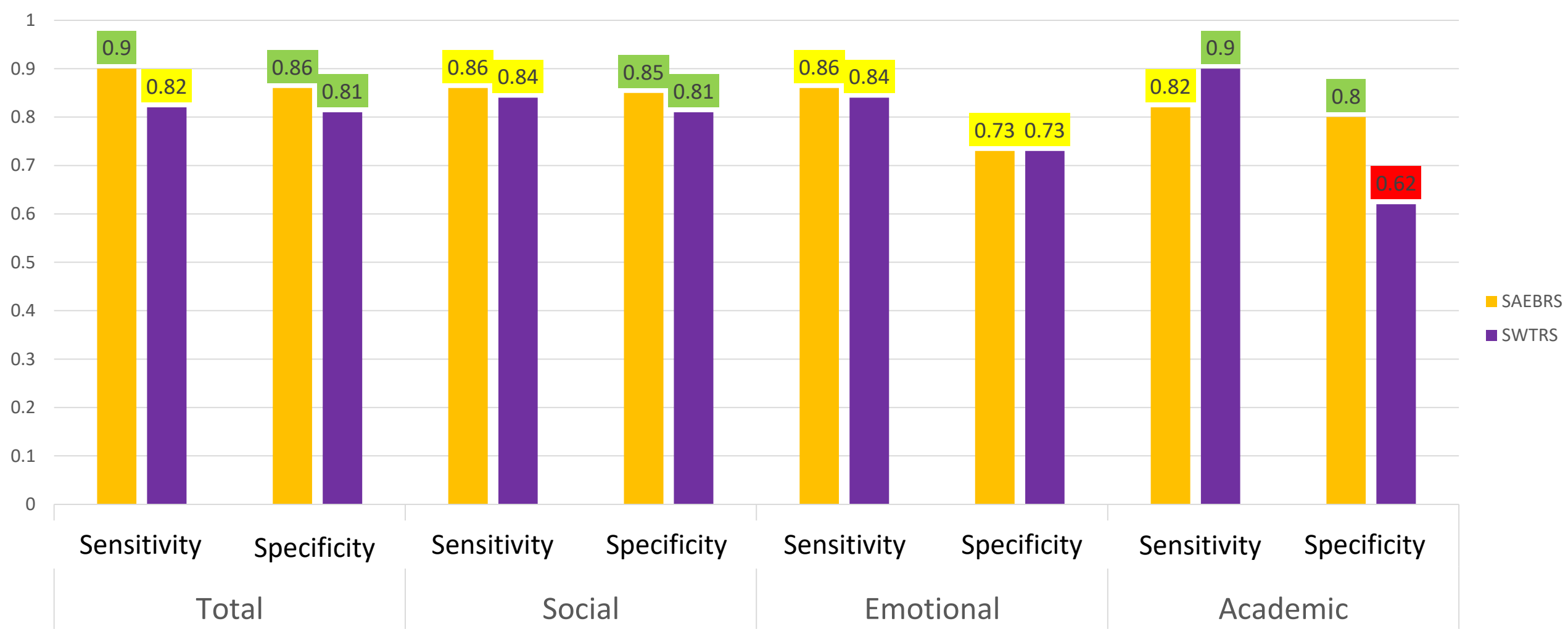*Note*: All *p* < .05 after Holm-Bonferroni correction

Q3:
SDQ-T Risk Detection
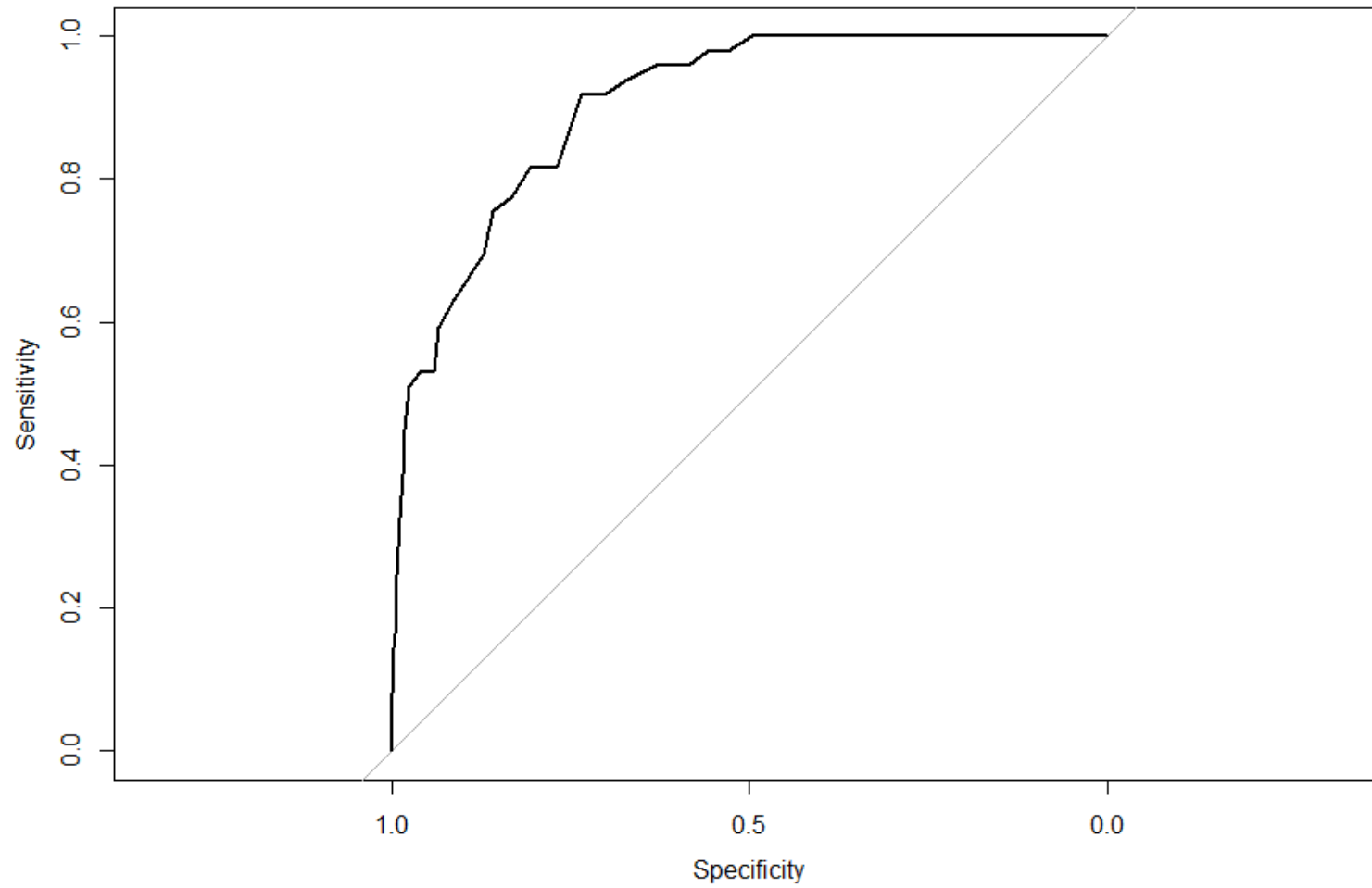
SWTRS TWB ROC Curve
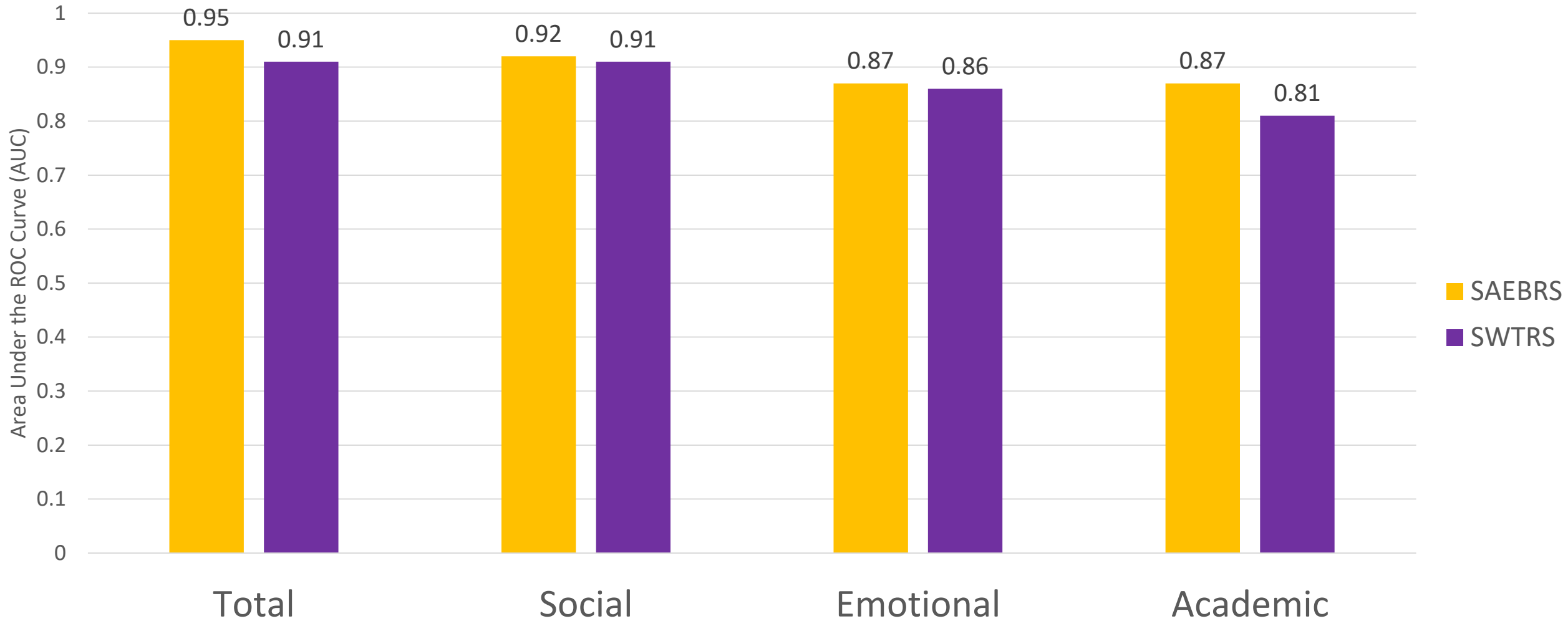
SWTRS and SAEBRS
AUC

