Hammill Institute on Disabilities

Remedial and Special Education
2019, Vol. 40(2) 67–73

DOI: 10.1177/0741932517753821
rase.sagepub.com
SAGE

Validation of a Measure of Sustainability of School-Wide Behavior Interventions and Supports

Angus Kittelman, MA[1], Katherine W. Bromley, MA[1], Sterett H. Mercer, PhD[2], and Kent McIntosh, PhD[1]

Abstract

The purpose of this study was to evaluate the stability and validity of a measure assessing sustainability of school-wide behavior interventions, the *School-Wide Universal Behavior Support Sustainability Index: School Teams* (SUBSIST). We assessed the stability of the SUBSIST across 3 years of measurement. We also assessed the convergent validity of two key SUBSIST factors, Team Use of Data and District Capacity Building, with more direct measures of these constructs. Results showed stability of the measure across multiple years and statistically significant correlations when compared with other external measures. Implications for research and practice are discussed.

Keywords

sustainability, school-wide behavior interventions, data-based decision making, capacity building, convergent validity

There has been an increased focus from states and districts on prioritizing the sustained implementation of effective practices in schools (Klingner, Boardman, & McMaster, 2013). Accordingly, there has been renewed attention to empirical research identifying critical variables that affect the continued use of effective practices in schools (Coffey & Horner, 2012; Pinkelman, McIntosh, Raspica, Berg, & Strickland-Cohen, 2015; Sanford DeRousie & Bierman, 2012). To identify these variables, research-validated measures are needed to quantify them and assess the extent to which they predict the sustained use of effective practices in schools. Such measures, if technically adequate, could be used in empirical research to identify the relative predictive power of variables identified through case study or qualitative research as influential in sustainability.

Measuring Sustainability of School Practices

To aid in predicting sustainability, McIntosh, Doolittle, Vincent, Horner, and Ervin (2009) developed a measure, titled the *School-Wide Universal Behavior Sustainability Index: School Teams* (SUBSIST), which identifies factors found to be important for predicting the sustained implementation of behavior support interventions delivered to all students in a school. To confirm the SUBSIST assessed these factors, McIntosh and colleagues (2013) conducted an exploratory factor analysis (EFA). The EFA indicated a four-factor structure with two school-level factors, School Priority and Team Use of Data, and two district-level factors, District Priority and District Capacity Building, from a sample of 217 schools implementing School-Wide Positive Behavior Interventions and Supports (SWPBIS; Horner et al., 2014). McIntosh and colleagues (2013) then created a combined structural equation model from these district- and school-level factors to assess their association with concurrent fidelity of implementation of SWPBIS. Together, these factors explained 45% of the variance in fidelity of implementation, and each factor was significantly correlated with fidelity. However, only Team Use of Data and District Capacity Building were statistically significant independent predictors associated with implementation fidelity.

To cross-validate and address potential measurement differences of the SUBSIST, Mercer, McIntosh, Strickland-Cohen, and Horner (2014) conducted a series of multiple-group confirmatory factor analyses (CFAs) to determine whether factor loadings and item thresholds were equivalent across year of implementation. The sample included 860 schools implementing SWPBIS grouped by stage of implementation (Adelman & Taylor, 1997): 209 schools in the initial implementation stage (implementing for

[1]University of Oregon, Eugene, USA
[2]The University of British Columbia, Vancouver, Canada

Corresponding Author:
Angus Kittelman, Department of Special Education, University of Oregon, Eugene, OR 97403, USA.
Email: angusk@uoregon.edu

0–1 years), 408 schools in the institutionalization stage (2–4 years), and 233 schools in the sustainability stage (5 or more years). Results supported strong measurement invariance for all but one item. In addition, results from a partially scalar model demonstrated no statistically significant differences in the SUBSIST factor means between the institutionalization and initial implementation groups, but higher School Priority ($d = .36$) and Team Use of Data ($d = .52$) in the sustainability group compared with the initial implementation group.

