



# Psychometric Properties of the *Bully-Free Schools Survey*

Prepared for:

Strategic Alternatives in Prevention Education  
(SAPE)



By:

Jim O'Neill, PhD and Angela O'Neill, MS, CHES  
O'Neill Consulting • 26641 Wall Street • Novi, MI 48374  
Voice: (734) 272-9002 • Fax: (248) 773-8326  
Web: [www.oneillconsult.net](http://www.oneillconsult.net)



August, 2011

## Table of Contents

<b>Executive Summary</b> .....	<b>2</b>
<b>Overview of BFSS</b> .....	<b>3</b>
<b>Methodology</b> .....	<b>4</b>
Design.....	4
Sample.....	5
Procedure .....	5
<b>Results</b> .....	<b>5</b>
Item Development and Refinement .....	6
Content Validity Analysis.....	6
Reliability and Construct Validity Analyses .....	6
School Staff Feedback .....	9
<b>Conclusions, Limitations, and Recommendations</b> .....	<b>9</b>
<b>References</b> .....	<b>9</b>
<b>Exhibits</b>	
Exhibit 1: BFSS Dimensions, Descriptions, and Item Number(s).....	3
EXHIBIT 2: BFSS Psychometric Properties Tested.....	4
EXHIBIT 3: Characteristics of the Samples.....	5
EXHIBIT 4: Test-Retest and Internal Consistency Reliability Results.....	6
EXHIBIT 5: Correlation between Scores on BFSS Target of Bullying/Aggression, Reduced Aggression/ Victimization Scale (Orpinas & Horne, 2006), and Rosenberg Self-Esteem Scale (1965).....	7
EXHIBIT 6: Mean BFSS PHYSICAL Aggression Target Scores by Gender, Grade Level, Suspension History, and Grades Earned.....	7
EXHIBIT 7: Mean BFSS RELATIONAL Aggression Target Scores by Gender, Grade Level, Suspension History, and Grades Earned.....	7
EXHIBIT 8: BFSS STAFF Version: Teacher Comments.....	8
<b>Appendices</b> .....	<b>10</b>
Appendix A: BFSS Instruments	
Appendix B: Teacher Comments	

The contents of this report were developed under a grant from the Department of Education. However, those contents do not necessarily represent the policy of the Department of Education, and you should not assume endorsement by the Federal Government.

**Suggested Reference:**

O’Neill, J.M., & O’Neill, A.K. (2011). *Psychometric Properties of the Bully-Free Schools Survey (BFSS)*. Novi, MI: O’Neill Consulting.

## Executive Summary

The Bully-Free Schools Survey (BFSS), a building-level needs assessment tool developed by Dee Lindenberger and colleagues of Strategic Alternatives in Prevention Education (SAPE), was revised in 2010 to incorporate new constructs and best practices in measuring bullying and other forms of aggression.

As part of a federally funded project of the Michigan Department of Education to provide no-cost, online needs assessment tools to all school districts in the State, the reliability and validity of the BFSS was tested in a sample of 1155 Michigan students in grades 5, 7, 9, and 11 from 23 buildings; and 169 Michigan school staff from 24 buildings. The methods used for the study were consistent with conventional strategies and standards for survey development.

Prior to testing, the BFSS was refined via **content validity analysis** by a panel of researchers and practitioners with expertise in school-based bullying/aggression. Panelists' ratings of the relevance of BFSS items and representativeness of BFSS dimensions were typically very high. The majority of their feedback was used to improve the readability of BFSS items and the representativeness of BFSS dimensions.

Results from testing student and staff samples showed that the **internal consistency and test-retest reliability** estimates for the BFSS were very high for each grade level surveyed, suggesting that the BFSS is a reliable instrument.

Several types of **construct validity** were tested empirically: convergent, discriminant, and concurrent. As predicted, the results showed that the BFSS (a) correlated with another measure of bullying/aggression (convergent validity); (b) did not correlate highly with a measure of self-esteem, a different construct (discriminant validity); and (c) distinguished between several subgroups known to exhibit different levels of bullying/aggression (concurrent validity).