SWTRS and SAEBRS Sensitivity and Specificity

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

SPEC: ≥ .80 = optimal, ≥ .70 = acceptable, and ≥ .60 = borderline
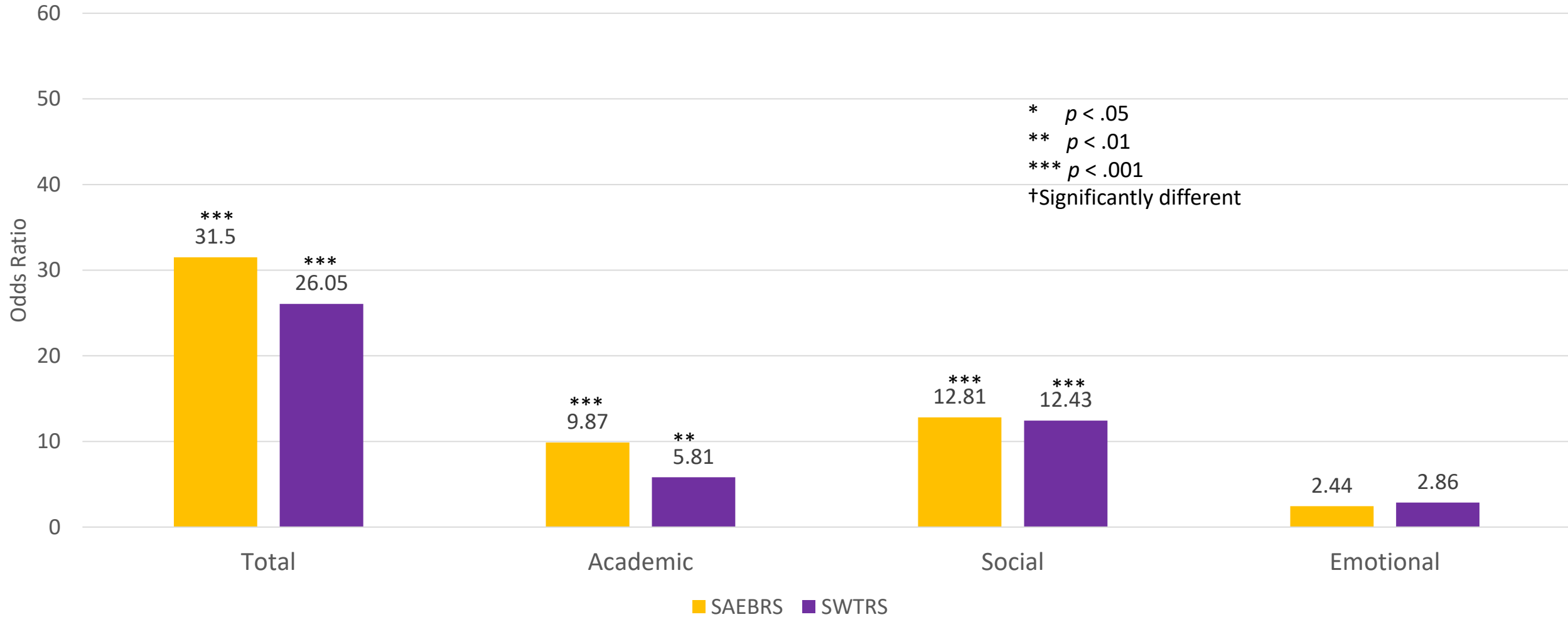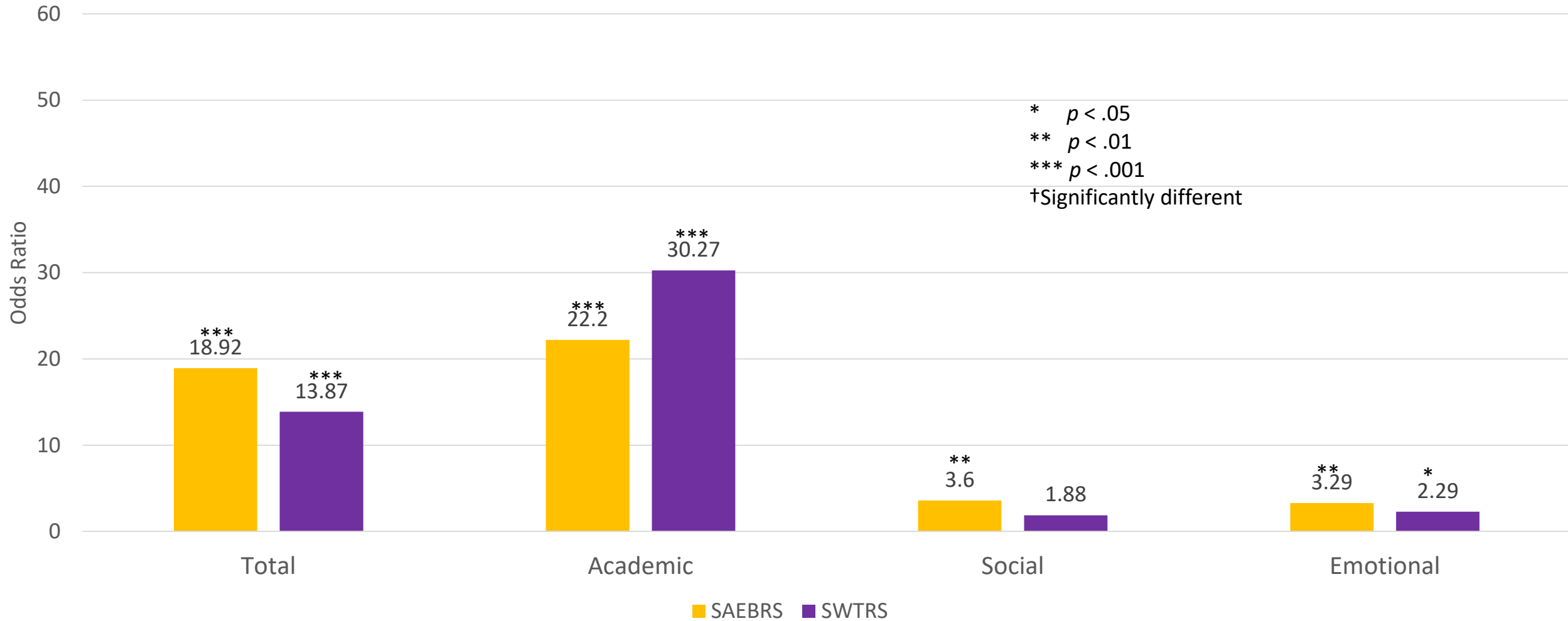
Q4:
Criterion Validity