Validation of the SUBSIST

Although the previous studies conducted by McIntosh et al. (2013) and Mercer et al. (2014) provide strong support for the four-factor structure of the SUBSIST, additional investigation of the stability and validity of the SUBSIST is warranted. For example, although the SUBSIST has identified factors important for the sustained implementation of SWPBIS at one point in time, additional validation is needed to understand whether sustainability, as measured by these critical factors, is stable over time across multiple years of measurement (McIntosh et al., 2013; Mercer et al., 2014). We define measurement stability as the extent to which a measure generates similar mean scores for the two school- and district-level constructs as measured on the SUBSIST across multiple years; if scores are highly stable across the years, then there may be limited need for school teams to reassess unless there has been a substantial change in the school or district context. In addition, the indirect, self-report nature of the SUBSIST (e.g., school SWPBIS team members self-reporting the extent to which the team uses data for decision making) could be a limitation if such self-reports do not correspond with more direct assessments of related variables (e.g., actual data use by these teams). Correspondence between the SUBSIST and more direct assessments of related variables would attest to the convergent validity of the SUBSIST measure. Specifically, the purpose of the current study was to assess the following research questions:

Research Question 1: To what extent are SUBSIST scores stable across multiple years of measurement?
Research Question 2: To what extent do SUBSIST factor scores at Year 1 and growth vary by implementation stage at Year 1?
Research Question 3: To what extent do two SUBSIST factors, Team Use of Data and District Capacity Building, correspond to more direct measures of these constructs?

Method

Participants and Setting

Participants for this study included a sample of school SWPBIS team members or district coaches representing their school in 14 U.S. states implementing SWPBIS, collected as part of a larger study and used previously to validate the SUBSIST (Mercer et al., 2014). Various samples used in the analyses in this study ranged from 454 to 788 schools. For the larger sample of schools, 212 (24.7%) were in the planning year or had been implementing for 1 year (initial implementation), 411 (47.8%) had been implementing for 2 to 4 years (institutionalization), and 237 (27.6%) had been implementing for 5 or more years (sustainability).

Measures

SUBSIST. The SUBSIST is a 39-item self-administered measure used to assess the critical factors that predict the sustained implementation of school-wide interventions when implemented with adequate fidelity. The SUBSIST contains both school-level (e.g., data are reviewed regularly at team meetings, the school team implementing SWPBIS meets at least monthly, the school administrators describe SWPBIS as a top priority for the school) and district-level (e.g., there are adequate district resources [funding and time] allocated for SWPBIS, school teams and new personnel are provided with professional development in SWPBIS at least yearly) items. Respondents are expected to rate the extent to which critical contextual variables perceived to be important for the sustainability of SWPBIS are implemented in their schools and districts based on a 4-point scale (1 = *not true* to 4 = *very true*). Initial validation of the SUBSIST was reported in McIntosh and colleagues (2011), which provided strong evidence of content validity (content validity index = .95) through an expert panel of reviewers and moderate to strong evidence of concurrent validity in comparison with the School-Wide Evaluation Tool (SET; Sugai, Lewis-Palmer, Todd, & Horner, 2001), a well-established measure that assesses the fidelity of implementation of SWPBIS in schools. In addition, McIntosh and colleagues (2011) reported that the SUBSIST had strong internal consistency ($\alpha = .77–.94$), test–retest reliability ($r = .96$), and interrater reliability ($r = .95$).

