FURTHER VALIDATION OF A MEASURE TO ASSESS SCHOOL-WIDE POSITIVE BEHAVIOUR SUPPORT

by

Amanda E. Hume

B.A., University of California at Berkeley, 2007

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

(School Psychology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2011

Abstract

This study assessed aspects of construct validity of the School-wide Universal Behaviour Sustainability Index- School Teams (SUBSIST), a measure evaluating critical features of the school context related to sustainability of SWPBS systems. Two-hundred and seventeen schools were administered the measure. Construct validity of the measure was examined by determining whether SUBSIST scores were associated with features that are predictive of sustained SWPBS implementation, and whether the SUBSIST differentiated between different types of schools. First, analyses were performed to identify any differences in school features based on sustainability score. Results showed that higher sustainability scores were associated with increased frequency of school team meetings, presentation of SWPBS data to school staff, access to an external coach/consultant, and greater number of years sustaining SWPBS. Second, a two-step cluster analysis was performed to classify schools based on responses to the 39 SUBSIST items. A two cluster solution was obtained, with schools in one cluster (n = 139) obtaining significantly higher scores on SUBSIST items than schools in the other cluster (n = 78). The most critical item that contributed to cluster formation was use of data for decision making. These results are discussed with regard to previous and future research, limitations, and implications for sustaining SWPBS systems.

Preface

This thesis was conducted by the graduate student, under the advisement of her research supervisor. The graduate student obtained data from a large study where her responsibilities included aiding in data collection and writing, and taking a lead role in analysis. Ethics approval was required by the UBC Behavioural Research Ethics Board (BREB) to conduct this research. The UBC BREB certificate number is H08-01010.

In regard to this thesis, the graduate student was primarily responsible for the design, analysis, and writing. Thus, this thesis represents her work as the lead researcher and author.

Table of Contents

Abstract ..... **ii**

Preface ..... **iii**

Table of Contents ..... **iv**

List of Tables ..... **vi**

Acknowledgments ..... **vii**

1. Introduction **1**

1.1 Sustaining effective practices 1

1.2 Processes underlying sustainability 3

1.3 A measure assessing sustainability factors 6

1.4 Construct validity 8

1.5 The current study 9

1.6 Research questions 9

2. Method **11**

2.1 Participants and settings 11

2.2 Measure 11

2.3 Procedure 13

2.4 Analysis 14

2.4.1 Differences in relation to sustainability score 14

2.4.1.1 Correlation procedure 14

2.4.1.2 Mean comparisson procedure 14

2.4.2 Cluster analysis 16

2.4.2.1 Cluster analysis procedure 16

2.4.2.2 Formation of clusters 14

2.4.2.3 Evaluating significance of the cluster solution 16

2.5 Missing data 21

3. Results **22**

3.1 Relation between sustainability scores and school features 22

3.1.1 Years implementing 22
3.1.2 Mean score differences 23
3.2 Cluster analysis 24
3.3 Critical items 26
3.4 Criterion analyses 30
3.4.1 Chi-square tests of independence 32
3.4.2 Independent-samples t-tests 36
4. Discussion 40
4.1 Overview 40
4.2 Construct validity of the SUBSIST 41
4.3 Processes underlying sustainability 44
4.4 Limitations and future research 45
4.5 Implications for practice 47
4.6 Conclusions 48
References 49

List of Tables

Table 3.1 Results of *t*-tests by variable 24

Table 3.2 SPSS two-step auto-clustering results 25

Table 3.3 Predictor importance values, means, and standard deviations of items 26

Table 3.4 Critical items with predictor importance statistics > .60 and cluster group comparisons 29

Table 3.5 Proportions of variables by clusters 31

Table 3.6 Proportions of clusters by urbanicity 33

Table 3.7 Proportions of clusters by grades served 34

Table 3.8 Proportions of clusters by Title I eligibility 34

Table 3.9 Proportions of clusters by meeting frequency 35

Table 3.10 Proportions of clusters by frequency of data presentation 35

Table 3.11 Proportions of clusters by access to external coach. 36

Table 3.12 Mean percent of students receiving free or reduced priced lunch 37

Table 3.13 Mean number of students enrolled 38

Table 3.14 Mean years of SWPBS implementation by cluster 39

Acknowledgments

I wish to thank my Committee for their support and guidance, particularly my research supervisor, Dr. Kent McIntosh. Kent, I am grateful for all the opportunities and knowledge you have provided me the past three years, as well as for your frequent encouragement, positive outlook, and many edits, your support has been invaluable.

I would also like to thank Dr. Anita Hubley for her wealth of knowledge and copious notes about measurement and statistics. Thanks to Dr. Bruno Zumbo for his expert knowledge and feedback in the areas of statistics and research design. Many thanks to my cohort (a.k.a., wonderful friends!) for their frequent humour, distraction, and motivation; knowing I always had someone who would understand what I was talking about made all the difference in the world. Finally, thanks to my parents for their love, encouragement, and unwavering belief in me.

1. Introduction

1.1 Sustaining effective practices

Although there has been significant attention focused on identifying and implementing evidence-based practices in schools, the task of sustaining these practices is often overlooked in both research and practice (McIntosh, Filter, Bennett, Ryan, & Sugai, 2010; Vaughn, Klingner, & Hughes, 2000). Because sustainability is so critical to ensuring that students have continuing access to evidence-based practices over time, the field of education could benefit from greater attention to researching the construct of sustainability and how it may be promoted.

There are many benefits when effective school-based practices are sustained. Outcomes for sustaining practices include reductions in long-term costs and improved student outcomes (Filter et al., 2007; Horner et al., 2009; McIntosh, Horner, & Sugai, 2009). The sustained use of evidence-based practices can improve academic, social, and emotional outcomes (Schaughency & Ervin, 2006; Walker, 2004). Potential benefits for staff include increased job satisfaction. Factors cited in the literature that are likely to result in increased job satisfaction among teachers include reduction in workload, having experienced leadership, and opportunities for staff collaboration (Sargent & Hannum, 2005). These themes are consistent with outcomes realized when sustaining evidence-based practices (Foorman & Moats, 2004; Schaughency & Ervin, 2006).

There are also large costs associated with abandoning existing practices to implement new programs. The financial cost of new school reform initiatives is

undeniably greater than the cost of sustaining already implemented practices (Latham, 1988). In addition to financial costs, there are social costs that result from the continual turnover of practices. A long-term study of teachers' responses to reform initiatives found that teachers are likely to return to traditional practices because the demands of reform efforts can become too burdensome (Tyack & Cuban, 1995). Valli and Buese (2007) referred to a process where teachers were expected to comply with the increasing demands of innovations--under the same or worsening conditions--as "intensifying" teacher roles. This process results in increased demands on teachers and feelings of frustration, which at best can be viewed as an annoyance, and at worst can lead to staff turnover or opposition to beneficial change. Clearly, sustaining valued practices benefits staff by relieving them of the burden of having to continually adopt new practices. Due to the financial and social costs associated with program turnover, it is not in the best interest of students and staff to continually implement new practices before attempting to sustain current ones that have led to improved outcomes (Coburn, 2003; McLaughlin & Mitra, 2001).

Although it is easy to see the benefits of sustaining effective practices, there is only a small research base exploring sustainability. Difficulties in sustaining evidence-based practices in school settings have been well documented (Fuchs & Fuchs, 1998; Santangelo, 2009), but there is little research documenting the mechanisms by which sustainability occurs. Because of the numerous barriers that frequently derail efforts to sustain practices in schools, it is important to explore how these barriers can be

overcome. Thus, to promote sustainable practices, it is important first to determine how and why sustainability occurs.

1.2 Processes underlying sustainability

McIntosh, Horner, and Sugai (2009) proposed a contextually-based model of sustainability that can be applied to any school-based practice. Sustainability begins when an institution (e.g., school, district) identifies valued outcomes to achieve through a change process (priority). Next, the institution will identify and adopt practices that will produce valued outcomes (effectiveness) and implement critical features of the practice with fidelity (efficiency). The process is refined through continuous regeneration, which the researchers define as the combination of data-based problem solving, continuous measurement, and capacity building. Continuous regeneration of practices is an iterative process that focuses on the use of data to enhance priority, effectiveness, and efficiency, which allows practices to continue to produce valued outcomes.

McIntosh et al.'s (2009) model of sustainability has particular application to the practice of School-wide Positive Behaviour Support (SWPBS). SWPBS is a research validated approach that has been shown to decrease problem behaviours and improve academic outcomes. Through SWPBS, school personnel establish a positive social climate, where expectations for student behaviours are well-defined, directly taught, and consistently acknowledged and monitored by staff (Sugai & Horner, 2009; Sugai et al., 2000). Sustainability of SWPBS practices is of particular interest because of its expansion in schools across North America over the past decade (Sugai & Horner, 2009).