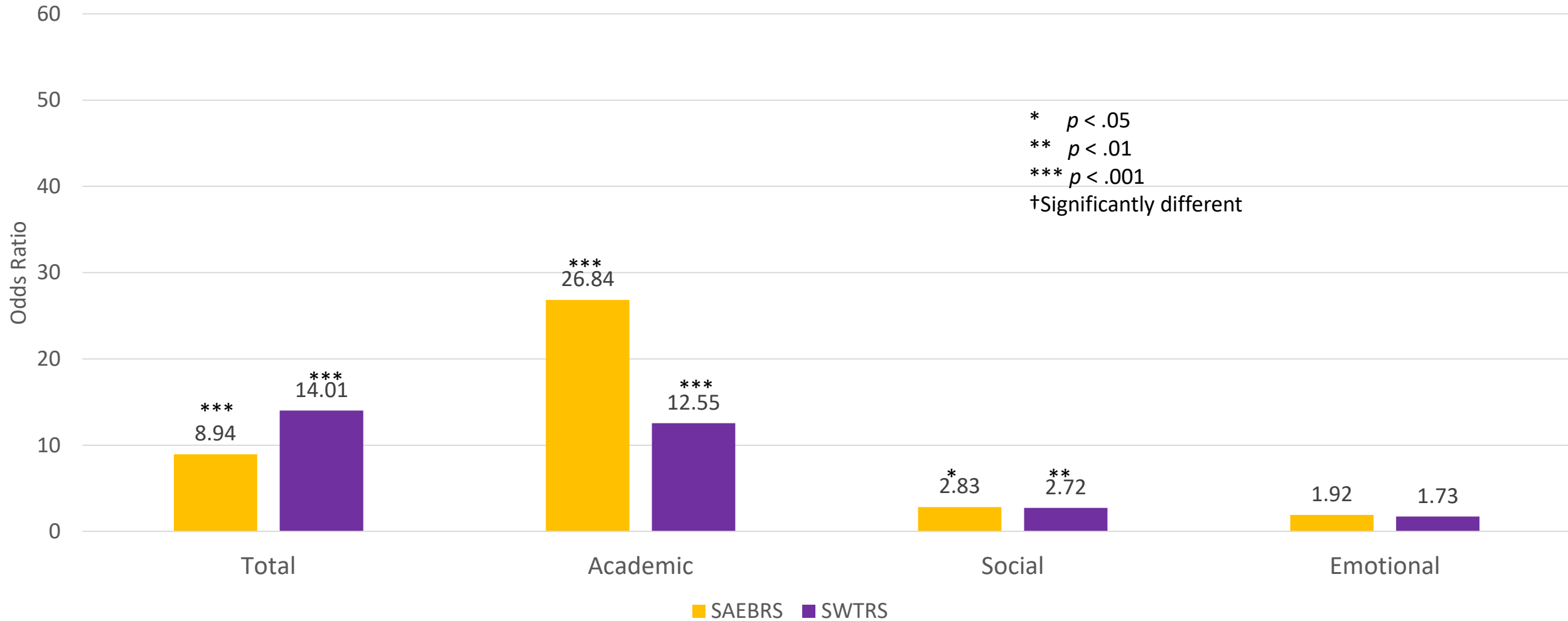# SDQ-T
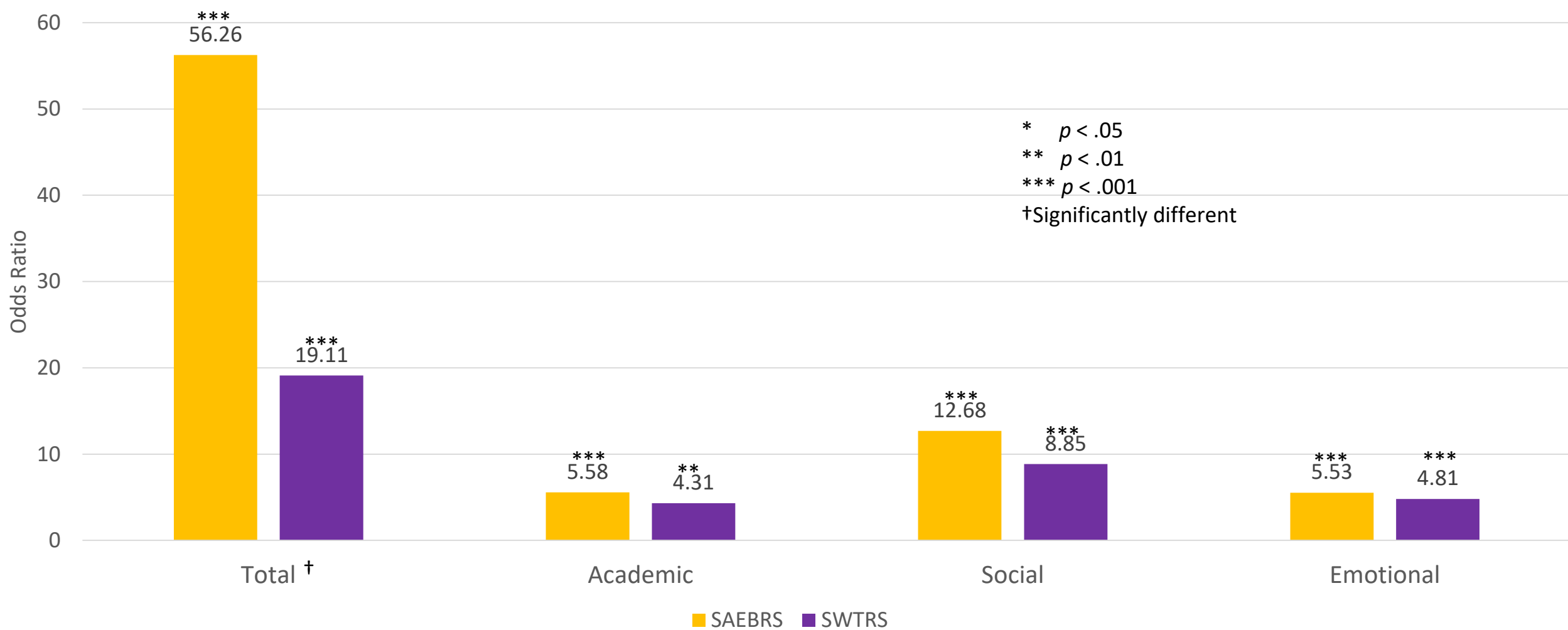
SWTRS and SAEBRS T1 Risk → T1 SDQ-T Risk