Generation of school discipline reports. Generation of school discipline reports was recorded using the *School-Wide Information System* (SWIS) Suite, a web-based information system responsible for collecting and summarizing student discipline data for decision making (May et al., 2013). Behavior data entered into SWIS include student office discipline referrals (ODRs), which allow school teams to generate basic reports to analyze overall school-wide referral patterns (e.g., number of referrals per month, referrals by location, referrals by problem behavior) and individual student patterns (i.e., number of referrals, when and where referrals occurred, type of referral [i.e., major or minor]). Student discipline information is then used by school teams for continuous data-based decision making to

better support students at the school-wide, small group, and individual student levels. When using SWIS for decision making, school teams generate reports of their data, and counts of reports generated are logged to indicate use of data for decision making. To calculate generation of school discipline reports, we recorded instances of schools generating SWIS reports by month in Year 1 of completing the SUBSIST. Using this information, we created a summary measure, generation of school discipline reports, of the proportion of months over Year 1 that at least one SWIS data report was generated in the school. This measure aligned closely with one of the 11 items in the SUBSIST Team Use of Data factor ("Data are reviewed regularly at team meetings") and generally to four others (e.g., "There is regular measurement of student outcomes [e.g., ODRs, achievement data, school safety surveys, student/parent satisfaction surveys]").

Formal SWPBIS training and coaching/technical support. Formal SWPBIS training and coaching/technical support were measured using the *Access to District Expertise and Professional Training* (ADEPT; McIntosh, Strickland-Cohen, & Horner, 2012), an online tool used to track district support in implementing and sustaining practices, such as SWPBIS. At the end of each school year, school teams tally the days spent in the following activities: (a) formal SWPBIS training focused on implementing SWPBIS (e.g., conferences, in-services) and (b) coaching/technical support from an external coach (e.g., attendance at team meetings, phone or email consultation). To calculate formal SWPBIS training, we first created a mean score by averaging those attended by one SWPBIS team member, the entire SWPBIS team, the building administrator, and all school staff for each of the 3 years. Using the three mean scores, we estimated a composite score for formal SWPBIS training across the 3 years of data collection. Similarly, we estimated a composite score for coaching/technical support using the cumulative count of coaching hours offered across the 3 years.

We assessed the technical adequacy of formal SWPBIS training and coaching/technical support as measured by the ADEPT in a subsample of the larger group that completed the ADEPT multiple times. Test–retest reliability was assessed by the same individuals after a 2-week latency (n ranged from 20–28). Test–retest reliability for formal SWPBIS training (one person, team, administrator, whole staff) was $r = .92$, formal SWPBIS training (team) was $r = .90$, formal SWPBIS training (administrator) was $r = .93$, and coaching/technical support was $r = .63$. Interrater reliability was assessed between individuals who worked within the same school (e.g., team facilitator and administrator; n ranged from 12–13). Interrater reliability for formal SWPBIS training (one person, team, administrator, whole staff) was $r = .86$, formal SWPBIS training (team) was $r = .76$, formal SWPBIS training (administrator) was $r = .87$, and coaching/technical support was $r = .83$. These measures aligned closely with two of the three items in the SUBSIST District Capacity Building factor ("School teams and new personnel are provided with professional development in SWPBIS at least yearly" and "The school team has regular access to district SWPBIS expertise [e.g., external/district coaches or consultants]").

Data Analysis

Analyses for Research Question 1. To examine the stability of the SUBSIST survey over multiple years of measurement, we used a subsample of schools with SUBSIST scores in Year 1 (2012–2013 academic year, $n = 788$), Year 2 (2013–2014 academic year, $n = 635$), and Year 3 (2014–2015 academic year, $n = 526$). The subsample of schools was selected based on schools with SUBSIST data during Year 1, with decreasing numbers in Years 2 and 3, due to attrition across the 3 years of data collection. We calculated composite scores for the two school-level (i.e., School Priority, Team Use of Data) and two district-level constructs (i.e., District Priority, District Capacity Building). We conducted descriptive analyses and bivariate correlations of the four composite scores across the 3 concurrent years of data collection.