Much of the current knowledge regarding sustainability of SWPBS has been generated through conceptual models and analysis of single-case studies (Taylor-Greene et al., 1997). The need for large-scale analysis of factors promoting sustainability of SWPBS has been recognized and is just beginning to take shape (Sugai & Horner, 2006).

One large-scale analysis was conducted using 117 schools across six U.S. states to identify features important to sustaining SWPBS practices (Coffey & Horner, in press). Results of this quantitative study showed that administrative support, communication, and data-based decision making assisted schools in the sustained implementation of SWPBS over at least five years. Additionally, the study found that schools that were successfully sustaining SWPBS were most likely to have the following critical features of SWPBS in place: (1) teaching behaviour expectations, (2) a reward system, and (3) a system of monitoring and decision making.

Single-case and descriptive studies have also contributed to the literature regarding factors related to implementation and sustainability. Taylor-Greene et al. (1997) conducted a single-case design study in a moderate sized rural middle school in its first year of SWPBS implementation. Their findings stressed the importance of school preparedness; having structures in place to support the implementation of SWPBS practices. Staff buy-in during pre-implementation and early implementation stages was a key component in both implementing and sustaining SWPBS practices. In a follow-up article, they highlighted that sustaining SWPBS systems was achieved through systems-level structures, including defined improvement goals, administrative support, teamwork, program reinforcement, and formative evaluation (Taylor-Greene & Kartoub, 2000).

Research using qualitative methodology has provided rich information regarding factors that are thought to promote or impede sustainability of SWPBS. A study by Kincaid, Childs, Blasé, and Wallace (2007) gathered statements from 70 school personnel (representing 26 schools and 18 school districts) implementing SWPBS for at least one year. Six areas were consistently noted as promoting or impeding initial implementation: administrative support, faculty buy-in, philosophical differences, staff training, student training, and reward systems.

Sustainability of SWPBS has also been examined in the context of individual student support systems in SWPBS. Using a qualitative design, Bambara, Nonnemacher, and Kern (2009) interviewed 25 participants, representing five different stakeholder groups (teachers, school administrators, parents, external facilitators, and internal facilitators). All participants were members of their school SWPBS team. From the responses, several factors were identified as enablers to sustainability, including school culture, building administrator leadership and support, efficient use of time, capacity building (ongoing professional development), and family and student (stakeholder) involvement.

Research has also highlighted numerous potential barriers that threaten to derail the continual success of existing initiatives. Lohrmann, Forman, Martin, and Palmieri (2008) used interviews with 14 technical assistance providers who had been assisting schools or districts with implementing and sustaining SWPBS for at least one year to identify five common barriers to sustaining SWPBS practices. Barriers included lack of administrative direction and leadership, skepticism that universal intervention is

necessary, hopelessness about change, philosophical differences, and staff feelings of disenfranchisement from each other, the administrator, or the mission of the school. Barriers cited in other studies include student and staff turnover, scarce access to resources, poor contextual fit of the practice with school priorities, and failure to produce desired outcomes (Han & Weiss, 2005; McLaughlin & Mitra, 2001; Valli & Buese, 2007). Identifying potential barriers school teams face in attempting to sustain SWPBS may help in the process of sustainability (Curtis, Castillo, & Cohen, 2008).

Though these studies provide some initial insight into factors regarding sustainability, the majority of research on the literature base includes anecdotal accounts of instances where practices have failed to sustain. There remains little large scale research focusing on the mechanisms and conditions that promote sustainability (Gersten, Chard, & Baker, 2000). Efforts to identify conditions that promote sustainability are necessary to help the growing numbers of newly implementing and currently sustaining schools continue to sustain SWPBS at high levels of fidelity.

1.3 A measure assessing sustainability factors

To explore the factors related to sustainability on a large scale, a quantitative measure was created to assess sustainability of school-based behaviour interventions. The *School-wide Universal Behaviour Sustainability Index - School Teams* (SUBSIST; McIntosh, Doolittle, Vincent, Horner, & Ervin, 2009) is a research instrument designed to evaluate schools' capacity to sustain SWPBS efforts at the universal tier. The SUBSIST aims to

identify variables and features found in the context of the school environment that enhance or impede sustainability. An individual with knowledge of a school's SWPBS systems is asked to rate, using a 4-point scale from *not true* to *very true,* the extent that features hypothesized to affect sustainability are currently in place at their specific school. Content for the measure was based on the results of a comprehensive literature review, the authors' extensive experiences with SWPBS, and the results of Coffey and Horner's (in press) study on sustainability of SWPBS (McIntosh et al., in press).

The survey was initially designed with 49 to 50 items (depending on the question asked), and initial psychometric properties are based on the initial survey. The content was validated through ratings from an expert panel indicating strong content validity for measuring sustainability (content validity index = .95). Results from a pilot study indicated strong inter-rater reliability (r = .95) and two-week test-retest reliability (r = .96). Additionally, the study found strong internal consistency within each subscale (Cronbach's alpha range = .77 to .94), and moderate convergent validity (r = .68) with the School-wide Evaluation Tool (SET; McIntosh et al., in press; Sugai, Lewis-Palmer, Todd, & Horner, 2001). A more recent study has indicated a four-factor structure, with strong concurrent validity with measures of SWPBS fidelity of implementation (McIntosh, Mercer, Hume, Frank, & Turri, 2011).

1.4 Construct validity

The SUBSIST is hypothesized to evaluate features of the environment that are necessary to describe the construct of sustainability, or the potential for sustained implementation of a practice. To make inferences regarding the appropriateness and usefulness of the SUBSIST's utility in explaining sustainability, it is important that the measure show evidence of construct validity. According to Cohen and Swerdlik (2009), "construct validity is a judgment about the appropriateness of inferences drawn from test scores regarding individual standings on a variable called a construct." Evidence of construct validity can be shown by demonstrating that scores on the measure are related to other variables in expected directions or vary as a function of membership in different groups (e.g., schools more or less likely to sustain SWPBS).

One way of providing construct validity is by showing the SUBSIST can differentiate between schools. Evidence for construct validity can be provided by demonstrating that features associated with sustainability vary with schools' SUBSIST scores in expected directions. Variables that are expected to positively affect sustainability include number of years implementing, and various school team actions, such as regular team meetings, presentation of data, and access to coaching. Schools implementing school-based interventions for more years and reporting more frequent meetings, presentation of data, and access to coaching are hypothesized to have higher sustainability scores (Adelman & Taylor, 2003; McIntosh et al., 2011; Taylor-Greene & Kartoub, 2000).

1.5 The current study

The purpose of this study was to assess the construct validity of the SUBSIST measure for assessing variables related to sustained implementation. The current study builds on research assessing the psychometric properties of the SUBSIST, including internal consistency, interrater reliability, test-retest reliability, content validity, concurrent validity, and factor analysis (McIntosh et al., in press; McIntosh et al., 2011). The aim of this study was to examine additional aspects of the measure's construct validity, namely whether schools differ based on their SUBSIST scores.

To determine whether the measure can be used to assess sustainable practices of SWPBS, schools' SUBSIST sustainability scores were evaluated in relation to school features that are hypothesized to be related to sustained implementation (years implementing SWPBS and school team actions). To further assess whether the measure can differentiate between different types of schools, schools were classified into clusters (subgroups) based on responses to the SUBSIST items.

1.6 Research questions

This study examined the following research questions:

1) Is the SUBSIST sustainability score associated with other variables theoretically related to sustainability of SWPBS?

2) Can schools be classified into valid clusters based on their responses to the SUBSIST measure?

2. Method

2.1 Participants and settings

Participants from 217 schools in 14 U.S. states were included in the study. Participants were school team members or external personnel that have knowledge of their school's SWPBS systems. The majority (43%) of participants identified themselves as a school team leader/facilitator (also known as internal coach), followed by school administrator (32%), school team member (12%), external/district coach (9%), and not specified (4%).

Each of the participants completed the survey in regards to their particular school. Of the 217 schools represented in the study, 50% were elementary schools (Grades K-3, 5, or 6), 16% were middle schools (Grades 6-8), 5% were high schools (Grades 9-12), 4% were Grades K-8, 1% were Grades K-12, 1% were Pre-K, and 23% were unspecified. The schools represented a range of communities (15% urban, 25% small/large city, 28% suburban, 4% semi-rural, 11% rural, and 18% unspecified). The median category of students receiving free and reduced priced lunch was 50-74%. On average, schools had been implementing SWPBS for 5.4 school years, with a standard deviation of 3.2 years and a range from 1 to 15 years.