As shown in the following summary of results, the BFSS had acceptable levels of reliability and validity for all six psychometric properties tested, and exceeded the minimal level for all but one property. The most common rating was "exemplary," the highest level for survey instruments (Robinson, Shaver, & Wrightsman, 1991). In addition, the BFSS appears to be superior to most other bullying/aggression survey tools available (Hamburger, Basile, & Vivolo, 2011), either in its breadth of dimensions measured and/or its level of reliability and validity.

BFSS psychometric property tested	BFSS reliability/ validity rating*	Evidence
<b>Reliability</b>		
Test-retest (stability)	<b>Moderate</b> ★★	BFSS score correlations (r) across a one-month period ranged from .28 to .44 for all dimensions.
Internal consistency	<b>Exemplary</b> ★★★★	Alpha coefficients for each BFSS multi-item scale were usually .80 or better, with all scales reaching at least .70
<b>Validity</b>		
Content	<b>Exemplary</b> ★★★★	Content experts rated 97% of BFSS items as "necessary" and 90% of BFSS dimensions as "completely representative." Most feedback from experts was used to improve the readability of BFSS items and representativeness of BFSS dimensions.
Construct: Convergent	<b>Minimal</b> ★	Students' scores on BFSS Target of Bullying/Aggression were highly correlated with their scores on the Reduced Aggression/Victimization Scale (developed by Orpinas & Horne, 2006).
Construct: Discriminant	<b>Moderate</b> ★★	Students' scores on a measure of self-esteem, the Rosenberg Self-Esteem Scale (Rosenberg, 1965), which is a different construct than bullying/aggression, were not highly correlated with their scores on BFSS Target of Bullying/Aggression.
Construct: Concurrent	<b>Exemplary</b> ★★★★	Students' scores on the BFSS Target of Bullying/Aggression were higher for students in younger grade levels, with a history of suspension, and who reported low school achievement. In addition, males were more likely to be a target of students' physical bullying/aggression, whereas females were more often the target of students' relational aggression.

\*Range of acceptable levels of reliability and validity: Minimal, Moderate, Extensive, and Exemplary (Robinson, et al., 1991).

Feedback from teachers who completed the BFSS and/or administered the BFSS to their students was mostly positive. However, the majority of fifth-grade teachers felt the comprehension level for their students was too difficult and the survey length was too long, suggesting the need to modify the BFSS for use in upper elementary grades.

###

The purpose of this report is to provide information about the psychometric properties (i.e., reliability and validity) of the revised Bully-Free Schools Survey (BFSS). This project is part of an effort to provide the BFSS as a no-cost, online needs assessment tool to all school districts in Michigan, which is funded by the Safe and Supportive Schools (S3) grant awarded by the U.S. Department of Education to the Michigan Department of Education.

## Overview of BFSS

The Bully-Free Schools Survey (BFSS) was developed in 2005 by Dee Lindenberger and colleagues of Strategic Alternatives in Prevention Education (SAPE), a group of regional health consultants throughout Michigan who assist schools with prevention and health promotion initiatives. It is designed to provide building-level needs assessment information to elementary and secondary schools as they work to create environments that reduce bullying and other forms of aggression. These problem behaviors have been shown to adversely impact learning (Glew et al., 2005) and academic achievement (Lacey & Cornell, 2011).

The BFSS was revised in 2010 to incorporate new constructs and best practices in assessing bullying/aggression, including the work of Stan Davis, who created the *Stop Bullying Now* intervention and collaborates on the *Youth Voice Project* with Dr. Charisse Nixon (from The Pennsylvania State University – Erie).

The BFSS is completed by school staff and students in grades 5 through 12. The school staff questionnaire consists of eight dimensions; the student version has nine. See Exhibit 1 for BFSS dimensions and descriptions. A copy of each version is in Appendix A.