Analyses for Research Question 2. To examine differences in the four SUBSIST factor scores across the three stages of implementation groups (i.e., initial implementation, institutionalization, and sustainability) across time, we selected a subsample of schools ($n = 506$) within districts with at least two schools in each district who completed the SUBSIST. Three-level hierarchical linear growth modeling (HLM) was used (Raudenbush & Bryk, 2002) to examine mean differences in the four SUBSIST factor composite scores by implementation stage at Year 1.We evaluated the extent to which the intercept and slope of each SUBSIST factor varied by stage of implementation at Year 1 by fitting a series of conditional growth models in which stages were included as categorical predictors (i.e., institutionalization vs. initial implementation and sustainability vs. initial implementation) of slope and intercept. Level 1 of the model was time (i.e., Year 1, Year 2, and Year 3), Level 2 was school, and Level 3 was district. We used HLM because it allows partitioning of the variance attributable to districts and schools implementing SWPBIS over time. In the current study, the intraclass correlation coefficients (ICCs) or the degree of variance at Level 3 ranged from .093 to .190, suggesting HLM analyses were necessary to account for nesting of schools within districts. In addition, HLM uses Maximum Likelihood (ML) estimation to account for data that are missing at Level 1.

Analyses for Research Question 3. To examine the extent to which Team Use of Data and District Capacity Building

Table 1. Descriptive Statistics and Intercorrelations for School- and District-Level Composite Scores Across 3 Years.

Variable	*n*	*M*	*SD*	% missing	1	2	3	4	5	6	7	8	9	10	11	12
1. SP Yr1	788	3.43	0.40	8.4	—											
2. SP Yr2	635	3.48	0.41	26.2	.52	—										
3. SP Yr3	526	3.48	0.41	38.8	.38	.57	—									
4. TD Yr1	785	3.37	0.55	8.7	.72	.44	.32	—								
5. TD Yr2	636	3.43	0.54	26.2	.39	.72	.47	.53	—							
6. TD Yr3	525	3.43	0.55	39.0	.29	.44	.74	.43	.60	—						
7. DP Yr1	787	3.05	0.63	8.5	.56	.28	.24	.47	.28	.23	—					
8. DP Yr2	635	3.13	0.65	26.2	.32	.54	.34	.35	.48	.30	.44	—				
9. DP Yr3	526	3.10	0.69	38.8	.30	.33	.57	.28	.32	.52	.40	.52	—			
10. DC Yr1	781	3.30	0.69	9.2	.50	.23	.20	.54	.28	.22	.49	.27	.26	—		
11. DC Yr2	635	3.31	0.71	26.2	.24	.41	.26	.36	.52	.31	.28	.54	.35	.36	—	
12. DC Yr3	526	3.27	0.72	38.8	.20	.23	.48	.28	.28	.50	.21	.39	.58	.32	.48	—

Note. All intercorrelations were significant at $p < .01$. SP = School Priority; Yr1 = year 1; Yr2 = year 2; Yr3 = year 3; TD = Team Use of Data; DP = District Priority; DC = District Capacity Building.

corresponded to more direct measures of those constructs, a subset of the total sample was used to calculate bivariate correlations between these SUBSIST factors and SWIS (n = 454) and ADEPT (n = 501) scores using Mplus 7.4 (Muthén & Muthén, 2015). The subsets of schools were selected for each of these analyses from the total sample based on whether (a) schools completed the SUBSIST and generated school reports using SWIS or (b) completed the SUBSIST and the ADEPT. Consistent with the analyses conducted by McIntosh and colleagues (2013), items from the SUBSIST were specified as ordered categorical variables using theta parameterization and the mean- and variance-corrected weighted least squares (WLSMV) estimator due to negative skew and limited response options for each item (Muthén & Muthén, 2015). The COMPLEX option in Mplus was used to adjust standard errors and chi-square tests of model fit due to the district-level clustering of schools (Asparouhov, 2005). SUBSIST factors were estimated using the four-factor SUBSIST model (McIntosh et al., 2013) using Year 1 data, and composite scores were calculated for SWIS data use and ADEPT formal SWPBIS training and coaching/technical support using 3 years of data.