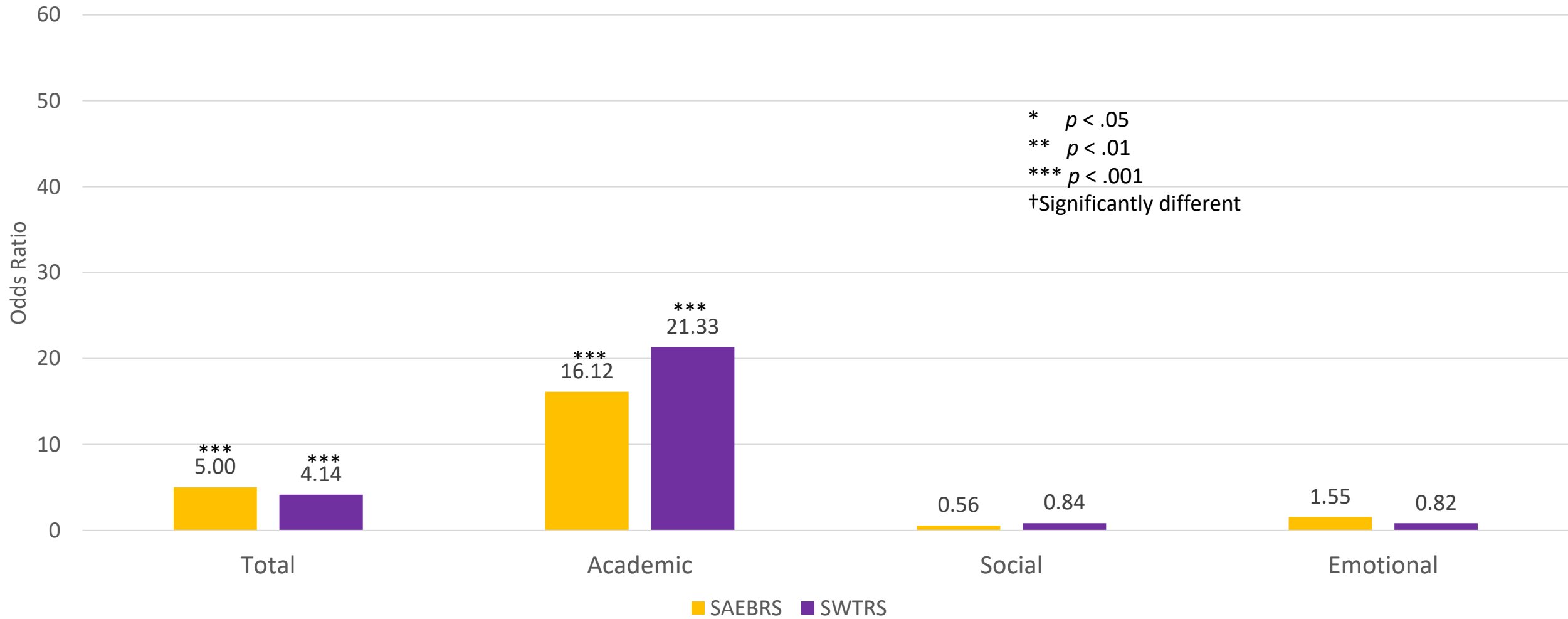SWTRS and SAEBRS T1 Risk → T2 SDQ-T Risk

TIME ON-TASK (TOT)
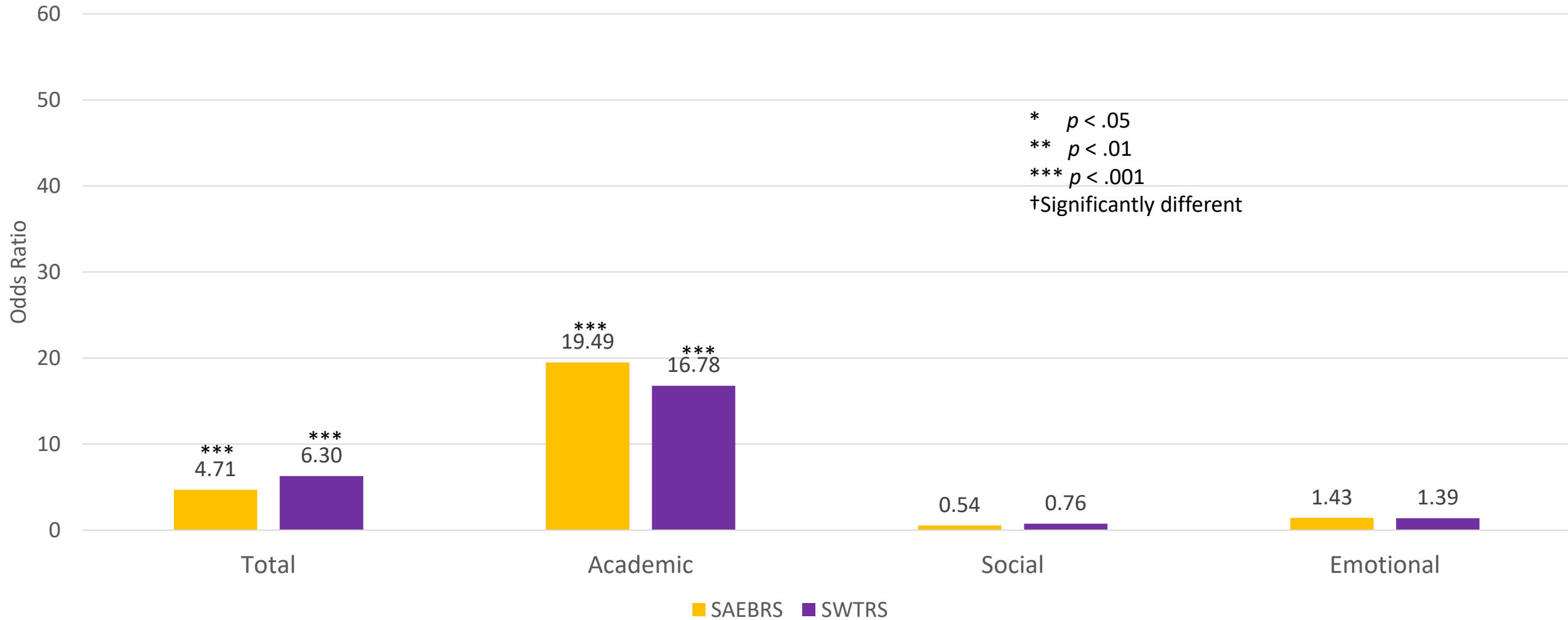
SWTRS and SAEBRS T1 Risk → T1 TOT Risk