2.2 Measure

The SUBSIST is composed of 39 items hypothesized to be critical features of the SWPBS systems or school context related to sustainability, such as "SWPBS (also known

as School-wide PBS, PBIS, or EBS) serves a critical need for the school" and "There are adequate district resources (funding and time) allocated for SWPBS." These items are organized into seven subscales, including Priority, Building Leadership, External Leadership, Effectiveness, Efficiency, Use of Data, and Capacity Building, plus a sustainability score, which is an overall item average. In keeping with results from a recent validation study (McIntosh et al., 2011), 10 items were removed from the survey for analyses due to item redundancy and poor factor loading for one subscale.

The SUBSIST takes approximately 40 to 60 minutes to complete. For each of the 39 items, respondents answer the question "To what extent is this statement true for your school right now?" Item responses include; 1, *Not True*; 2, *Partially True*; 3, *Mostly True*; and 4, *Very True*. There is also the option to respond *Don't Know/NA*. The SUBSIST also includes 12 open-ended and multiple-choice questions regarding school demographics and actions of the school SWPBS team, such as "What is your school's enrollment?" and "How often does your school SWPBS team meet (during the school year)?". Demographic and school team action variables were also used as variables in the study.

Sustainability scores are calculated for each participating school by averaging their 39 SUBSIST item scores. This score is an estimate of the presence of research-supported features of the school environment that are predictive of sustainability of SWPBS. Higher sustainability scores are hypothesized to be associated with greater likelihood of sustained SWPBS practices. For the current sample, the internal consistency for the SUBSIST sustainability score was .94, with inter-item correlations ranging from

-.02 to .60, with an average of .30. The average sustainability score was 3.26, with a standard deviation of .47.

2.3 Procedure

Recruitment to complete the SUBSIST measure began in October, 2009. Participants were recruited through two methods. First, state SWPBS coordinators were contacted, requesting that they forward an invitation to participate to school and district personnel who might be interested in participating. Second, the authors invited a list of 383 schools reporting at least five years of fidelity of implementation data to the U.S. Center on Positive Behavioral Interventions and Supports to participate. Emails or hard copy letters (depending on what contact information was available) were sent to the 383 schools on the recruitment list. Participants were invited to access a web-based version of the survey or complete a paper version. Participants were entered into a draw for a $100 gift certificate. Follow-up invitations (including a hard copy of the measure) were sent approximately one month after the initial invitation. A second follow-up hard copy invitation was sent six weeks later.

A total of 217 surveys were included in the analysis. Of the 217 school personnel that completed surveys, 79 respondents were invited by a statewide coordinator. From the list of schools the researchers invited, 123 responded, a response rate of 32%, which is above the suggested minimum criterion of 30% for survey responses (Dillman, 1983). All surveys were completed during the 2009-2010 school year. Additionally, responses from

the 15 SUBSIST pilot study participants (McIntosh et al., in press), who completed the surveys during the 2008-2009 school year, were included in the study.

2.4 Analysis

2.4.1 Differences in relation to sustainability score

Sustainability scores were evaluated in relation to years implementing SWPBS and school team actions to assess whether differences in sustainability scores were associated with differences in school-based variables. The following variables were evaluated:

(1) Years implementing. Number of years the school has been implementing SWPBS.

(2) Meetings. Frequency of SWPBS school team meetings (i.e., at least once per month or fewer than once per month).

(3) Data presented. Frequency that data are presented to all school personnel (i.e., at least once per month or fewer than once per month).

(4) External coach. Access to an external coach, facilitator, or consultant dedicated to supporting SWPBS at least one hour per month (i.e., access or no access).

2.4.1.1 Correlation procedure

A Pearson's product-moment correlation coefficient was calculated to determine the size and direction of the relationship between the sustainability score and years implementing SWPBS. Scatterplots and tests of normality were produced to determine if

variables met assumptions of normality, linearity, and homoscedasticity. In the instance of a variable not meeting assumptions, both parametric (Pearson's) and non-parametric (Spearman's rho) correlations were produced. If the decision to retain or reject the null hypothesis were equivalent for both correlations, results of parametric correlations were reported for ease of interpretation. If the decision to retain or reject the null hypothesis differs based on the type of procedure conducted, the result of the non-parametric procedure were reported. Cohen's (1988) criteria for effect sizes were used to indicate whether the correlation was small (r = .10 to .29), medium (r = .30 to .49), or large (r >.50).

2.4.1.2 Mean comparison procedures

Independent-samples *t*-tests were performed on categorical variables (meetings, data presented, and external coach) to compare differences in sustainability scores across two response categories for each variable. Before conducting each analysis, the researcher determined whether assumptions of normality and homogeneity of variances had been met. Skewness and kurtosis values, the Shapiro-Wilk statistic, and Levene's test for equality of variances were considered. Additionally, histograms were used to inspect the shape of variable distribution by cluster group. In the instance of a variable not meeting the assumption of normality or homogeneity, the alternative non-parametric Mann-Whitney U Tests were also conducted. When the decision to retain or reject the null hypothesis was equivalent for both parametric and non-parametric analyses, results of parametric (*t*-test) procedures were reported for ease of interpretation. If the decision to retain or reject the null hypothesis differed based on the type of procedure conducted,

the result of the non-parametric (Mann Whitney U) procedures were reported. An effect size statistic (Cohen's *d*) was calculated for each finding and criteria for interpretation were applied as follows: .20 (small), .50 (medium), and .80 (large; Cohen, 1988). To correct for an increase in Type I error, a Bonferroni correction was used based on the number of comparisons conducted ($n = 3$, $\alpha = .017$).

2.4.2 Cluster analysis

Cluster analysis refers to a family of statistical techniques that can be used to group (or classify) individual cases based on their similarities. The number of clusters does not need to be predefined, meaning that the analysis can produce a classification scheme for previously unclassified data. There are numerous reasons to perform a cluster analysis, including the desire to distinguish between several subgroups found within a population, construct a scheme for classifying entities, and test hypothesized subgroups believed to exist within a population (Lorr, 1983).

2.4.2.1 Cluster analysis procedure

A two-step cluster analysis was performed using SPSS (v. 19). Schools ($n = 217$) served as the unit of analysis for clustering purposes. The 39 SUBSIST items were the variables used to form clusters of schools. Two-step cluster analysis was selected because of its relative advantages over other cluster analysis techniques. Two-step cluster analysis was designed to accommodate very large datasets and is the recommended analysis when over 1000 entities are present (SPSS, 2004). Although only 217 schools were included in this analysis, preclustering the data was a desirable approach given the large number of items in the measure. Another advantage of two-step cluster analysis is that it generates a

standardized predictor importance value for each variable, which allows assessment of the relative contribution of each variable to cluster formation. A traditional hierarchical factor analysis does not provide specific information regarding the predictor importance of each variable. A two-step cluster analysis procedure was also preferred because the optimal number of clusters was unknown, and two-step procedures do not require *a priori* specification (Norusis, 2010).

The cluster analysis was performed using the procedure recommended by Norusis (2010). Before the analyses were conducted, all 39 items were standardized by z-score. The first step in two-step cluster analysis is preclustering, or initially grouping entities in a sequential fashion based on distance criterion to reduce the size of the data matrix. The distance measure used to group cases was log-likelihood, which is recommended when variables are not considered completely independent from one another. In this sample, the SUBSIST Pearson's correlations between items ranged from -.02 to .60. After preclustering data, each precluster was treated as a single entity. Next, preclusters were grouped into clusters using an agglomerative hierarchical technique. Schwarz's Bayesian Information Criterion (BIC) was used to produce the optimal number of clusters. The maximum number of clusters was limited to 15. Outliers were placed in their own cluster (according to the 25% differentiation option) to prevent inflation of the overall number of clusters.

2.4.2.2 Formation of clusters

SPSS uses an algorithm evaluating change in BIC values to create the optimal number of clusters. Formulas for the calculation of BIC values and the Log-likelihood distance measures are given by Chiu et al. (2001).

For each SUBSIST item, a standardized *t* statistic was calculated to compare the cluster mean to the overall mean. Items with a statistically significant *t* statistic are considered critical in the formation of clusters (Norusis, 2010). A predictor importance value (*p*-value) for each item is produced based on the *t* statistic distribution; the predictor importance value signifies the relative contribution of each item to the cluster formation (SPSS, 2010). The greater an item's predictor importance value, the more critical that item was in formation of the cluster solution.