Results

SUBSIST Stability

Table 1 and Figure 1 present the means, standard deviations, and intercorrelations of the dependent variables used to examine the extent to which the SUBSIST measure was stable across multiple years of measurement. We found all intercorrelations using composite factor scores to be statistically significant ($p < .01$). For School Priority, we found that scores were highly consistent from Year 1 (M = 3.43, SD = 0.40) to Year 3 (M = 3.48, SD = 0.41) and moderately correlated (r = .38–.57, $p < .01$). For Team Use of Data, scores were also highly consistent from Year 1 (M = 3.37, SD = 0.55) to Year 3 (M = 3.43, SD = 0.55) and moderately correlated (r = .43–.60, $p < .01$). Concerning the two district-level factors, District Priority was highly consistent from Year 1 (M = 3.05, SD = 0.63) to Year 3 (M = 3.10, SD = 0.69) and moderately correlated (r = .40–.52, $p < .01$), and District Capacity Building was also highly consistent from Year 1 (M = 3.30, SD = 0.69) to Year 3 (M = 3.27, SD = 0.72) and modestly correlated (r = .32–.48, $p < .01$).

SUBSIST Factor Variation by Stage of Implementation

The results of the conditional growth models are presented in Tables 2 and 3. We retained the random effects from the unconditional growth models and evaluated the differences in intercepts and slopes by stage of implementation. Inspection of the stage predictors indicated that schools implementing SWPBIS for 5 or more years showed higher School Priority and Team Use of Data scores at Year 1 compared with schools initially implementing SWPBIS, although inspection of pseudo r^2 values (School Priority, r^2 = .003; Team Use of Data, r^2 = .018) indicates these differences were trivial. All other effects were nonsignificant.

SUBSIST Relation to More Direct Measures

We found that Year 1 SUBSIST Team Use of Data was modestly and statistically significantly correlated with generation of discipline data reports in all implementation groups (initial implementation: r = .31, p = .008; institutionalization: r = .21, p = .001; sustainability: r = .20, p = .003; n = 454 schools). Also, we found that SUBSIST reports of District Capacity Building were positively correlated with