### EXHIBIT 1: BFSS Dimensions, Descriptions, and Item Number(s) (See Appendix A for a copy of each version)

	BFSS Dimension	Description	STUDENT BFSS Item(s)	STAFF BFSS Item(s)
Prevention of Bullying/Aggression at School	1. School Climate to Prevent Bullying/Aggression	Student and staff perceptions of student connectedness, respect, structure, and trust at school.	1; 2a – 2f (7 items)	1a – 1f (6 items)
	2. Classroom-Based Approaches to Prevent Bullying/Aggression	Perceptions of the frequency of teachers' offerings of and students' participation in classroom-based lessons, activities, or discussions in the past 30 days that were designed to prevent or reduce bullying/aggression.	3a – 3h (8 items) (student participation)	2a – 2h (8 items) (teacher offerings)
	3. Behaviors to Stop Bullying/Aggression	Perceptions of the type and frequency of pro-social behaviors by students and staff in the past 30 days to stop bullying/aggression at school.	7a – 7k (11 items)	5a – 5h (8 items)
Bullying/Aggression at School	4. Witness of Student-to-Student Bullying/Aggression	Student and staff respondents' perceptions of the type and frequency of student-to-student bullying/aggression they witnessed at school in the past 30 days.	4a – 4o; 5a – 5j (25 items)	3a – 3o; 4a – 4j (25 items)
	5. Witness of Staff-to-Student Bullying/Aggression	Student and staff respondents' perceptions of the type and frequency of staff-to-student bullying/aggression they witnessed at school in the past 30 days.	8a – 8j; 9a – 9j (19 items)	6a – 6j; 7a – 7j (19 items)
	6. Target of Student-to-Student Bullying/Aggression	Student respondents' perceptions of the type, frequency, location, circumstance, and psychological impact of bullying/aggression they experienced from other students in the past 30 days.	10a – 10o; 11a – 11j; 12 - 14 (28 items)	N/A
Reactions to Bullying/Aggression at School	7. Reactions to Student-to-Student Bullying/Aggression	Student respondents' perceptions of their own, other students', and adults' reaction to being a target of bullying/aggression at school in the past 30 days.	15a – 15j; 16a – 16l; 17a – 17j (32 items)	N/A
	8. Staff Intolerance of Staff-to-Student Bullying/Aggression	Staff respondents' perceptions of the acceptability and consequences of staff-to-student bullying/aggression at school.	N/A	8 – 9 (2 items)
Suggestions to Stop Bullying/Aggression at School	9. Desired School Response to Student-to-Student Bullying/Aggression	Student and staff respondents' suggested school response to student-to-student bullying/aggression at school.	18a – 18o; 19a – 19j (25 items)	10a – 10o; 11a – 11j (25 items)
	10. Suggestions to Stop School-Based Bullying/Aggression	Respondents' perceptions of effective strategies, programs, practices, etc. to prevent bullying/aggression at school.	20 (1 item)	12 (1 item)

## Methodology

### Design

To ensure measurement quality, the BFSS was tested for its reliability and validity. **Reliability** refers to the degree to which a measure is consistent or stable. Using a car analogy, a reliable gas gauge consistently reads empty when the tank is empty. If the BFSS is reliable, the “gauge” will consistently measure levels of bullying/aggression any time it is used. **Validity**, on the other hand, is the degree to which a measure accurately measures what it’s supposed to measure. For example, a gas gauge is supposed to indicate how much gas is in the tank, not how much oil or water. A valid BFSS instrument should measure bullying/aggression, not some other construct, such as self-esteem.

As shown in Exhibit 2, the reliability and validity of the BFSS was tested using a variety of approaches, based upon conventional methods for evaluating the psychometric properties of survey instruments (Robinson et al., 1991).

### EXHIBIT 2: BFSS Psychometric Properties Tested

BFSS Psychometric property tested	Why this test?	What is the desired result for the BFSS?
<b>Reliability</b>		
Test-retest (stability)	To see if BFSS is stable in its measurement from one time to the next. It is tested by calculating a correlation between a person’s BFSS scores collected on two different occasions.	Highly correlated BFSS scores between testing time 1 and 2 (one month apart)
Internal consistency	The BFSS contains several scales that consist of multiple items. A person’s responses on these items should correlate, meaning that they are internally consistent. For example, there are seven BFSS items that measure school climate, all of which should correlate with each other.	High correlations between BFSS items within each multi-item scale.
<b>Validity</b>		
Content	To determine the degree to which items of an assessment instrument are relevant to and representative of the targeted construct for a particular assessment purpose. This assessment is conducted by researchers and practitioners who are considered experts in bullying/aggression.	High ratings from content experts for relevance of BFSS items and representativeness of BFSS dimensions
Construct: Convergent	To establish whether scores on different measures of the same construct are correlated. For example, a person’s scores on the BFSS should correlate with that person’s scores on a different bullying survey.	High correlation between students’ scores on the BFSS Target of Bullying/Aggression scale and a different bullying measure (developed by Orpinas & Horne, 2006).
Construct: Discriminant	To test whether scores on measures of two (or more) different constructs are not highly correlated. For example, a person’s scores on the BFSS should not correlate with that person’s scores on a self-esteem scale.	No or low correlation between students’ scores on BFSS Target of Bullying/Aggression scale and a self-esteem scale (developed by Rosenberg, 1965).
Construct: Concurrent	To see if the BFSS can distinguish between groups that it should theoretically be able to distinguish between. For example, bullying/aggression are more prevalent in elementary and middle school grades than high school grades, so scores on the BFSS should be higher for elementary and middle school students than their high school counterparts.	Significantly higher scores on BFSS Target of Bullying/Aggression scale for students in Gr. 5/7 students (vs Gr. 9/11 students), with a history of suspension, and who reported low school achievement. Higher scores on physical bullying/aggression for males and higher scores for relational aggression for females.