2.4.2.3 Evaluating significance of the cluster solution

A drawback of cluster analysis is that it always produces clusters, whether or not they are valid (Dubes & Jain, 1979). Therefore, significance tests of the cluster solution were conducted. Following recommendations by Milligan (1996), statistical validity of the cluster solution was tested using both internal and external criterion analysis. Internal criterion analyses use information obtained from within the clustering process to assess how well the variables cluster. The analyses examined the structure silhouette measure of cohesion and separation provided by SPSS v.19 (Kaufman & Rousseeuw, 1990). This measure assesses the closeness of variables within clusters to closeness between clusters.

External criterion analyses are post hoc tests for significant differences between the clusters using variables not included in the cluster analysis. In this study,

demographic variables and school team actions were used for the external criterion analyses, as they were not used in the analysis. The majority of demographic information was provided by participants; however, some information was supplemented through the National Center for Education Statistics (NCES). The following school demographics were used to describe the clusters:

(1) Urbanicity. Description of the neighbourhood locale (rural/semi-rural, suburban, small/medium/large city).

(2) Grade. Grade levels served by the school. Grade levels were grouped into categories according to the following criteria: Elementary (schools serving any combination of Grades pre-K - 6), Middle School (schools serving any combination of Grades 5 - 8), High School (schools serving any combination of Grades 9 - 12), and Other (K-8, K-12).

(3) Enrollment.[1] Number of students enrolled at the school.

(4) Free or reduced lunch.[1] Percent of students receiving free or reduced lunch.

(5) Title I.[1] Eligibility for Title I funding.

(6) Years implementing. Number of years the school has been implementing SWPBS.

[1] Obtained from NCES database.

Additionally, the following dichotomous school team actions were used to describe the clusters:

(7) Meetings. Frequency of SWPBS school team meetings (at least once per month or fewer than once per month).

(8) Data presented. Frequency that data are presented to all school personnel (at least once per month or fewer than once per month).

(9) External coach. Access to an external coach, facilitator, or consultant dedicated to supporting SWPBS at least one hour per month (access or no access).

Nonparametric chi-squared statistics were calculated for categorical dependent variables (urbanicity, grade, Title I, meetings, data presented, and external coach). A chi-squared (χ^2) statistic compares the distributions of observed data within each cluster and determines the level of consistency between clusters. Independent *t*-tests, or the equivalent nonparametric alternative, were calculated to compare mean differences between cluster groups for continuous dependent variables (enrollment, free or reduced lunch, and years implementing).

Chi-square tests of independence were performed following the procedure recommended by Gravetter and Wallnau (2004). For each analysis, an evaluation of the 'minimum expected cell frequency' was performed. To satisfy the minimum cell frequency assumption, it was expected that at least 80% of the cells contained a minimum of five schools. A Yates' Correction for Continuity was performed when reporting a 2 x 2 table to correct for overestimation of the chi-square value. An *a priori* alpha value of .05 was set. An effect size value was reported for each analysis. For 2 x 2 analyses (Title I, meetings, data presented, and external coach), a Φ correlation coefficient was reported. Cohen's (1988) criteria of .10 for a small effect, .30 for a medium effect, and .50 for a

large effect were applied when interpreting the Φ coefficient. For tables larger than 2 x 2 (urbanicity and grade), Cramer's V, which takes into account the degrees of freedom, was reported. Criteria for interpreting the Cramer's V effect size were as follows: .01 (small), .30 (medium), and .50 (large).

Independent-samples *t*-tests were performed to compare the mean scores between the two cluster groups for specified variables. Procedures outlined above (Mean score differences, 2.4.1.2) were used to conduct *t*-tests. An effect size statistic (Cohen's *d*) was calculated for each finding and criteria for interpretation were applied as follows: .20 (small), .50 (medium), and .80 (large; Cohen, 1988). To correct for an increase in Type I error, a Bonferroni correction was used based on the number of comparisons conducted ($n = 3$, $\alpha = .017$).

2.5 Missing data

Missing data represented 6% of the expected responses for items included in the analysis. Of those missing data, 41% were due to selection of the option "Don't know/NA", and 59% were left blank (no response was provided). Both types of missing data were treated equivalently, as recommended for survey measures by Enders (2010). Multiple imputation procedures were used to replace missing data using the multiple imputation software NORM (Schafer, 1999). Multiple imputation is a statistical procedure that simulates multiple versions of a datasets then combines those datasets to produce estimates of the missing data (Baraldi & Enders, 2010).

3. Results

3.1 Relation between sustainability scores and school features

3.1.1 Years implementing

The relation between sustainability score and years implementing SWPBS was investigated by correlating the variables. Preliminary analyses were performed to investigate the assumptions of normality, linearity, and homoscedasticity. In the case of the sustainability score variable, a statistically significant Shapiro-Wilk statistic ($p <$ .001), skew values greater than 1, and analysis of boxplots, Q-Q plots and histograms, revealed a negatively skewed distribution. The years implementing variable met the above assumptions.

To investigate the relations between variables, Pearson's product-moment correlation coefficients and Spearman's nonparametric correlation coefficients were calculated. After calculating each coefficient, the decision to retain or reject the null hypothesis was equivalent, and correlation coefficients were nearly identical. Therefore, for ease of interpretation, the Pearson's product-moment correlation coefficient is reported.

There was a statistically significant positive correlation between sustainability score and years implementing, r =.37, n =188, $p < .001$. The coefficient of determination (R^2) = .137, indicates that approximately 14% of the variance between sustainability score and years implementing is shared. Additionally, a large effect was indicated (d = .80).

3.1.2 Mean score differences

Three dichotomous school team action variables (meetings, data presented, and external coach) were analyzed, using independent-samples *t*-tests, for differences in sustainability scores. *T*-tests and non-parametric Mann-Whitney U tests were conducted for each analysis. For each analysis, equivalent results were found between parametric and non-parametric procedures, therefore, for ease of interpretation, results of *t*-tests are presented below.

Independent-samples *t*-tests were conducted to compare SUBSIST sustainability scores across categorical variables (see Table 3.1). Results of the test for frequency of team meetings indicated a statistically significant difference between groups (higher sustainability scores for schools meeting at least monthly), $t\ (194) = 4.91$, $p < .001$, with a large effect ($d = 0.93$). Results of the test for frequency of presenting data indicated a statistically significant difference between groups (higher sustainability scores for schools presenting data at least monthly), $t\ (191) = 4.65$, $p < .001$, with a medium effect ($d = 0.71$). Finally, results of the test for access to external support indicated a statistically significant difference between groups (higher sustainability scores for those reporting access to coaching), $t\ (201) = 3.98$, $p < .001$, with a medium effect ($d = 0.56$).

Table 3.1 Results of *t*-tests by variable.

Variable	*M (SD)*	*t*	*d*
Meetings		4.91***	0.93
Fewer than once per month	2.90 (.55)		
At least monthly	3.35 (.41)		
Data presented		4.65***	0.71
Fewer than once per month	3.13 (.52)		
At least monthly	3.44 (.34)		
External coach		3.98***	0.56
No access	3.10 (.49)		
Access	3.36 (.44)		

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

3.2 Cluster analysis

The SPSS two-step auto-clustering results used for the determination of numbers of clusters are shown in Table 3.2. The number of clusters was automatically determined using a two phase estimator. The BIC values were computed in the first phase. The maximum number of clusters is set equal to the number of clusters when the ratio

$$\frac{BIC_{cluster_k}}{BIC_{cluster_1}}$$

is for the first time smaller than c_1=0.04 (SPSS, 2010). As shown in Table 3.2, this criterion was met with three clusters. The second phase uses the ratio change in distance for k clusters, which is defined as

$$R(k) = \frac{d_{k-1}}{dk},$$

where d_{k-1} is the distance when k clusters are merged to k−1 clusters. The number of clusters is defined at the point where a remarkable shift of the ratio change occurs. As shown in Table 3.2, the two largest ratios of distance measures were obtained for the two and three cluster solutions. The ratio of R(3)/R(2) is 0.26 and smaller than the threshold value of c_2=1.15 (SPSS, 2010). Hence, the two, instead of three, cluster solution was selected as the solution that best fit the data.

Table 3.2 SPSS two-step auto-clustering results

No. of clusters	Schwartz Bayesian criterion (BIC)	BIC change[a]	Ratio of BIC changes[b]	Ratio of distance measures
1	6266			
2	5626	-640	1.00	4.15
3	5790	164	-0.26	1.08

a The change results from the previous number of clusters in the table.
b The ratio of changes is with respect to the change at the two clusters.
c The ratios of distance measures are based on the current number of clusters against the previous number of clusters.

3.3 Critical items

The predictor importance value, which indicates how well the variable can differentiate between clusters, range from 0, item contributed little to cluster formation, to 1, item contributed significantly to cluster formation. Predictor importance value (*p* –value), mean, and standard deviation for each SUBSIST item are listed in Table 3.3.