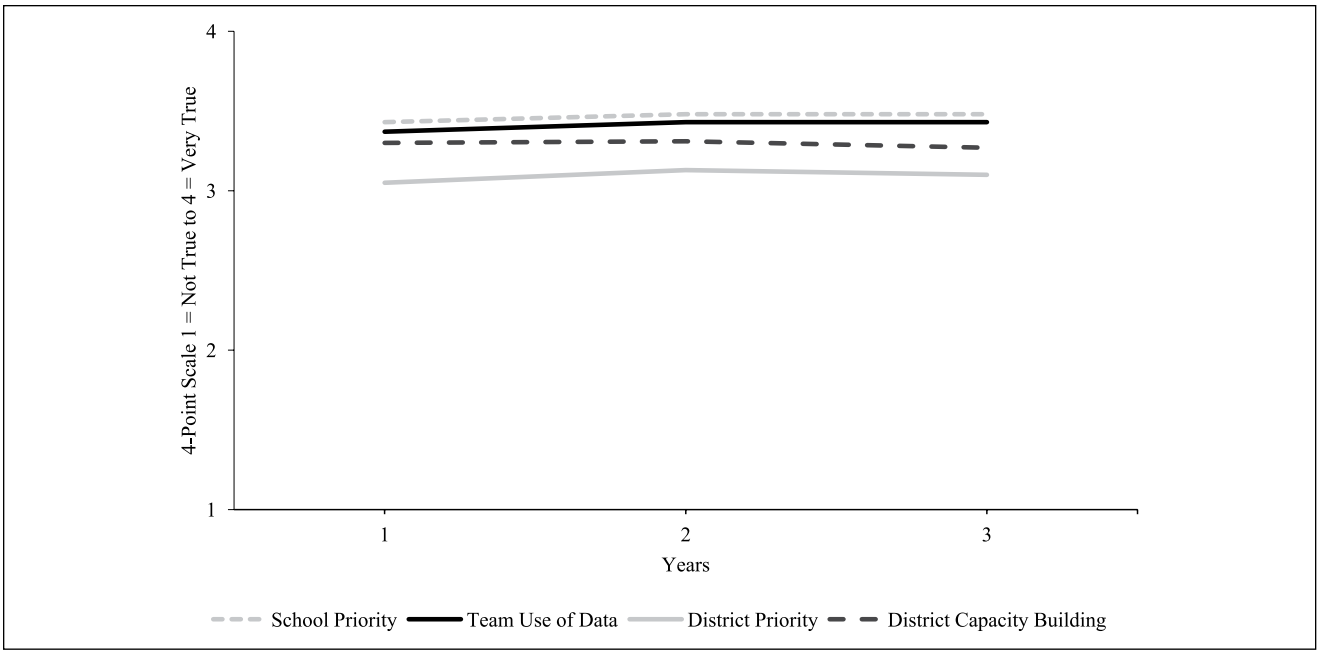

Figure 1. Evidence of the stability of the four-factor structure using composite scores of the SUBSIST across 3 concurrent years of data collection.
Note. SUBSIST = School-Wide Universal Behavior Support Sustainability Index: School Teams.

Table 2. Conditional Growth Model Predicting School-Level SUBSIST Composite Scores From Time and Implementation Stage (Centered at Time 1).

	School Priority					Team Use of Data				
Parameter	Coefficient	*t*	*df*	*SE*	*p*	Coefficient	*t*	*df*	*SE*	*p*
Intercept	3.37	80.84	59	0.04	<.01	3.26	55.02	59	0.06	<.01
Slope	−0.01	−0.37	441	0.03	.71	<−0.01	−0.06	59	0.04	.95
Inst.	0.01	0.18	441	0.05	.86	0.07	1.40	323	0.05	.16
Sust.	0.11	2.15	441	0.05	.03	0.18	2.30	323	0.08	.02
Inst. × Slope	0.05	1.67	441	0.03	.10	0.02	0.54	323	0.03	.59
Sust. × Slope	0.02	0.44	441	0.04	.66	0.02	0.40	59	0.05	.69

Note. School Priority pseudo r^2 = .003; Team Use of Data pseudo r^2 = .018. SUBSIST = School-Wide Universal Behavior Support Sustainability Index: School Teams; Inst. = institutionalization versus initial implementation stage; Sust. = sustainability versus initial implementation stage.

Formal SWPBIS Training over the 3 years (r = .30, p = .008) and Coaching/Technical Support (r = .36, p < .001; n = 501 schools). Although correlations were modest, they provide some convergent evidence that school SWPBIS team members' reports of practices related to data use (on the Team Use of Data factor) are related to their actual use of school data and that team members' reports of District Capacity Building corresponded to more direct reports of district-level trainings and availability of coaching.

Discussion

Prior research studies on the SUBSIST (a) provide strong evidence of the four-factor structure across schools at various stages of SWPBIS implementation and with varied student populations (McIntosh et al., 2013; Mercer et al., 2014) and (b) demonstrate that schools in the sustainability stage of implementation (5 or more years) report higher levels of School Priority and Team Use of Data than initially implementing schools (Mercer et al., 2014). To provide additional validation of the SUBSIST, this study sought to (a) evaluate the stability of the SUBSIST across multiple years of measurement, (b) evaluate the extent to which the intercept and slope of each SUBSIST factor varied by implementation stage at Year 1, and (c) provide further evidence of the convergent validity of the SUBSIST through concurrent associations of SUBSIST scores when compared with more precise measures of both school- (i.e., Team Use

Table 3. Conditional Growth Model Predicting District-Level SUBSIST Composite Scores From Time and Implementation Stage (Centered at Time 1).