SWTRS and SAEBRS T1 Risk → T2 TOT Risk

READING ACHIEVEMENT

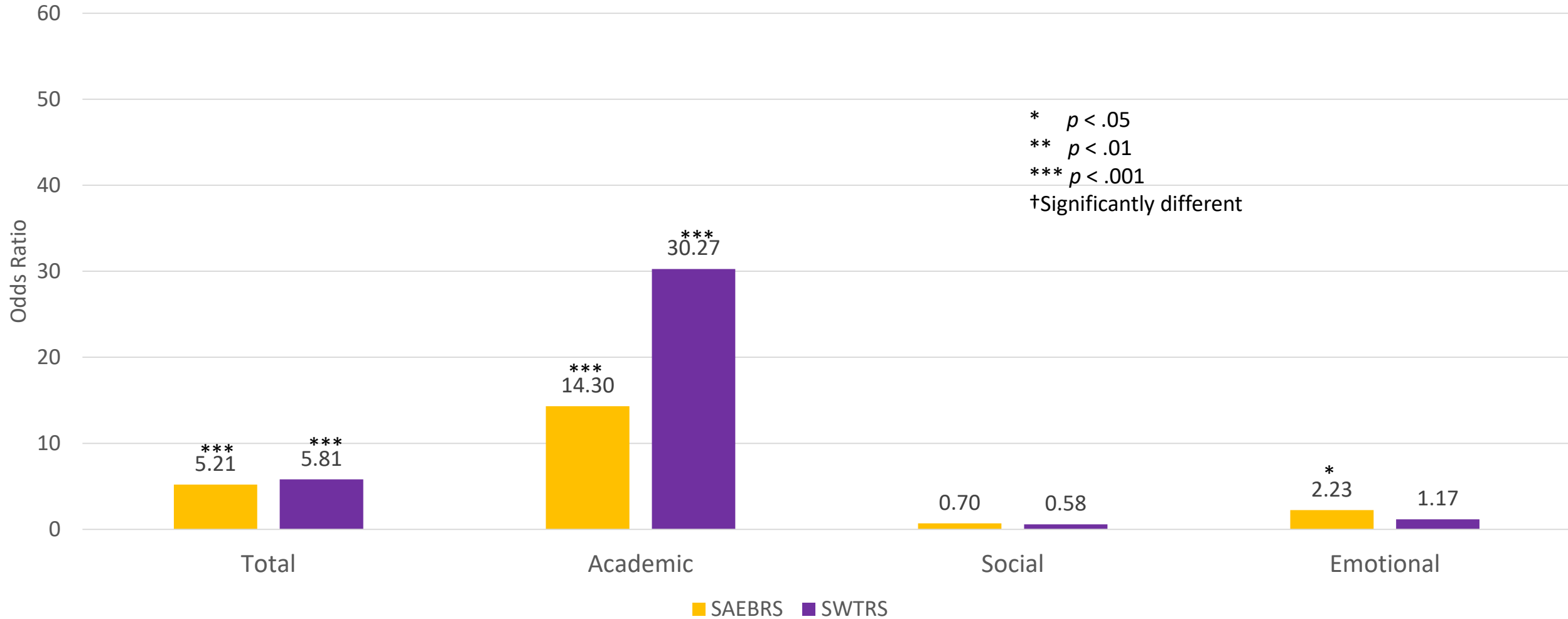SWTRS and SAEBRS T1 Risk → T1 Below Grade-Level Reading Risk

SWTRS and SAEBRS T1 Risk → T2 Below Grade-Level Reading Risk

MATH ACHIEVEMENT

SWTRS and SAEBRS T1 Risk → T1 Below Grade-Level Math Risk
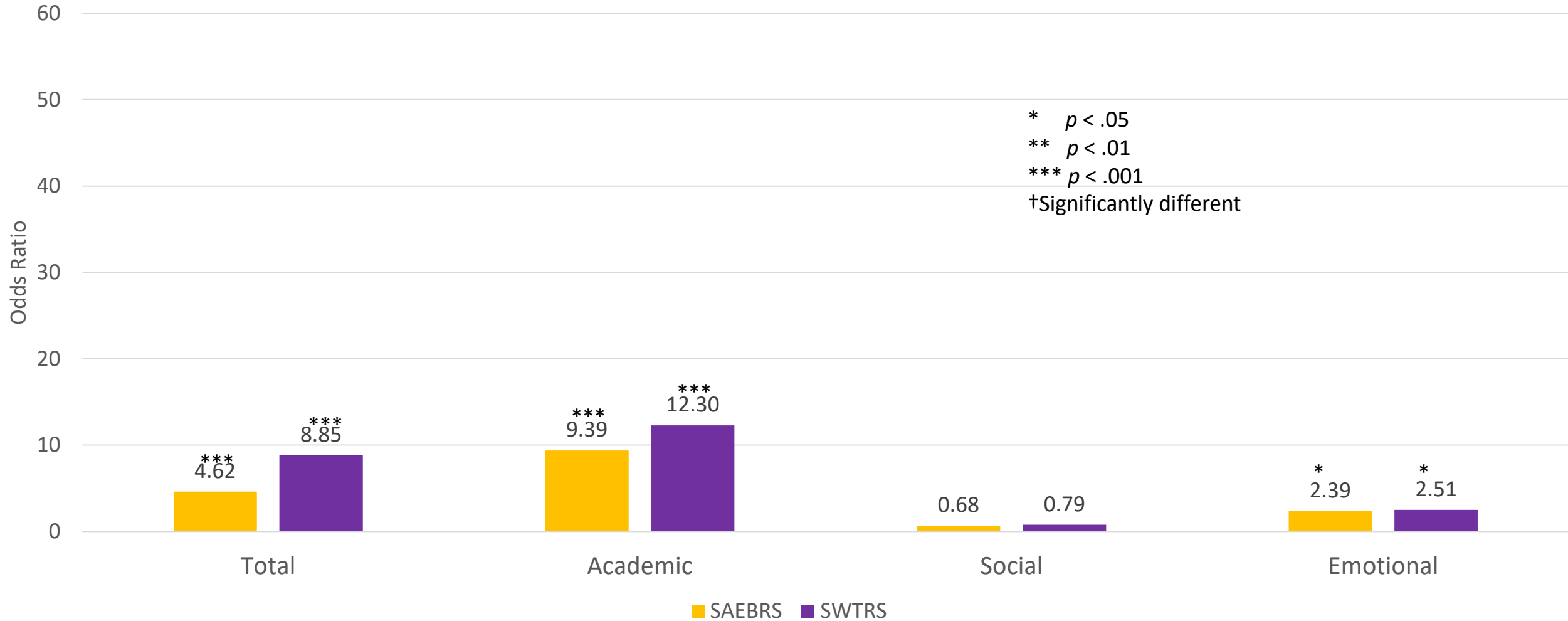
SWTRS and SAEBRS T1 Risk → T2 Below Grade-Level Math Risk