Testing the psychometric properties of the BFSS involved several steps: (a) item development and refinement (December, 2010 – January, 2011), (b) content validity analysis (February – March, 2011), (c) reliability and construct validity analysis (April – July, 2011), and (d) school staff feedback (April – June, 2011). The results from these steps are provided in the Results section.

Reliability analysis included administration of the BFSS to the same respondents twice, separated by one month, which occurred between April and June, 2011. In addition to the BFSS, respondents also completed (a) an additional bullying/aggression tool, the Reduced Aggression/Victimization Scale (RAVS) (Orpinas & Horne, 2006), to test convergent validity; (b) a measure of a different construct, self-esteem, using the Rosenberg Self-Esteem Scale (RSES) (Rosenberg, 1965), to test discriminant validity; and (c) several demographic items (i.e., gender, grade, ethnicity, discipline history, and academic performance) to test concurrent validity. Both the RAVS and RSES have been shown by their respective developers to have adequate reliability and validity.

## Sample

Both the staff and student samples were selected by convenience using a recruitment e-mail developed by the S3 evaluators that was forwarded to school personnel throughout the state by SAPE and MDE staff members. Informed consent to participate was provided to staff as part of the survey. Parental consent for student participation was collected by classroom teachers. Students without consent did not participate in the survey; instead, they were given an alternative classroom activity during the survey session.

A total of 169 staff from 24 buildings completed the survey on both testing periods. Most (96%) were teachers, with some support staff (3%) and administrators (1%). There were 1155 student respondents in Grade 5, 7, 9, and 11 from 23 buildings who completed the survey twice. Building names are withheld in order to protect confidentiality of the school districts. Students in other grades may have completed the survey, but were not included in the analysis due to low sample size. Details about each sample are provided in Exhibit 3.

### EXHIBIT 3: Characteristics of the Samples

Staff Sample (n 169)	n	(%)	Student Sample (n 1155)	n	(%)
<b>Gender</b>			<b>Gender</b>		
Female	128	75.8%	Female	562	48.7%
Male	39	24.2%	Male	593	51.3%
<b>Race/Ethnicity</b>			<b>Race/Ethnicity</b>		
White	162	96.1%	White	738	63.9%
Other	7	3.9%	African-American	176	15.2%
<b>Years working in a school setting</b>			Latino/Hispanic		
1 – 5 yrs	28	16.8%	Arab/Chaldean	113	9.9%
6 – 10 yrs	39	22.8%	Other	14	1.1%
11 – 15 yrs	39	22.8%	<b>Grade Level</b>		
16 – 20 yrs	35	20.8%	5	168	14.5%
21+ yrs	28	16.8%	7	472	40.9%
			9	366	31.7%
			11	149	12.9%
			<b>Grades earned in school (self-report)</b>		
			Mostly As/Bs	824	70.2%
			Mostly Cs	159	13.5%
			Mostly Ds/Fs	67	5.7%
			<b>Discipline record (self-report)</b>		
			Ever suspended from school	379	32.0%
			Ever expelled from school	58	5.0%

## Procedure

The BFSS was conducted using an online portal (from surveymonkey.com) available to schools between April and June, 2011. Staff completed their version of the survey twice, about one month apart (average = three weeks, five days) using their office or personal computer. The survey took approximately 20 minutes for each respondent to complete. They were paid \$25 dollars to participate and an additional \$60 to coordinate the survey administration for their respective classroom(s). After completion of the survey, teachers were asked to provide feedback in an e-mail sent to the S3 evaluator.