	District Priority					District Capacity Building				
Parameter	Coefficient	*t*	*df*	*SE*	*p*	Coefficient	*t*	*df*	*SE*	*p*
Intercept	3.01	39.97	59	0.08	<.01	3.24	29.17	59	0.11	<.01
Slope	−0.06	−1.24	441	0.05	.22	−0.05	−0.57	59	0.08	.57
Inst.	0.01	0.10	441	0.08	.92	−0.04	−0.42	382	0.11	.68
Sust.	0.03	0.43	441	0.08	.67	0.08	0.70	382	0.12	.49
Inst. × Slope	0.08	1.38	441	0.06	.17	0.07	0.86	382	0.86	.39
Sust. × Slope	0.10	1.88	441	0.06	.06	<−0.01	<−0.01	382	0.09	1.00

Note. District Priority pseudo $r^2 = 0$; District Capacity Building pseudo $r^2 = .002$. SUBSIST = School-Wide Universal Behavior Support Sustainability Index: School Teams; Inst. = institutionalization versus initial implementation stage; Sust. = sustainability versus initial implementation stage.

of Data) and district-level (i.e., District Capacity Building) constructs.

As evidenced in Figure 1, mean scores for all four of the latent school- and district-level constructs were found to be highly stable, which attests to the consistency of the SUBSIST scores across multiple concurrent years of measurement. We also discovered that when stages of implementation were included as school-level predictors of slope and intercept in a series of conditional growth models, results indicated that schools implementing SWPBIS for 5 or more years had slightly higher School Priority and Team Use of Data scores in the first year of completing the SUBSIST survey compared with schools initially implementing SWPBIS. In addition, using Year 1 SUBSIST factor scores, we found statistically significant, modest convergent validity estimates when comparing the key school-level SUBSIST factor (i.e., Team Use of Data) to a more direct measure of this construct (i.e., generation of school discipline reports), and SUBSIST reports of the key district-level construct (i.e., District Capacity Building) with more direct reports of formal SWPBIS training and coaching/technical support using the ADEPT. The results showing somewhat stronger relations for the training and coaching measures than data use are not entirely surprising given that the SUBSIST Team Use of Data factor measures a construct broader than simply generating data reports.

Limitations and Future Research

The results of the current findings provide support for the stability of the SUBSIST measure; however, the results should be interpreted in light of a few key limitations. First, the amount of missing data, particularly in Years 2 and 3, was substantial and may have biased the interpretation of the results to some degree. Future research should investigate whether the SUBSIST measure is sensitive to identify which schools are most in danger of abandoning SWPBIS implementation and which criteria are most predictive. Second, data were collected for schools at three points in time; however, some schools in the sample had been implementing SWPBIS for multiple years before the start of the study. Although SUBSIST factor composites were stable overall, the analyses did not account for patterns of change that might exist based on context, such as years of implementation. For future research, results of the current study would be strengthened by matching year of implementation across waves of data collection.

Implications for the Improvement of Practice

Our main objective for this study was to validate additional aspects of a measure assessing critical factors important for the sustained implementation of school-wide behavior practices. Further validation ensures school teams have access to a more rigorously validated measure to assess whether critical features for sustaining school-wide behavior practices are in place in their school. If critical features are not in place, the SUBSIST has the potential to help teams identify what steps teams can take to ensure these features develop. For example, if a SUBSIST score is low on the construct of Team Use of Data, it could indicate for the external coach to provide more support to strengthen their data-based decision making. In addition, because the SUBSIST was found to be relatively consistent and stable across years, school teams implementing SWPBIS may not need to reassess using the SUBSIST at a high frequency, unless significant change occurs within the school or district (e.g., administrator or SWPBIS team turnover, changes in funding, loss of district support). Last, convergence between team-reported and more direct measures of factors like Team Use of Data strengthen the validity of decisions like providing more external coaching for teams reporting low Team Use of Data.

Conclusion

To support recent empirical efforts focused on identifying factors important for schools and districts to sustain

practices, we continued to evaluate the technical adequacy of a measurement tool used to assess important school- and district-level constructs related to the sustained implementation of school-wide behavior practices in schools. Overall, we found the SUBSIST measure to be both stable and consistent in assessing these important constructs over time. We also found constructs measured on the SUBSIST to be related to more direct measures schools are using, which further attests to the validity of the measure. In summary, this study provides further rigorous examination and validation of a measure used to identify factors important for the sustained implementation of school-wide behavior support practices.