BEHAVIORAL INTERVENTION
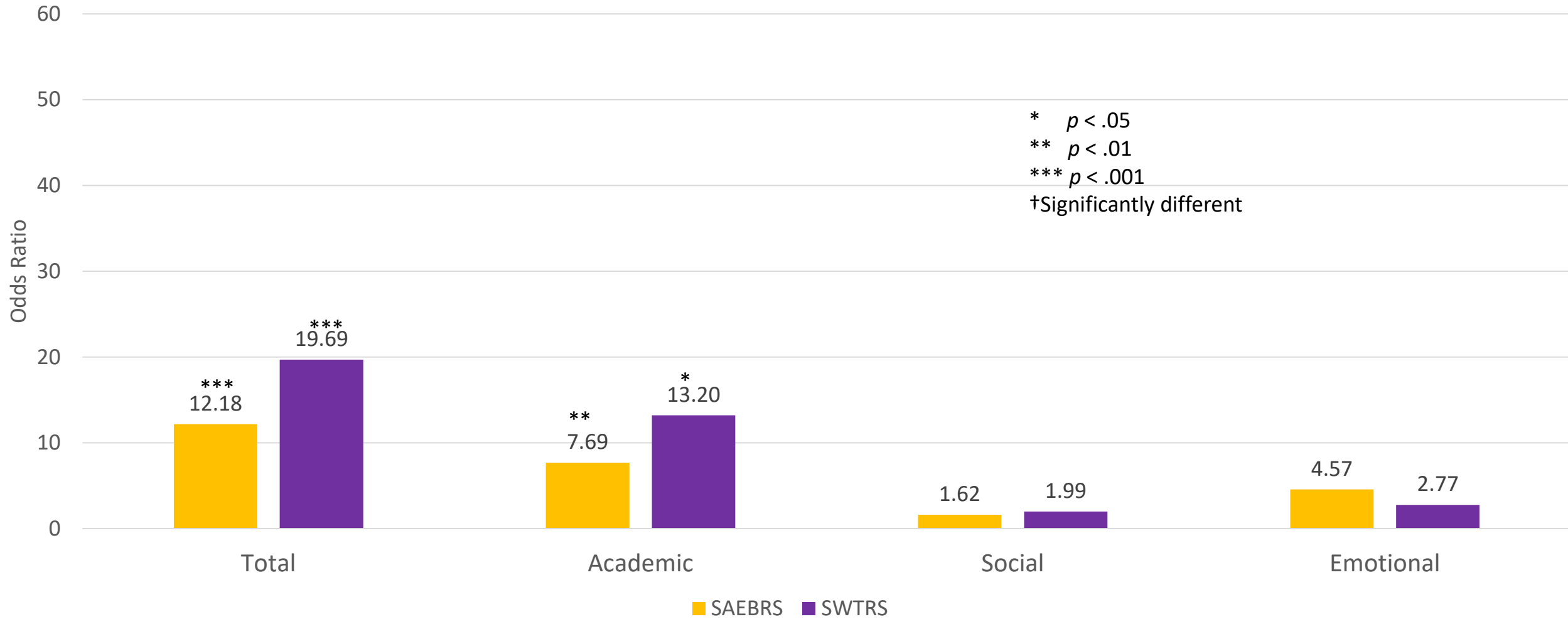
SWTRS and SAEBRS T1 Risk → T1 Behavioral Intervention
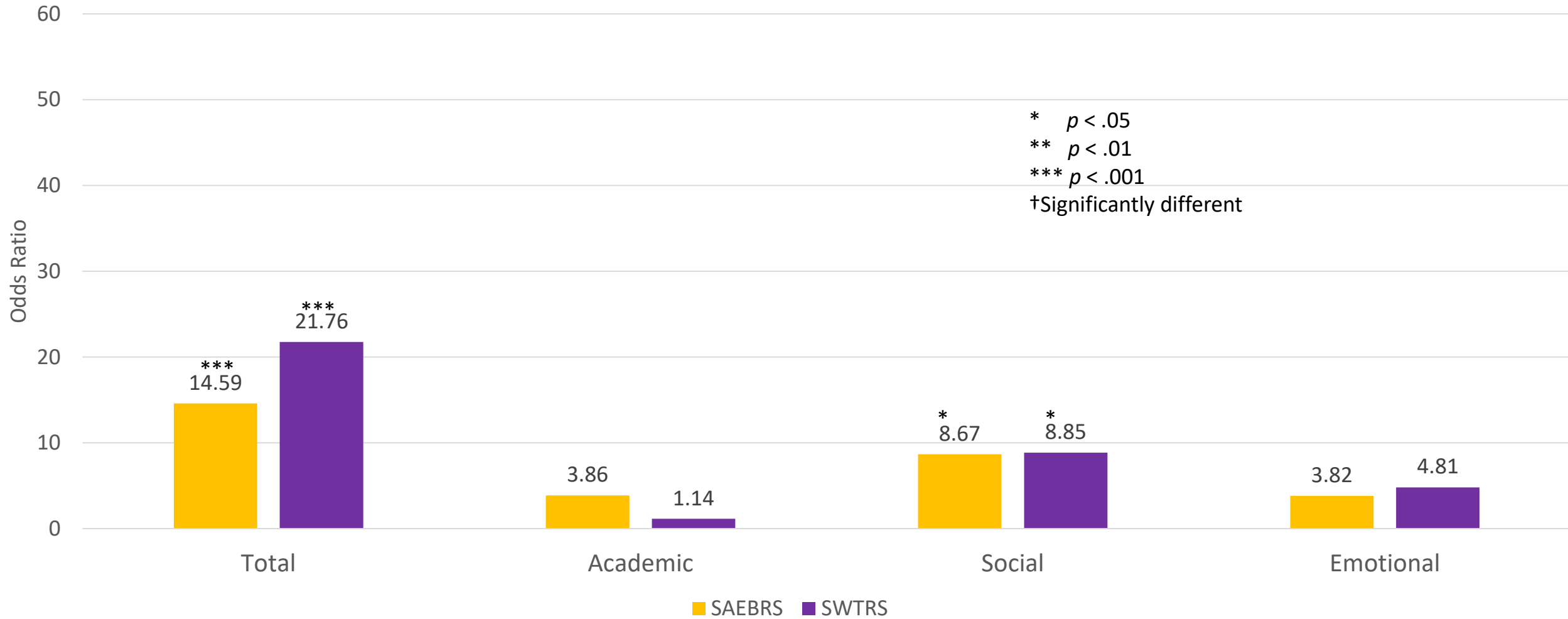
SWTRS and SAEBRS T1 Risk → T2 Behavioral Intervention