Table 3.3 Predictor importance values, means, and standard deviations of items.

SUBSIST Item	Predictor Importance	Mean	Standard Deviation
Data are used for problem solving, decision making and action planning (to make SWPBS more effective &/or efficient)	1.00	3.45	0.82
The school team implementing SWPBS is well organized and operates efficiently	0.67	3.40	0.78
School personnel perceive SWPBS as effective in helping them achieve desired outcomes	0.66	3.21	0.76
SWPBS has been expanded to other areas (e.g., classrooms, buses, students with intensive needs, parenting workshops)	0.62	3.22	0.89
SWPBS is implemented with fidelity (i.e., it is used as intended)	0.58	3.27	0.78
The school administrators describe SWPBS as a top priority for their school	0.58	3.46	0.77
SWPBS is considered to be a typical operating procedure of the school (it has become "what we do here/what we've always done)	0.58	3.40	0.83
Data are reviewed regularly at team meetings	0.56	3.42	0.88
The school team implementing SWPBS is knowledgeable and skilled in SWPBS	0.55	3.46	0.71

SUBSIST Item	Predictor Importance	Mean	Standard Deviation
SWPBS has been integrated into new school or district initiatives (e.g., renamed to meet new needs, shown how it can meet the goals of the new initiatives as well)	0.55	3.39	0.78
A vast majority of school personnel (80% or more) support SWPBS	0.55	3.50	0.71
Data are presented to all school personnel at least four times a year	0.54	3.19	1.01
There is regular measurement of student outcomes (e.g., ODRS, achievement data, school safety survey, student/parent satisfaction survey)	0.53	3.38	0.80
SWPBS is embedded in school and/or district policy (e.g., school improvement plans, mission/value statements)	0.53	3.39	0.85
There is regular measurement of fidelity of implementation (e.g.. team checklist. SET, Benchmarks of Quality)	0.51	3.18	0.98
School personnel celebrate the positive effects of SWPBS at least yearly	0.51	3.25	0.94
Needs assessments (e.g., EBS/PBS self assessment survey) are conducted	0.43	3.31	0.92
The school administrators actively support school personnel when implementing and aligning initiatives (e.g., shield staff from competing demands, change language to align SWPBS with new initiatives) to allow SWPBS to occur	0.41	3.30	0.84
SWPBS addresses outcomes that are highly valued by school personnel	0.40	3.44	0.69
SWPBS has a “crossover effect” in other areas (e.g., improved academic achievement scores, attendance)	0.40	3.09	0.87
SWPBS is effective for a large proportion of students	0.39	3.58	0.69
The school team implementing SWPBS meets at least monthly	0.38	3.54	0.88
Data are presented at least once per year to key stakeholders outside of the school (e.g., district officials, school boards, community agencies/groups	0.36	2.79	1.15

Variable	Predictor Importance	Mean	Standard Deviation
SWPBS is promoted and visible to important organizations (e.g., school boards, community agencies, businesses, parent groups)	0.32	2.81	0.94
SWPBS becomes easier to use with continued experience	0.32	3.53	0.69
School teams and new personnel are provided with professional development in SWPBS at least yearly	0.31	3.15	0.97
The district administration actively supports SWPBS (e.g., describes SWPBS as top priority, provides clear direction)	0.31	3.13	0.90
Materials related to SWPBS (e.g., handbook, posters) can be used or adapted with ease across years	0.29	3.51	0.65
All school personnel have a basic understanding of SWPBS (i.e., know the critical features and practices)	0.29	3.50	0.70
Data collected for SWPBS are easy to collect and do not interfere with teaching	0.29	3.27	0.80
The practices and strategies of SWPBS are evidence-based (i.e., there is published research documenting their effectiveness)	0.27	3.51	0.73
A school administrator regularly attends and participates in SWPBS team meetings	0.25	3.53	0.78
SWPBS serves a critical need for the school	0.20	3.59	0.64
The school team is connected to a “community of practice” (e.g., network of other SWPBS schools in district, local/regional conferences)	0.19	3.01	1.02
SWPBS is cost-effective (in terms of money and effort)	0.14	3.32	0.71
The school team has regular access to district SWPBS expertise (e.g., external/district coaches or consultants)	0.13	3.27	0.97
There are adequate district resources (funding and time) allocated for SWPBS	0.12	2.63	0.98
Parents are actively involved in the SWPBS effort (e.g., as part of SWPBS team or district committee)	0.08	2.12	0.87
State/provincial officials actively support SWPBS (e.g., promotion, publicity, providing infrastructure)	0.06	2.71	1.08

Mean response values and standard deviations are provided for the four items with predictor importance values greater than .60, given in Table 3.4. Response pattern by cluster is provided. Cluster one includes 78 schools and is characterized by schools reporting lower average ratings across sustainability items (2.55 to 2.79). This group was named *Low Sustainability Schools*. Cluster two contains 139 schools and is characterized by schools reporting higher average ratings across sustainability items (3.54 to 3.87). This group was named *High Sustainability Schools.* Mean comparisons across cluster group produced statistically significant differences ($p < .001$) for each of the four items with predictor importance values >.60; SUBSIST critical features were reported as more present for *High Sustainability Schools* than *Low Sustainability Schools*.

Table 3.4 Critical items with predictor importance statistics > .60 and cluster group comparisons.

Item	Low Sustainability Schools		High Sustainability Schools		*t*
	Mean	**Standard Deviation**	**Mean**	**Standard Deviation**	
Data are used for problem solving, decision making and action planning (to make SWPBS more effective &/or efficient)	2.71	0.88	3.87	0.34	13.83***
The school team implementing SWPBS is well organized and operates efficiently	2.79	0.86	3.74	0.46	10.61***
School personnel perceive SWPBS as effective in helping them achieve desired outcomes	2.62	0.69	3.54	0.58	10.46***

Item	Low Sustainability Schools		High Sustainability Schools		*t*
	Mean	Standard Deviation	Mean	Standard Deviation	
SWPBS has been expanded to other areas (e.g., classrooms, buses, students with intensive needs, parenting workshops)	2.55	0.94	3.60	0.59	10.10***

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

3.4 Criterion analyses

Internal criterion analysis of the two cluster solution was assessed by examining the structure silhouette of cohesion and separation (Kaufman & Rousseeuw, 1990). The coefficient ranges between -1, suggesting a very poor model, and 1, suggesting an excellent model. SPSS (v.19) categorizes models as poor (< .2, data does not exhibit adequate structure), fair (.2 - .5, reasonable partitioning of data), and good (> .5 excellent partitioning of data) based on how closely related cases belonging to the same cluster are, in comparison to their relation to cases belonging to other clusters. The two cluster solution in this study had a coefficient of .3, a fair rating. This result suggests that although there was a moderate degree of relation between cases belonging to opposing clusters, there was satisfactory distance between them to infer that there are substantial differences between cluster groups.

In addition to internal criterion analysis, tests of external validity were conducted. Demographic variables and school team actions that were not used in the cluster analysis

were employed to determine whether meaningful differences existed between the two clusters for these variables. External variables are presented by cluster in Table 3.5.

Table 3.5 Proportions of variables by clusters.

Variable	Low Sustainability Schools	High Sustainability Schools
Urbanicity		
Rural/Semirural	14.3%	20.0%
Suburban	34.9%	33.0%
Sm/Med/Large City	50.8%	47.0%
Grades served		
Elementary school	66.7%	65.5%
K-8 & K-12	6.4%	5.0%
Middle school	19.2%	26.6%
High school	7.7%	2.9%
Title I eligible		
Yes	62.3%	69.6%
No	37.7%	30.4%
Meeting frequency		
At least one meeting per month	75.0%	93.8%
Fewer than one meeting per month	25.0%	6.3%

Variable	Low Sustainability Schools	High Sustainability Schools
Data presented		
At least one meeting per month	27.5%	52.4%
Fewer than one meeting per month	72.5%	47.6%
External coach		
Yes	49.3%	64.0%
No	50.7%	36.0%

3.4.1 Chi-square tests of independence

Chi-square tests of independence were performed to determine if the clusters varied significantly by categorical dependent variable (see Tables 3.6 to 3.11). Results of the test for urbanicity (see Table 3.6) showed no statistically significant proportions between *Low Sustainability Schools* and *High Sustainability Schools*, χ^2 (2, n = 178) = 0.91, p = .64, with a small effect (Cramer's V = .07). Results of the test for grades served (see Table 3.7) were not significant, χ^2 (3, n = 217) = 3.83, p = .28, with a small effect (Cramer's V = .13). Results of the test for Title I eligibility (with Yates Continuity Correction; see Table 3.8) were also not significant, χ^2 (1, n = 212) = 0.87, p = .35, with a small effect (Φ = -.08).