Authors' Note

The opinions expressed are those of the authors and do not represent views of the Institute of Education Sciences or the U.S. Department of Education.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R324A120278 to University of Oregon.

ORCID iD


Sterett H. Mercer http://orcid.org/0000-0002-7940-4221

References

Adelman, H. S., & Taylor, L. (1997). Toward a scale-up model for replicating new approaches to schooling. *Journal of Educational and Psychological Consultation*, *8*, 197–230.

Asparouhov, T. (2005). Sampling weights in latent variable modeling. *Structural Equation Modeling*, *12*, 411–434. doi:10.1207/s15328007sem1203_4

Coffey, J. H., & Horner, R. H. (2012). The sustainability of schoolwide positive behavior interventions and supports. *Exceptional Children*, *78*, 407–422.

Horner, R. H., Kincaid, D., Sugai, G., Lewis, T., Eber, L., Barrett, S., . . . Johnson, N. (2014). Scaling up school-wide positive behavioral interventions and supports: Experiences of seven states with documented success. *Journal of Positive Behavior Interventions*, *16*, 197–208. doi:10.1177/1098300713503685

Klingner, J. K., Boardman, A. G., & McMaster, K. L. (2013). What does it take to scale up and sustain evidence-based practices? *Exceptional Children*, *79*, 195–211.

May, S. L., Ard, W. I., Todd, A. W., Horner, R. H., Glasgow, A., Sugai, G., & Sprague, J. R. (2013). *School-wide information system*. Eugene: Educational and Community Supports, University of Oregon. Available from http://www.pbisapps.org

McIntosh, K., Doolittle, J., Vincent, C. G., Horner, R. H., & Ervin, R. A. (2009). *School-wide universal behavior sustainability index: School teams*. Vancouver, Canada: The University of British Columbia.

McIntosh, K., MacKay, L. D., Hume, A. E., Doolittle, J., Vincent, C. G., Horner, R. H., & Ervin, R. A. (2011). Development and initial validation of a measure to assess factors related to sustainability of school-wide positive behavior support. *Journal of Positive Behavior Interventions*, *13*, 208–218. doi:10.1177/1098300710385348

McIntosh, K., Mercer, S. H., Hume, A. E., Frank, J. L., Turri, M. G., & Mathews, S. (2013). Factors related to sustained implementation of schoolwide positive behavior support. *Exceptional Children*, *79*, 293–311.

McIntosh, K., Strickland-Cohen, M. K., & Horner, R. H. (2012). *Access to District Expertise and Professional Training (ADEPT)* (Unpublished research measure). University of Oregon, Eugene.

Mercer, S. H., McIntosh, K., Strickland-Cohen, M. K., & Horner, R. H. (2014). Measurement invariance of an instrument assessing sustainability of school-based universal behavior practices. *School Psychology Quarterly*, *29*, 125–137.

Muthén, L. K., & Muthén, B. O. (2015). *Mplus user's guide* (7th ed.). Los Angeles, CA: Author.

Pinkelman, S., McIntosh, K., Raspica, C., Berg, T., & Strickland-Cohen, M. K. (2015). Perceived enablers and barriers related to sustainability of school-wide positive behavioral interventions and supports. *Behavioral Disorders*, *40*, 171–183.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: SAGE.

Sanford DeRousie, R. M., & Bierman, K. L. (2012). Examining the sustainability of an evidence-based preschool curriculum: The REDI program. *Early Childhood Research Quarterly*, *27*, 55–65.

Sugai, G., Lewis-Palmer, T. L., Todd, A. W., & Horner, R. H. (2001). *School-Wide Evaluation Tool (SET)*. Eugene, OR: Educational and Community Supports. Available from http://www.pbis.org