Students completed the survey twice, approximately one month apart (average = three weeks, four days), in a computer lab or in the classroom equipped with mobile computers. Their classroom teacher administered the survey and utilized standardized instructions for conducting the survey and addressing student questions (e.g., avoid telling students how to answer the items). The survey took approximately 20 minutes for each student to complete. An exception was among 5<sup>th</sup>-grade students, who took an average of 40 minutes to complete the survey and needed considerably more assistance from the classroom teacher than students in older grades. Students did not receive an incentive for participation and were told that their participation, or non-participation, did not affect their grade in class.

## Results

The results are organized by each major step of the survey development process; (a) item development and refinement, (b) content validity analysis, (c) reliability and construct validity analysis, and (d) school staff feedback.



## Item Development and Refinement

Under the leadership of Dee Lindenberger, SAPE originally developed the BFSS in 2005 and revised it in 2010 in order to incorporate new constructs and best practices in assessing bullying and other forms of aggression. Included in the revision were scales adapted from Stan Davis, who created the *Stop Bullying Now* intervention and collaborates on the Youth Voice Project with Dr. Charisse Nixon (from The Pennsylvania State University – Erie). The revised version consisted of 143 items and 86 items on the student and staff versions, respectively.

## Content Validity Analysis

After the item development phase, the BFSS version was evaluated by a panel of researchers and practitioners who were considered experts on school-based bullying/aggression. The panel was comprised of two university researchers, a school social worker, an Intermediate School District administrator, and an MDE administrator.

In March 2011, each panelist conducted an independent review of the BFSS. They evaluated the content validity of the instruments in three ways, based upon a strategy used by Haynes et al. (2003):

**A. Item-level rating.** Each item was rated for its relevance to the construct (e.g., School Climate to Prevent Bullying/Aggression) using a three-point Likert scale: “not necessary” (0), “useful, but not essential” (1), and “essential” (2). Virtually all items (97%) were rated as “essential” (2), suggesting that the BFSS items had high content validity. A few items (3.6%) were rated as “useful, but not essential,” or “not necessary” (2.7%).

**B. Dimension-level rating.** Each dimension (e.g., School Climate to Prevent Bullying/Aggression) was rated for its representativeness of bullying/aggression using the following four-point Likert scale: “not at all” (0), “somewhat” (1), “very much” (2), and “completely” (3). All but one (or 90%) of the dimensions was rated as “completely representative,” suggesting that the BFSS had high content validity. One dimension, Suggestions to Stop School-Based Bullying/Aggression, was rated as “very much representative.”

**C. Qualitative feedback.** Panelists were asked to provide comments that would serve to improve the content validity of the BFSS. Analysis of the results showed that over 90% of the comments focused on improving the readability of an item or the representativeness of a dimension.

These results suggest high content validity for the BFSS. The version finalized through this process was used to test reliability and construct validity. This version is described in Exhibit 1 and provided in Appendix A.

## Reliability and Construct Validity Analyses

### Reliability Analysis

Presented in Exhibit 4 are the test-retest and internal consistency reliability results for each BFSS dimension except #10: Suggestions to Stop School-Based Bullying/Aggression. Reliability for that dimension was not appropriate, because the responses were qualitative. Based upon rating criteria used by Robinson and colleagues (1991), the internal consistency of the BFSS ranged from extensive (Cronbach’s  $\alpha$  range = .70 to .79) or exemplary (Cronbach’s  $\alpha$  range = .80 or higher). One-month test-retest coefficients (using Pearson Product-Moment formula) for all dimensions ranged from .41 to .83, which is well above the minimum of .30 expected for moderate reliability. Reliability estimates by grade level and gender showed no significant differences ( $p < .05$ ). Overall, these results indicate strong reliability for the BFSS.