Results of the test for frequency of team meetings (with Yates Continuity Correction; see Table 3.9) indicated a statistically significant difference between clusters,

with a higher proportion of *High Sustainability Schools* meeting at least monthly, χ^2 (1, n = 196) = 12.4, $p < .001$, with a borderline medium effect (Φ = -.27). Likewise, results of the test for frequency of data presented to staff (with Yates Continuity Correction; see Table 3.10) indicated a statistically significant difference between clusters (higher proportion of *High Sustainability* schools presenting data at least monthly), χ^2 (1, n = 193) = 10.2, p = .001, with a small to medium effect (Φ = -.24). Finally, results of the test for access to external support (see Table 3.11) indicated a statistically significant difference between clusters (more *High Sustainability Schools* reporting access to coaching), χ^2 (1, n = 203) = 9.76, p = .002, with a small to medium effect (Φ = -.23).

Table 3.6 Proportions of clusters by urbanicity.

	Location				
Cluster	Rural/Semirural	Suburban	Small/Med/Large City	χ^2	Cramer's V
Low Sustainability Schools	9 (14.3%)	22 (34.9%)	32 (50.8%)	0.91	.07
High Sustainability Schools	23 (20.0%)	38 (33.0%)	54 (47.0%)		

Note. n = 212. Percent observed within cluster appear in parentheses below group frequencies. * $p < .05$.

Table 3.7 Proportions of clusters by grades served.

	Grades Served					
Cluster	Elementary	K-8, K-12	Middle School	High School	χ^2	Cramer's V
Low Sustainability Schools	52 (66.7%)	5 (6.4%)	15 (19.2%)	6 (7.7%)	3.83	.13
High Sustainability Schools	91 (65.5%)	7 (5.0%)	37 (26.6%)	4 (2.9%)		

Note. n = 217. Percent observed within cluster appear in parenthesis below group frequencies. * $p < .05$.

Table 3.8 Proportions of clusters by Title I eligibility.

	Title 1 Eligible			
Cluster	Yes	No	χ^2	Φ
Low Sustainability Schools	48 (62.3%)	29 (37.7%)	0.87	-.08
High Sustainability Schools	94 (69.6%)	41 (30.4%)		

Note. n = 212. Percent observed within cluster appear in parenthesis below group frequencies. * $p < .05$.

Table 3.9 Proportions of clusters by meeting frequency.

	Frequency of SWPBS School team Meetings			
Cluster	At least one meeting per month	Fewer than one meeting per month	χ^2	Φ
Low Sustainability Schools	51 (75.0%)	17 (25.0%)	12.4***	-.27
High Sustainability Schools	120 (93.8%)	8 (6.3%)		

Note. n = 196. Percent observed within cluster appear in parenthesis below group frequencies. *p < .05. **p < .01. ***p < .001.

Table 3.10 Proportions of clusters by frequency of data presentation.

	Frequency SWPBS data is presented to all school personnel			
Cluster	At least one time per month	Fewer than one time per month	χ^2	Φ
Low Sustainability Schools	19 (27.5%)	50 (72.5%)	10.2**	-.24
High Sustainability Schools	65 (52.4%)	59 (47.6%)		

Note. n = 193. Percent observed within cluster appear in parenthesis below group frequencies. *p < .05. **p < .01.

Table 3.11 Proportions of clusters by access to external coach.

	Access to at least one hour of SWPBS external support per week			
Cluster	Yes	No	χ^2	Φ
Low Sustainability Schools	36 (49.3%)	37 (50.7%)	9.76**	-.23
High Sustainability Schools	94 (64.0%)	36 (36.0%)		

Note. n = 203. Percent observed within cluster appear in parenthesis below group frequencies. *p < .05. **p < .01.

3.4.2 Independent-samples t-tests

Three demographic variables (free or reduced lunch, enrollment, and years implementing) were analyzed for differences by cluster group using the appropriate comparison procedures. A comparison of the mean percent of students receiving free or reduced priced lunch in *High Sustainability Schools* and *Low Sustainability Schools* was used to determine whether a statistically significant difference existed between the groups. A non-significant Shapiro-Wilk statistic (p = .200) suggested that the assumption of normality had not been violated. Analyses of the skew, kurtosis, and box-plots further indicated a normal distribution of the free or reduced lunch variable. Additionally, an examination of Levene's homogeneity of variance revealed that the assumption had been met. Based on meeting the assumptions of normality and homogeneity of variances, an independent-samples *t*-test was used. There was no statistically significant difference in

the percent of students receiving free or reduced priced lunch between cluster groups, t (210) = 0.78, p = .44. Additionally, the Cohen's d (.11) indicated a small effect.

Table 3.12 Mean percent of students receiving free or reduced priced lunch.

Cluster	*M*	*SD*	*t*
Low Sustainability Schools (n = 76)	51.4%	28.9%	0.78
High Sustainability Schools (n = 136)	48.4%	26.0%	

Note. *p < .05.

A comparison of mean enrollment numbers in *High Sustainability Schools* and *Low Sustainability School* was used to determine whether a statistically significant difference existed between the groups. A statistically significant Shapiro-Wilk statistic (p < .001) suggested that the assumption of normality had been violated. Analyses of the skew, kurtosis, and box-plots further indicated that enrollment was not normally distributed. A non-significant Levene's test suggested equality of variances between the two groups' distributions. In response to a violation of normality, the examiner conducted an independent-samples t-test and a Mann-Whitney U Test; the latter of which is more robust to violations of the assumption of normality. Both procedures produced results that were not statistically significant, therefore, the parametric analysis is reported for ease of interpretation.

An independent-samples *t*-test revealed no statistically significant difference in the mean number of students enrolled in *High Sustainability Schools* and *Low Sustainability Schools*, t (210) = 0.07, p = .94. Additionally, a Cohen's d of .01 indicates that there was a negligible difference between clusters.

Table 3.13 Mean number of students enrolled.

Cluster	*M*	*SD*	*t*
Low Sustainability Schools (n = 76)	575	328	0.07
High Sustainability Schools (n = 136)	572	344	

Note. * p < .05.

A comparison of the mean years *High Sustainability* and *Low Sustainability* schools have been implementing SWPBS was used to determine whether a statistically significant difference existed between the groups. A statistically significant Shapiro-Wilk statistic (p < .001) suggested that the assumption of normality may have been violated. However, analyses of the skew, kurtosis, and box-plots indicated a normally distributed variable. An examination of Levene's homogeneity of variance for both samples revealed that the assumption of homogeneity had been met. An independent-samples *t*-test was conducted to compare the mean numbers of years implementing SWPBS between clusters. There was a statistically significant difference in mean years implementing between clusters, t (186) = -4.94, p < .001. Additionally, the Cohen's d (.76) indicated a medium to large effect.

Table 3.14 Mean years of SWPBS implementation by cluster.

Cluster	*M*	*SD*	*t*
Low Sustainability Schools (*n* = 70)	3.99	2.73	-4.94***
High Sustainability Schools (*n* = 118)	6.23	3.16	

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

4. Discussion

4.1 Overview

The purpose of the present study was to obtain additional data to evaluate the validity of the SUBSIST measure for measuring sustainability of school-based interventions. Individuals with knowledge of their school's SWPBS systems completed the measure; in total, 217 schools participated in the study. The aims of the research questions posed were to determine whether the measure was capable of differentiating between different types of schools based on overall sustainability score and responses to SUBSIST items. Differentiation between different types of schools provides evidence of construct validity for the measure.

Associations between sustainability scores and school features were considered by performing Pearson's correlations and independent-samples *t*-tests. Analyses supported that higher sustainability scores were associated with higher ratings of school team actions (i.e., more frequent school team meetings, more frequent data sharing with school staff, and access to consultation), and more years implementing SWPBS. Significantly higher sustainability scores were found for schools that reported having school teams that met and presented data to school staff at least once per month, as well as school reporting access to an external coach or consultant. Additionally, years sustaining SWPBS was significantly correlated with sustainability score.

A two-step cluster analysis was conducted to determine whether schools could be classified into subgroups based on responses to the SUBSIST items. Internal and external

criterion analyses of the cluster solution were used to determine whether the solution was valid. Results of the analysis supported a two cluster solution, with *High Sustainability Schools* reporting significantly higher SUBSIST scores across items. Internal criterion analysis of the cluster solution produced a "fair" rating, suggesting reasonable partitioning of data but that caution should be applied interpreting the results. Results of external criterion analyses were consistent with sustainability score analyses; specifically *High Sustainability Schools* had significantly higher reports of school team actions (i.e., meetings, data present, external consultant) and years sustaining SWPBS than *Low Sustainability Schools*. Further analyses revealed that statistically significant differences in demographics did not exist between schools across cluster groups.