ACADEMIC INTERVENTION
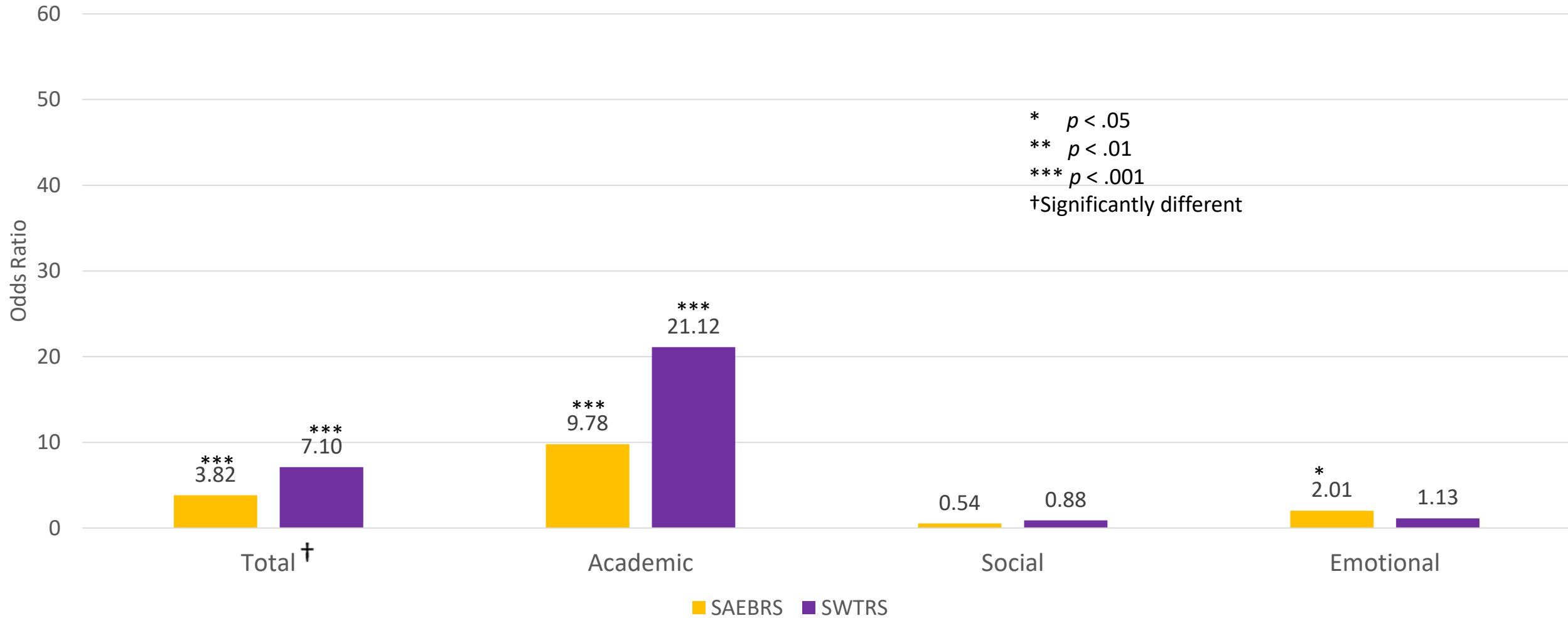
SWTRS and SAEBRS T1 Risk → T1 Academic Intervention
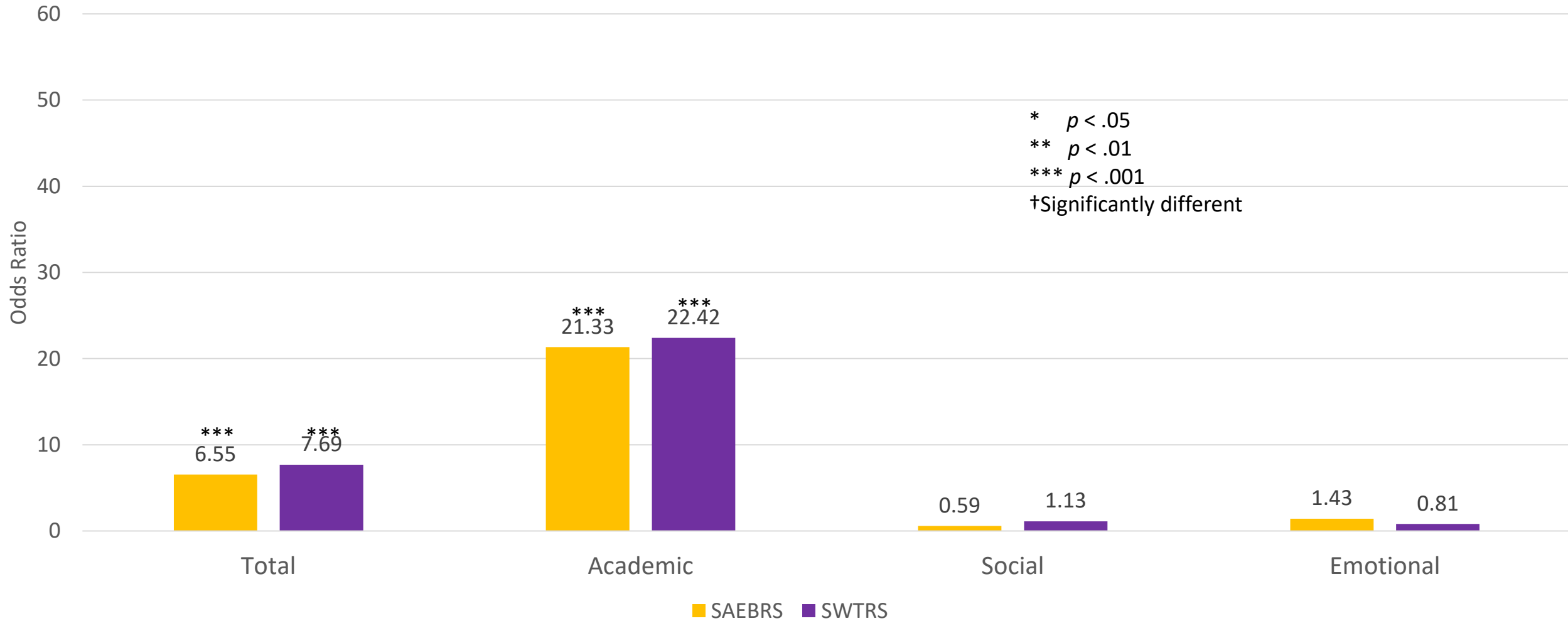
SWTRS and SAEBRS T1 Risk → T2 Academic Intervention

Summary & Conclusions

Summary

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

Summary

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

Overfitting SWTRS
- Replication samples / "Traning" and "Test" data sets needed

Staggered T1/T2 Data Collection
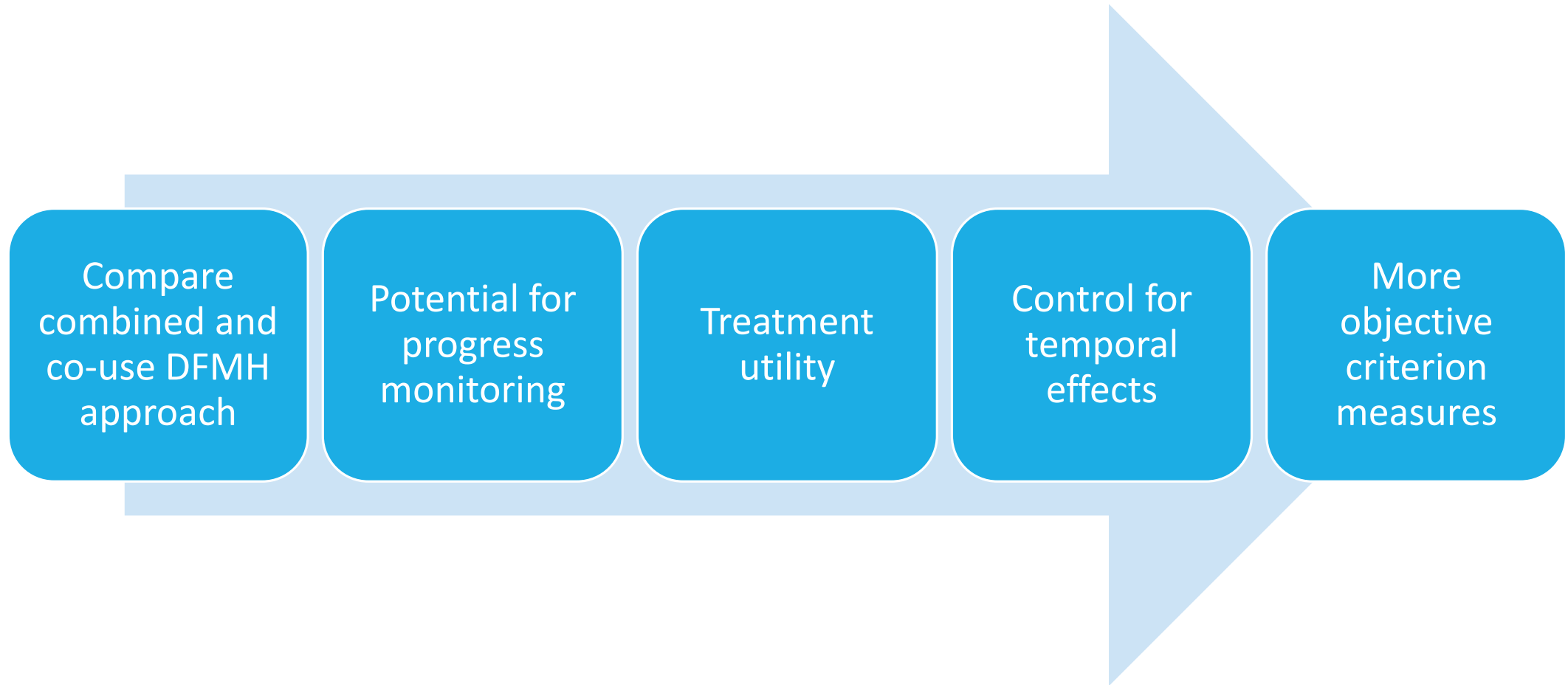- Only reflects 1st and 2nd instances of screening

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

References

1. Centers for Disease Control and Prevention. (2013). *Mental health surveillance among children: United States*, *2005–2011*. Washington, DC: Author.