### EXHIBIT 4: Test-Retest and Internal Consistency Reliability Results

BFSS Dimension (# items)	STUDENT BFSS (n 1555)			STAFF BFSS (n 169)		
	# of items	One month test retest (r)	Internal consistency ( $\alpha$ )	# of items	One month test retest (r)	Internal consistency ( $\alpha$ )
1. School Climate to Prevent Bullying/Aggression	7	.41	.79	6	.58	.77
2. Classroom-Based Approaches to Prevent Bullying/Aggression	8	.46	.91	8	.63	.90
3. Behaviors to Stop Bullying/Aggression	11	.50	.93	8	.69	.95
4. Witness of Student-to-Student Bullying/Aggression	25	.54	.95	25	.76	.93
5. Witness of Staff-to-Student Bullying/Aggression	19	.55	.94	19	.64	.77
6. Target of Student-to-Student Bullying/Aggression	28	.63	.95	N/A	N/A	N/A
7. Reactions to Student-to-Student Bullying/Aggression	32	.47	.92	N/A	N/A	N/A
8. Staff Intolerance of Staff-to-Student Bullying/Aggression	N/A	N/A	N/A	2	.83	.89
9. Desired School Response to Student-to-Student Bullying/Aggression	25	.61	.97	25	.76	.95

Note. N/A = Dimension consists of one item only and/or it is not included on the student or staff version.

## Construct Validity Analysis

Three types of construct validity were tested: Convergent, discriminant, and concurrent. These tests were conducted for the student version only, due to the limited sample size expected for staff.

For analysis of **convergent validity**, it was expected that scores on the BFSS Target of Bullying/Aggression scale would correlate highly with those of another bullying/aggression scale, the Reduced Aggression/Victimization Scale (RAVS) (Orpinas & Horne, 2006). Shown in Exhibit 5 is the correlation (computed using the Pearson-Product Moment formula) between BFSS Target of Bullying/Aggression and the RAVS. As expected, scores on these two measures of bullying/aggression victimization are significantly correlated. Additional analyses showed that these results did not differ by grade level or gender ( $p < .05$ ).

**Discriminant validity** was assessed by correlating scores on the BFSS Target of Bullying/Aggression scale with a measure of self-esteem (Rosenberg, 1965). Although self-esteem is related with bullying/aggression victimization (Seals & Young, 2003), these two constructs are different, so the correlation between them should be relatively low. As shown in Exhibit 5, results confirmed this prediction, as self-esteem was only marginally related to being a target of bullying. Additional analyses showed that these results did not differ by grade level or gender ( $p < .05$ ).

### EXHIBIT 5: Correlation between Scores on BFSS Target of Bullying/Aggression, Reduced Aggression/Victimization Scale (Orpinas & Horne, 2006), and Rosenberg Self-Esteem Scale (1965)

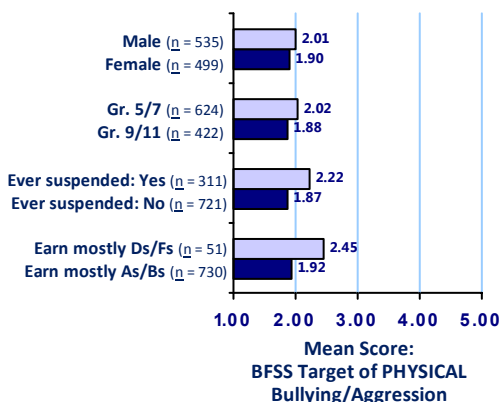
BFSS Dimension	Reduced Aggression/ Victimization Scale	Rosenberg Self- Esteem Scale
BFSS Target of Student-to-Student Bullying/Aggression	.46**	-.19*

Note.  $n = 115$ . \* $p < .05$  \*\* $p < .01$ .

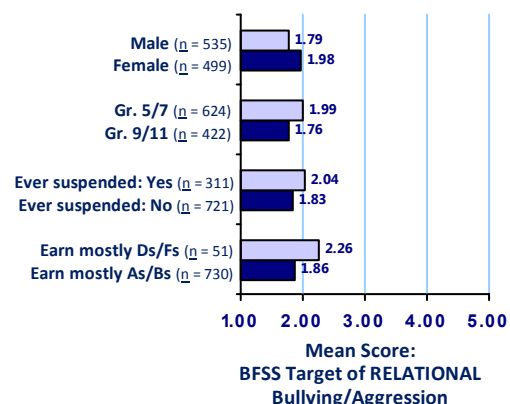
A **concurrent validity** analysis was conducted by testing the differences in scores on the BFSS Target of Bullying/Aggression scale between subgroups for gender (male, female), grade level (5, 7, 9, 11), discipline history (ever suspended: yes, no), and grades earned (Mostly As/Bs, Mostly Ds). Researchers (e.g., Nansel, 2001) have shown that higher levels of physical bullying/aggression (e.g., punching or kicking) are found among students who are male, are in elementary/middle school (versus high school), have a history of suspension, and earn lower grades (mostly Ds/Fs vs. mostly As/Bs). In addition, higher levels of relational aggression (e.g., rumors or social exclusion) are found among the same subgroups, except that females are more likely to be targets than males (e.g., Cook et al., 2010).