The results of this study provide further evidence of construct validity for the SUBSIST measure to evaluate schools' capacity to sustain SWPBS efforts at the universal tier. Specifically, findings support that SUBSIST sustainability scores are associated with features predictive of sustained implementation of SWPBS, and that SUBSIST scores are capable of differentiating between different types of schools.

4.2 Construct validity of the SUBSIST

By demonstrating associations between the SUBSIST sustainability score and features that are predictive of sustained outcomes, it can be inferred that the SUBSIST is capable of differentiating between schools that are more or less likely to sustain based on overall sustainability score. In addition to number of years implementing SWPBS, features

considered were selected based on prior research that documents that use of data for decision-making, prioritizing SWPBS, and having active school teams, are all associated with a greater chance of achieving sustained implementation (Coffey & Horner, in press; McIntosh et al., 2011).

Findings that SUBSIST sustainability scores are strongly associated with all four predictive indicators of sustainable outcomes provide evidence that the sustainability score can be considered a valid indicator of the construct of sustainability that the measure aims to assess. Although the three school team action variables considered in the analyses were directly assessed by items on the SUBSIST, each variable was only one item in the 39-item overall sustainability score. The results indicate that these variables seem to be pivotal to sustainability and related strongly to other sustainability features measured by the SUBSIST.

Further evidence of construct validity for the SUBSIST was provided by findings from the cluster analysis. It was hypothesized that the SUBSIST should be capable of differentiating between different types of schools (producing cluster groups) based on responses to specific items on the SUBSIST. The cluster analysis produced two specific clusters of schools. The distinguishing feature between the clusters was the extent to which schools reported having multiple features predictive of sustained practice in place. Items that had a critical role in determining clusters (e.g., use of data for decision making, efficiency of school teams, priority of SWPBS, and staff/administrator support) have been cited in previous literature as playing a role in sustainability of SWPBS (Bambara et al., 2009; Coffey & Horner, in press; Taylor-Greene & Kartoub, 2000). These same

factors were also hypothesized by McIntosh et al. (2009) as contributing to a model conceptualizing sustainability.

Results of the structure silhouette measure of cohesion and separation, the internal criterion analysis of the cluster solution, suggested that there was a moderate degree of relation between cases belonging to opposing clusters; however, there was satisfactory distance between the two clusters to infer that differences exist between cluster groups based on their responses to SUBSIST items. Results of the external criterion analysis of the cluster solution provided tentative evidence that differences existed between cluster groups. There were statistically significant differences found between cluster groups when evaluating factors predictive of sustained outcomes, including all three school team features and years implementing SWPBS. These results provide evidence that there are differences between cluster groups in regards to the schools' potential to sustain SWPBS. Interestingly, there were no differences found between cluster groups in regards to school demographics. A review of the literature considering the impact of school demographics relation to achieving sustained school practices shows some conflicting evidence of the impact of demographics (Pas, 2011; Simmons et al., 2002). Therefore, the non-significant findings of the demographic analyses do not provide definitive evidence for or against the validity of the cluster solution.

4.3 Processes underlying sustainability

This study also contributes to our understanding of the context of sustainability and offers insights into how the process of sustainability occurs in schools. First, it is important to make note of the high degree of relatedness between the 39 sustainability items, as evidenced by strong inter-correlation values between items. Moreover, the cluster analysis showed that the *High Sustainability Schools* had higher mean scores in 100% of SUBSIST items, representing a range of variables enhancing sustainability. Because items are likely to affect each other, the presence of some features likely effects the presence of others. This item cohesion lends support to the notion that the construct of sustainability is additive, meaning that the more environmental features that are in place, the more likely sustainability is to occur.

A second interesting finding is the importance of data to sustainability. The most significant contributor to the cluster analysis classification scheme was responses to the item "Data are used for problem solving, decision making and action planning (to make SWPBS more effective &/or efficient)." Use of data has frequently been cited in the literature as being a significant contributor to sustained SWPBS practices, and current research exploring factors predictive of sustainability of SWPBS systems suggest that school team functioning, most notably school team use of data for decision making, is the largest contributing factor to sustained implementation (Coffey & Horner, in press; Fullan, 2005; McIntosh et al., 2011). The result of this cluster analysis supports this finding, providing additional support for use of data as a critical mechanism in sustainability.

Finally, it is an interesting and encouraging finding that SUBSIST scores varied across schools based on responses to malleable, contextually-based variables, as opposed to less malleable demographic factors. Results of the external criterion analysis support Biglan's (2004) suggestion to adopt a contextually-based framework when studying human behaviour. When conceptualizing sustainability using a contextually-based framework, the core belief is that the construct can be altered by making changes to malleable variables that exist within the environment. This finding is promising because it suggests that schools have more control over their outcomes. If schools prioritize SWPBS and implement critical features related to sustainability, school demographics do not seem to have a restrictive effect on sustaining SWPBS.

4.4 Limitations and future research

The results of this study provide valuable contributions to the validation of the SUBSIST measure and adds to the growing research on sustained implementation of school-based practices; however, there are several limitations that must be considered. A central limitation of this study is that school team features used as predictive indicators of sustainability (features that were considered in relation to sustainability score and provided external criterion analysis of the cluster analysis solution) were assessed to some extent by items on the SUBSIST measure. Ideally, measures should be validated by relations to variables that are not included in the measure itself, and future analysis should consider other predictors of sustained implementation.

Another limiting aspect of this study was the fair internal criterion analysis rating for the cluster solution. Support for the cluster solution would be stronger had internal analysis shown greater separation between factors, and as a result, some caution should be used when interpreting the clusters.

Another limitation of the study was that the sample was overrepresented by elementary schools serving a high-needs student population (large representation of Title I schools, schools with high proportions of students receiving free or reduced priced lunch). A need for further investigation of sustainability in schools serving different student populations is needed to determine if results generalize across school and student populations.

Another limiting aspect of this study was that variables used for the external criterion analysis of the cluster solution were restricted to questions included in the SUBSIST survey or available from the NCES website. Additional variables that would be of interest to consider, in light of the findings of this and other studies, would be those related to the design and functioning of schools SWPBS systems; specifically, variables related to use of data for decision making and access to communities of practice. Specific questions that could be considered might include, whether schools have access to software or database systems to monitor and track school data, whether schools are connected to a community of practice at the district or statewide level, and whether communities of practice offer tangible resources (e.g., funding, FTE personnel, trainings) to support schools sustaining SWPBS, and if so, what kinds and how much are provided.

These limitations are important considerations and should be addressed in future research.

There is a need for replications of these findings to validate whether SUBSIST items are related to sustainability of SWPBS systems in practice. Particularly, long-term studies investigating the contextual features of school environments, particularly factors contributing to continuous regeneration of practices, are needed to determine specific practices that lead to sustained implementation.

4.5 Implications for practice

Findings from this study suggest that the SUBSIST could be used as a research or action-planning tool to assess whether school environments have features in place that promote sustainability. Conceptualizing the construct of sustainability through an additive framework would encourage the use of the sustainability score as a general indicator of potential to sustain SWPBS practices, as the score factors in the presence of a wide variety of features. However, findings also indicate that a particular emphasis should be placed on the use of data for decision making and prioritizing school team actions, as these features have strong associations with sustained implementation of SWPBS.

Another promising finding for schools is that sustainability was not shown to be significantly influenced by particular school demographics, including factors related to the student population being served, size, or urbanicity. In essence, results indicate that

any type of school can sustain SWPBS, as long as the school team attends to the critical features measured by the SUBSIST.

4.6 Conclusions

The results of this study contribute to growing research on sustainability of SWPBS and adds further validity for the use of the SUBSIST measure for assessing variables related to sustained implementation. Results suggest that the SUBSIST sustainability score is associated with features predictive of sustained implementation of SWPBS and that the measure is able to distinguish between schools that are more or less likely to sustain SWPBS based on responses to SUBSIST items. Both of these findings provide construct validity for the measure, suggesting that inferences made from the SUBSIST in regards to prediction of sustainable practices are useful, meaningful, and appropriate.

References

Adelman, H. S., & Taylor, L. (2003). On sustainability of project innovations as systemic change. *Journal of Educational and Psychological Consultation, 14*, 1-25.

Bambara, L., Nonnemacher, S., & Kern, L. (2009). Sustaining school-based individualized Positive Behavior Support: Perceived barriers and enablers. *Journal of Positive Behavior Interventions, 11*, 161-178.