2. Marmorstein, N. R., Iacono, W. G., & Malone, S. M. (2010). Longitudinal associations between depression and substance dependence from adolescence through early adulthood. *Drug and Alcohol Dependence*, *107*, 154–160. doi:10.1016/j.drugalcdep.2009.10.002

3. Rudolph, K. D., & Klein, D. N. (2009). Exploring depressive personality traits in youth: Origins, correlates, and developmental consequences. *Development and Psychopathology*, *21*, 1155–1180. doi:10.1017/S0954579409990095

4. Mathyssek, C. M., Olino, T. M., Velhurst, F. C., & van Oort, F. V. A. (2012). Childhood internalizing and externalizing problems predict the onset of clinical panic attacks over adolescence: The TRAILS study. *PLoS ONE*, *7*, e51564. doi:10.1371/journal.pone.0051564

5. Bradshaw, C. P., Buckley, J. A., & Ialongo, N. S. (2008). School-based service utilization among urban children with early onset educational and mental health problems: The squeaky wheel phenomenon. *School Psychology Quarterly*, *23*, 169–186. doi:10.1037/1045-3830.23.2.169

6. McIntosh, K., Flannery, K. B., Sugai, G., Braun, D. H., & Cochrane, K. L. (2008). Relationships between academics and problem behavior in the transition from middle school to high school. *Journal of Positive Behavior Interventions*, *10*, 243–255. doi:10.1177/1098300708318961

7. Roehrig, C. (2016). Mental disorders top the list of the most costly conditions in the United States: $201 Billion. *Health Affairs*, *35*(6), 1–6. doi:10.1377/hlthaff.2015.1659

8. Hoagwood, K., & Johnson, J. (2003). School psychology: A public health framework: I. From evidence-based practices to evidence-based policies. *Journal of School Psychology*, *41*, 3–21.

9. Burns, B. J., Costello, E. J., Angold, A., Tweed, D., Stangl, D., Farmer, E. M., & Erkanli, A. (1995). Children's mental health service use across service sectors. *Health Affairs*, *14*, 147–159. doi:10.1377/hlthaff.14.3.147

10. Wanzek, J., Otaiba, S. A., & Petscher, Y. (2014). Oral reading fluency development for children with emotional disturbance or learning disabilities. *Exceptional Children*, *80*, 187–204.

11. President's New Freedom Commission on Mental Health. (2003). *Achieving the promise: Transforming mental health care in America. Final report* (U.S. DHHS Pub. No. SMA-03–3832). Rockville, MD: U.S. Department of Health and Human Services.

12. Gresham, F. M. (2007). Response to Intervention and Emotional and Behavioral Disorders Best Practices in Assessment for Intervention. *Assessment for Effective Intervention*, *32*, 214–222.

References

13. Wilson J. M. & Jungner, F. (1968). Principles and practices of screening for diseases. Geneva: WHO.

14. Severson, H. H., Walker, H. M., Hope-Doolittle, J., Kratochwill, T. R., & Gresham, F. M. (2007). Proactive, early screening to detect behaviorally at-risk students: Issues, approaches, emerging innovations, and professional practices. *Journal of School Psychology*, *45*, 193–223.

15. Albers, C. A., & Kettler, R. J. (2014). Best practices in universal screening. In P. Harrison & A. Thomas (Eds.), *Best practices in school psychology: Data-based and collaborative decision making.* Bethesda, MD: The National Association of School Psychologists.

16. Dowdy, E., Ritchey, K., & Kamphaus, R. W. (2010). School-Based Screening: A Population-Based Approach to Inform and Monitor Children's Mental Health Needs. *School Mental Health*, *2*, 166-176. doi: 10.1007/s12310-010-9036-3

17. Stoiber, K. C. (2014). A comprehensive framework for multitiered systems of support in school psychology. In P. Harrison & A. Thomas (Eds.), *Best practices in school psychology: Data-based and collaborative decision making.* Bethesda, MD: The National Association of School Psychologists.

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

19. Ysseldyke, J., & Reschly, D. J. (2014). The evolution of school psychology: Origins, contemporary status, and future directions. In P. Harrison & A. Thomas (Eds.), *Best practices in school psychology: Data-based and collaborative decision making* (pp. 71–84). Bethesda, MD: The National Association of School Psychologists.

20. Greenspoon, P. J., & Saklofske, D. H. (2001). Toward an integration of subjective wellbeing and psychopathology. *Social Indicators Research*, *54*, 81–108. doi:10.1023/A:1007219227883

21. Suldo, S. M. & Shaffer, E. J. (2008). Looking beyond psychopathology: The dual-factor model of mental health in youth. *School Psychology Review*, *37*, 52–68. Retrieved from: https://www.researchgate.net/publication/228656864

22. Suldo, S., Thalji, A., & Ferron, J. (2011). Longitudinal academic outcomes predicted by early adolescents' subjective wellbeing, psychopathology, and mental health status yielded from a dual factor model. *The Journal of Positive Psychology*, *6*, 17–30. doi:10.1080/17439760.2010.536774

23. Kim, E. K., Furlong, M. J., Dowdy, E., & Felix, E. D., (2014). Exploring the relative contributions of the strength and distress components of dual-factor complete mental

24. Dowdy, E., Furlong, M., Raines, T. C., Bovery, B., Kauffman, B., Kamphaus, R. W., ... & Murdock, J. (2014). Enhancing school-based mental health services with a preventive and promotive approach to universal screening for complete mental health. *Journal of Educational and Psychological Consultation, 25*, 179–197. doi:10.1080/10474412.2014.929951

References

25. Furlong, M. J., You, S., Renshaw, T., Smith, D. C., & O'Malley, M. D. (2013). Preliminary development and validation of the Social and Emotional Health Survey for secondary school students. *Social Indicators Research*, *117*, 1011–1032. doi:10.1007/s11205-013-0373-0

26. Renshaw, T. L., Long, A. C. J., & Cook, C. R. (2014). Assessing adolescents' positive psychological functioning at school: Development and validation of the Student Subjective Wellbeing Questionnaire. *School Psychology Quarterly*. Advance online publication. doi:10.1037/spq0000088

27. Furlong, M. J., You, S., Renshaw, T. L., O'Malley, M. D., & Rebelez, J. (2013). Preliminary development of the Positive Experiences at School Scale for elementary school children. *Child Indicators Research*, *6*, 753–775. doi:10.1007/s12187-013-9193-7

28. Naglieri, J.A., LeBuffe, P.A., & Shapiro, V. (2010). *Devereux Student Strengths Assessment-mini*. Lewisville, NC: Kaplan Press.

29. 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. doi:10.1037/spq0000024

30. Roberson, A. J. (2016). Initial development and validation of the student wellbeing teacher-report scales (Unpublished master's thesis). Louisiana State University, Baton Rouge, Louisiana.

31. Kilgus, S. P., Eklund, K., von der Embse, N. P., Taylor, C. N., & Sims, W. A. (2016a). Psychometric defensibility of the Social, Academic, and Emotional Behavior Risk Screener (SAEBRS) Teacher Rating Scale and multiple gating procedure within elementary and middle school samples. *Journal of School Psychology*, *58*, 21–39. doi:10.1016/j.jsp.2016.07.001

32. Kilgus, S. P., Sims, W. A., von der Embse, N. P., & Taylor, C. N. (2016b). Technical adequacy of the Social, Academic, and Emotional Behavior Risk Screener in an elementary sample. *Assessment for Effective Intervention*, *42*, 46–59. doi:10.1177/1534508415623269

33. Krosnick, J. A., & Presser, S. (2010) Question and Questionnaire Design. In P. V. Marsden & J. D. Wright (Eds.), *Handbook of survey research* (pp. 263–313). Emerald Group Publishing.

34. American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

35. Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, *38*, 581–586.