In Exhibit 6 are results from comparing mean scores on BFSS physical bullying/aggression by gender, grade level, suspension history, and grades earned. Exhibit 7 shows the same comparisons for mean scores on BFSS relational bullying/aggression.

### EXHIBIT 6: Mean BFSS PHYSICAL Aggression Target Scores by Gender, Grade Level, Suspension History, and Grades Earned



### EXHIBIT 7: Mean BFSS RELATIONAL Aggression Target Scores by Gender, Grade Level, Suspension History, and Grades Earned





As expected, being a target of physical aggression occurred more often for respondents who were male ( $t = 1.98, p = .05$ ), were in 5<sup>th</sup>/7<sup>th</sup> grade ( $t = 2.21, p = .03$ ), had a history of school suspension ( $t = 5.25, p < .01$ ), and earned low grades ( $t = 3.79, p < .01$ ). For relational aggression, a similar pattern of statistically significant results was found ( $t > 2.67, p < .01$ ), except females were the targets more often than males ( $t = 3.38, p < .01$ ), as predicted.

Collectively, the observed results for convergent, discriminant, and concurrent validity suggest that the construct validity of the BFSS is well above the minimally acceptable criteria for survey instruments (Robinson et al., 1991). In addition, the quality of the BFSS appears to be superior to most other bullying/aggression survey instruments available (Hamburger, Basile, & Vivolo, 2011), either in the breadth of dimensions measured and/or the reliability and validity of the instrument.

### School Staff Feedback

Teachers provided qualitative feedback after completing the BFSS Staff version ( $n = 92$ ) and/or administering the BFSS Student version to their students ( $n = 37$ ).

Regarding the **BFSS Staff version**, teachers were asked the following questions after the first test was administered, to help understand the teacher's perspective of his/her experience in completing the staff survey:

- 1) *What is your full name, address of your school, and grade(s) that you teach?*
- 2) *Were any items on the survey hard to comprehend or difficult to answer? (Yes or No, and please provide specific item numbers if Yes)*
- 3) *From a teacher perspective, did you find the survey a helpful tool to assess the bullying climate at your school? (Yes, No, or I Don't Know)*
- 4) *Any additional comments?*

Across all grades (5<sup>th</sup>, 7<sup>th</sup>, 9<sup>th</sup>, and 11<sup>th</sup>), most teachers (92%) answered "No" when asked if any survey items were hard to comprehend or difficult to answer. Most teachers (89%) answered "Yes" or "I Don't Know" when asked if the survey was a helpful tool to assess the bullying climate at their school. Some teachers (13%) felt that some items were difficult to answer and/or had additional comments to share. The majority of these comments were provided by 5<sup>th</sup> and 7<sup>th</sup> grade teachers. Presented in Exhibit 8 are a few verbatim comments from teachers. All comments are provided in Appendix B.

#### EXHIBIT 8: BFSS STAFF Version: Teacher Comments

- *Question 5 (Behaviors to Stop Bullying/Aggression) about how often we responded in a particular way to a situation was difficult. It should read if we saw the situation how would we respond to it each time. Something about the wording was weird. I had a hard time answering this set.*
- *Interesting survey. It makes you stop and think about what is going on around you and in the regular running of the school day. I found a few items difficult to answer because I don't know if they directly applied to our school.*
- *Verbal bullying is harder to detect than physical bullying. Kids are sly about bullying when adults are not present and /or in earshot. I would be interested to know the kids take on bullying, as I am an adult and do not see or hear everything. Good survey.*
- *The survey questions were helpful in assessing bullying in our school.... raised my level of awareness just by having me think specifically about various examples of bullying.*
- *They were not hard to answer as to the fact that they occur, but some are hard to rate as to how often they occur rather monthly, weekly, or daily. Yes as I can see some of these incidents happening every day but it is so hard to address all of them properly and at the same time to get the needed teaching in.*
- *This is an important topic at our school, and thank you for the opportunity. It made me think about what I can do as a teacher to stop such behavior in my classroom.*
- *I think it is good to keep the focus on bringing attention to these types of incidents, and to concentrate on following through with consequences!*