Baraldi, A. N., & Enders, C. K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology, 48*, 5-37.

Biglan, A. (2004). Contextualism and the development of effective prevention practices. *Prevention Science, 5*, 15-21.

Chiu, T., Fang, D., Chen, Y., Wang, Y., & Jeris, C. (2001). *A robust and scalable clustering algorithm for mixed type attributes in large database environments.* Paper presented at the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher, 32*, 3-12.

Coffey, J., & Horner, R. H. (in press). The sustainability of school-wide positive behavioral interventions and supports. *Exceptional Children*.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, R. J., & Swerdlik, M. E. (2009). *Psychological testing and assessment: An introduction to tests and measurement* (7th ed.). New York, NY: McGraw-Hill Higher Education.

Curtis, M. J., Castillo, J. M., & Cohen, R. M. (2008). Best practices in system-level change. In A. Thomas & J. P. Grimes (Eds.), *Best practices in school psychology V* (pp. 887-901). Bethesda, MD: National Association of School Psychologists.

Dillman, D. A. (1983). Mail and other self-administered questionnaires. In P. Rossi, J. D. Wright & A. B. Anderson (Eds.), *The handbook of survey research* (pp. 359-377). New York: Academic Press.

Dubes, R., & Jain, A. K. (1979). Validity studies in clustering methodologies. *Pattern Recognition, 11*, 235-254.

Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford.

Filter, K. J., McKenna, M. K., Benedict, E. A., Horner, R. H., Todd, A. W., & Watson, J. (2007). Check in/Check out: A post-hoc evaluation of an efficient, secondary-level targeted intervention for reducing problem behaviors in schools. *Education and Treatment of Children, 30*, 69-84.

Foorman, B. R., & Moats, L. C. (2004). Conditions for sustaining evidence-based practices in early reading instruction. *Remedial & Special Education, 25*, 51-60.

Fuchs, L. S., & Fuchs, D. (1998). Treatment validity: A unifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research and Practice, 13*, 204-219.

Fullan, M. (2005). *Leadership and sustainability*. Thousand Oaks, CA: Corwin Press.

Gersten, R., Chard, D. J., & Baker, S. (2000). Factors enhancing sustained use of research-based instructional practices. *Journal of Learning Disabilities, 33*, 445-457.

Gravetter, F. J., & Wallnau, L. B. (2004). *Statistics for the behavioral sciences* (6th ed.). Belmont, CA: Wadsworth.

Han, S. S., & Weiss, B. (2005). Sustainability of teacher implementation of school-based mental health programs. *Journal of Abnormal Child Psychology, 33*, 665-679.

Horner, R. H., Sugai, G., Smolkowski, K., Eber, L., Nakasato, J., Todd, A. W. (2009). A randomized, wait-list controlled effectiveness trial assessing school-wide positive behavior support in elementary schools. *Journal of Positive Behavior Interventions, 11*, 133-144.

Kaufman, L., & Rousseeuw, P. J. (1990). *Finding groups in data: An introduction to cluster analysis*. New York: John Wiley and Sons.

Kincaid, D., Childs, K., Blase, K. A., & Wallace, F. (2007). Identifying barriers and facilitators in implementing schoolwide positive behavior support. *Journal of Positive Behavior Interventions, 9*, 174-184.

Latham, G. (1988). The birth and death cycles of educational innovations. *Principal, 68*, 41-43.

Lohrmann, S., Forman, S., Martin, S., & Palmieri, M. (2008). Understanding school personnel's resistance to adopting schoolwide positive behavior support at a universal level of intervention. *Journal of Positive Behavior Interventions, 10*, 256-269.

Lorr, M. (1983). *Cluster analysis for social scientists*. San Francisco, CA: Jossey-Bass Publishers.

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

McIntosh, K., Filter, K. J., Bennett, J. L., Ryan, C., & Sugai, G. (2010). Principles of sustainable prevention: Designing scale-up of school-wide positive behavior support to promote durable systems. *Psychology in the Schools, 47*, 5-21.

McIntosh, K., Horner, R. H., & Sugai, G. (2009). Sustainability of systems-level evidence-based practices in schools: Current knowledge and future directions. In W. Sailor, G. Dunlap, G. Sugai & R. H. Horner (Eds.), *Handbook of positive behavior support* (pp. 327-352). New York: Springer.

McIntosh, K., MacKay, L. D., Hume, A. E., Doolittle, J., Vincent, C. G., Horner, R. H. (in press). Development and validation of a measure to assess factors related to sustainability of school-wide positive behavior support. *Journal of Positive Behavior Interventions*.

McIntosh, K., Mercer, S. H., Hume, A. E., Frank, J. L., & Turri, M. G. (2011). Factors predicting sustained implementation of school-wide positive behavior interventions and supports. *Manuscript in preparation*.

McLaughlin, M. W., & Mitra, D. (2001). Theory-based change and change-based theory: Going deeper, going broader. *Journal of Educational Change, 2*, 301-323.

Milligan, G. W. (1996). Clustering validation: Results and implications for applied analysis. In P. Arabie, L. J. Hubert & G. De Soete (Eds.), *Clustering and classification* (pp. 341-375). River Edge, NJ: World Scientific Publishing Co.

Norusis, M. J. (2010). *PASW statistics 18 statistical procedures companion*. Upper Saddle River, NJ: Prentice Hall.

Pas, E. T. (2011, February). *State-wide scale-up of SW-PBIS in Maryland elementary schools: School and district factors related to adoption and implementation.* Paper presented at the National Association of School Psychologists, San Francisco, CA.

Santangelo, T. (2009). Collaborative problem solving effectively implemented, but not sustained: A case study for aligning the sun, the moon, and the stars. *Exceptional Children, 75*, 185-209.

Sargent, T., & Hannum, E. (2005). Keeping teachers happy: Job satisfaction among primary school teachers in rural northwest China. *Comparative education review, 49*, 173-204.

Schafer, J. L. (1999). *NORM: Multiple imputation of incomplete multivariate data under a normal model, version 2.* Software for Windows 95/98/NT, available from http://www.stat.psu.edu/~jls/misoftwa.html.

Schaughency, E., & Ervin, R. A. (2006). Building capacity to implement and sustain effective practices to better serve children. *School Psychology Review, 35*, 155-166.

Simmons, D. C., Kame'enui, E. J., Good, R. H., Harn, B. A., Cole, C., & Braun, D. (2002). Building, implementing, and sustaining a beginning reading model: Lessons learned school by school. In M. R. Shinn, H. M. Walker & G. Stoner (Eds.), *Interventions for academic and behavioral problems II: Preventive and remedial approaches* (pp. 403-432). Bethesda, MD: National Association of School Psychologists.

SPSS. (2004). Two step cluster analysis. Chicago.

SPSS. (2010). IBM SPSS statistics 19 algorithms. Chicago.

Sugai, G., & Horner, R. H. (2006). A promising approach for expanding and sustaining the implementation of school-wide positive behavior support. *School Psychology Review, 35*, 245-259.

Sugai, G., & Horner, R. H. (2009). Defining and describing schoolwide positive behavior support. In W. Sailor, G. Dunlap, G. Sugai & R. H. Horner (Eds.), *Handbook of positive behavior support* (pp. 307-326). New York: Springer.

Sugai, G., Horner, R. H., Dunlap, G., Hieneman, M., Lewis, T. J., Nelson, C. M. (2000). Applying positive behavior support and functional behavioral assessment in schools. *Journal of Positive Behavior Interventions, 2*, 131-143.

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 at http://www.pbis.org.

Taylor-Greene, S., Brown, D., Nelson, L., Longton, J., Gassman, T., Cohen, J. (1997). School-wide behavioral support: Starting the year off right. *Journal of Behavioral Education, 7*, 99-112.

Taylor-Greene, S., & Kartoub, D. T. (2000). Durable implementation of school-wide behavior support: The High Five Program. *Journal of Positive Behavior Interventions, 2*, 233-235.

Tyack, D. B., & Cuban, L. (1995). *Tinkering toward utopia: A century of public school reform*. Cambridge, MA: Harvard University Press.

Valli, L., & Buese, D. (2007). The changing roles of teachers in an era of high-stakes accountability. *American Educational Research Journal, 44*, 519-588.

Vaughn, S., Klingner, J., & Hughes, M. (2000). Sustainability of research-based practices. *Exceptional Children, 66*, 163-171.

Walker, H. M. (2004). Commentary: Use of evidence-based intervention in schools: Where we've been, where we are, and where we need to go. *School Psychology Review, 33*, 398-407.