Regarding the **BFSS Student version**, teachers were asked the following questions after the first test was administered, to help understand the teacher's perspective of his/her classroom's experience in completing the student survey:

- 1) *What is your full name, address of your school, and grade(s) that you teach?*
- 2) *How many of your students completed the survey? (Please include # from each classroom, if you teach more than one classroom)*
- 3) *Did your students have trouble comprehending, or have difficulty answering, any specific questions? (Yes or No, and please provide specific item numbers if Yes)*
- 4) *Any additional comments?*

Most 5<sup>th</sup> grade (63%) teachers responded that survey items were too complex for their students to comprehend; for example, items relating to "sexual orientation" were perceived as difficult. Also, there was a repeated concern that the length of the survey was excessive for 5th grade students. Most 7th, 9th and 11th grade teachers (82%) answered "No" when asked if any survey items were hard for their students' to comprehend or difficult to answer. Some 7th grade teachers (28%) were concerned that the survey was too long, as their students had trouble staying focused for the entire survey period. All comments are provided in Appendix B.

## Conclusions, Limitations, and Recommendations

1. The level of reliability and validity of the BFSS exceeded the minimally acceptable criteria for survey instruments.
2. The reliability and validity results were similar for males, females, and all grade levels tested (5, 7, 9, & 11), suggesting that BFSS is not influenced by these demographic factors.
3. The sample was limited mostly to White and African-American students from southern, lower Michigan. Although these groups and the region comprise the majority of students in Michigan, additional studies should be done to replicate these results with other racial/ethnic groups and in other parts of the State.
4. Feedback from fifth-grade teachers who administered the BFSS to their students suggest that the instrument should be revised for upper elementary grades. Major revisions to BFSS items or dimensions would require a re-test for reliability and validity prior to its use in the field.

## References

- Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7*, 238-247.
- Cook, C.R., Williams, Kirk R., Guerra, N.G., Kim, T.E., & Sadek, S. (2010). Predictors of bullying and victimization in childhood and adolescence: A meta-analytic investigation. *School Psychology Quarterly, 25*(2), 65-83.
- Lacey, A., & Cornell, D. (2011, August). *The Impact of Bullying Climate on Schoolwide Academic Performance*. Paper presented at the Annual Conference of the American Psychological Association, Washington, DC. News release available at: [http://www.eurekalert.org/pub\\_releases/2011-08/apa-bmc072711.php](http://www.eurekalert.org/pub_releases/2011-08/apa-bmc072711.php).
- Glew, G.M., Fan, M-Y., Katon, W., Rivara, F.P., & Kernic, M.A. (2005). Bullying, Psychosocial Adjustment, and Academic Performance in Elementary School. *Archives of Pediatric Adolescent Medicine, 159*, 1026-1031.
- Hamburger, M.E., Basile, K.C., Vivolo, A.M. (2011). *Measuring Bullying Victimization, Perpetration, and Bystander Experiences: A Compendium of Assessment Tools*. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Injury Prevention and Control.
- Nansel, T.R., Overpeck, M., Pilla, R. S., Ruan, W. J., Simons-Morton, B., & Scheidt, P. (2001). Bullying behaviors among U.S. youth: Prevalence and association with psychosocial adjustment. *Journal of the American Medical Association, 285*, 1094-2100.
- Orpinas, P., & Horne, A. M. (2006). *Bullying prevention: Creating a positive school climate and developing social competence*. Washington, DC: American Psychological Association.
- Robinson, J.P., Shaver, P.R., & Wrightsman, L.S. (1991). *Measures of personality and social psychological attitudes*. San Diego, CA: Academic Press, Inc.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Seals, D. & Young, J. (2003). Bullying and Victimization: Prevalence and Relationship to Gender, Grade Level, Ethnicity, Self-Esteem, and Depression. *Adolescence, 38*, 735-747.