CrossMark

The Functionality of Four Bullying Involvement Classification Schemas: Prevalence Rates and Associations with Mental Health and School Outcomes

Tyler L. Renshaw[1] · Anthony J. Roberson[1] · Kelsie N. Hammons[1]

Abstract The present study used the 2009–2010 sample of the Health Behavior in School-Aged Children Survey ($N = 12,642$) to investigate the incremental validity of four competing bullying involvement classification schemas, which differ as a function of relative rates of endorsing victimization and perpetration behaviors at school: the standard four-group model, an alternative four-group model, a nine-group model, and a sixteen-group model. The prevalence rates and associations with concurrent mental health and school functioning outcomes were explored for each of the four bullying involvement classification schemas, which were derived using a categorical analytic approach that relied on the same two self-report items—one targeting victimization and one targeting perpetration—in the HBSC self-report survey. Results indicated substantially different prevalence rates among the four schemas and that schemas which differentiated between higher frequencies of perpetration and victimization involvement were associated with greater odds of being at risk of psychological distress, psychological well-being, academic performance, and attitudes toward school. Findings also suggested that a continuous approach to deriving bullying involvement classifications may be useful, yet preliminary post hoc analyses indicated that this approach would yield redundant results given the parameters of the present study. Implications for school mental health practice and future research are discussed.

Keywords Bullying · Classification · Mental health · School functioning

✉ Tyler L. Renshaw
   trenshaw@lsu.edu

[1] Department of Psychology, Louisiana State University, 236 Audubon Hall, Baton Rouge, LA 70803, USA

Introduction

Over the last four decades, empirical research on school bullying and its relation to youth outcomes has grown substantially (Hymel & Swearer, 2015). Reviews consistently indicate that bullying involvement is a substantial predictor of mental health problems and poor school outcomes for students around the world (Arseneault, Bowes, & Shakoor, 2010; Jimerson, Nickerson, Mayer, & Furlong, 2012; Jimerson, Swearer, & Espelage, 2010). Definitions of bullying can vary somewhat, but most scholars agree that it is a multidimensional behavioral construct characterized by three key features: (1) harmfulness, (2) frequency, and (3) peer power differential (Olweus, 2010). For a behavior to be considered true "bullying," it must (a) injure another person in some way (e.g., physically, psychologically, socially), (b) be exhibited repeatedly over time, and (c) occur within a relational context in which the perpetrator has a distinct advantage (e.g., physically, psychologically, socially) over the victim. To date, empirical work regarding definitional issues has been overshadowed by investigations aiming to classify bullying involvement, determine prevalence rates, and then explore the relation of these classifications with valued youth outcomes. The overarching purpose of the present study was to extend this line of classification research by exploring the functionality of four competing schemas for categorizing youths' bullying involvement.

Classifying Bullying Involvement

Research regarding students' bullying involvement typically proceeds by, first, using youths' self-reports to classify them into one of four mutually exclusive categories of involvement and then, second, investigating the prevalence

rates and differential psychosocial outcomes associated with these classifications. These four classifications are commonly referred to as *uninvolved*, *victim*, *bully*, and *bully–victim* (e.g., Veenstra, Lindenberg, Oldehinkel, De Winter, & Ormel, 2005) and make up the only major schema, to date, for categorizing bullying involvement. For the purposes of the present study, we slightly renamed these four classifications as follows: *uninvolved*, *victim only*, *perpetrator only*, and *perpetrating victim*. By adding "only" to two categories and replacing the term "bully" with "perpetrator," we intended to clarify the distinguishing features of the classifications and avoid common confusions in interpretation. Previous studies not utilizing this four-group classification schema have typically assessed only one domain of bullying behavior—victimization or perpetration—and therefore used a partial but more limited schema (i.e., victims vs. non-victims, perpetrators vs. non-perpetrators; see Arseneault et al., 2010). Although all items used to measure youths' bullying involvement are arranged along response scales that are characterized by relative frequency, it is noteworthy that categorical approaches to data analysis are the norm in this line of research. While using a continuous analytic approach seems warranted given the nature of the data, most bullying involvement measures are single Likert-type items, not more robust Likert-type scales comprised of multiple internally consistent items, and have therefore been treated as ordinal data as opposed to interval or ratio data when used for classification purposes (for more on this methodological issue, see Carifio & Perla, 2008; Gadermann, Guhn, & Zumbo, 2012; Norman, 2010). Adopting this methodological rationale, the present study extends this categorical analytic tradition by exploring the functionality of four competing classification schemas, all derived from single Likert-type items for measuring youths' bullying involvement at school.

Research regarding the characteristics of youth involved in bullying has focused primarily on the association between particular classifications of involvement and the presence or absence of mental health symptoms and other undesirable life outcomes (e.g., Arseneault et al., 2010). For instance, compared to *uninvolved* students, those classified as *perpetrator only* have been shown to be more antagonistic and more aggressive and to exhibit a higher degree of power dominance over their peers—indicating elevated levels of externalizing behaviors and lower levels of prosocial behaviors (Cook, Williams, Guerra, Kim, & Sadek, 2010; Veenstra et al., 2005). On the other hand, youth identified as *victim only* have been found to exhibit higher levels of insecurity, anxiety, and depression and are less likely to attend school and more likely to perform poorly at school—indicating higher levels of internalizing behaviors (Cook et al., 2010; Haynie et al., 2001;

Kochenderfer & Ladd, 1996; Swearer, Espelage, Vaillancourt, & Hymel, 2010). *Perpetrator–victim* students appear to be the most at-risk group, as they exhibit the least prosocial behavior, are the most aggressive, are likely to experience the highest rates of internalizing symptoms, have the poorest academic performance, and are considered to be the most socially "dislikable" (Cook et al., 2010; Farmer et al., 2010; Veenstra et al., 2005). Taken together, the upshot of these studies is that students' involvement in bullying at any level is associated with poorer mental health and school functioning outcomes than being uninvolved in bullying and that greater levels of involvement (i.e., *perpetrating victim*) are associated with the poorest overall outcomes. The present study further explores this outcome trend by proposing classification schemas that account for greater frequencies of bullying involvement (as both a victim and perpetrator) and then investigating the concurrent validity of these enhanced schemas compared to the original four-group schema.

Measuring Bullying Involvement

Although research suggests meaningful differences between the mental health and school functioning of students involved in bullying at different levels, there is currently no "gold standard" measure or procedure for assessing bullying involvement, and several critical issues have been raised regarding the available methods. Many scholars, including Felix, Sharkey, Green, Furlong, & Tanigawa (2011), have raised this issue in the past, concluding that it is wise for researchers interested in assessing bullying involvement to familiarize themselves with the strengths and weaknesses of available methods. For example, Cornell and Cole (2012) note that the nature of definitions of bullying included (or not) in measures are likely to influence youths' response rates, offering support from Vailancourt et al. (2010), who found fewer endorsements of bullying involvement when definitions of bullying were explicitly stated prior to the items assessing involvement. Moreover, Sharkey, Dowdy, Twyford, & Furlong (2012) suggest that the timeframe provided for endorsing bullying involvement (e.g., within the past month, past few months, or past year) is likely to influence the accuracy of prevalence rates and that the anonymity (or lack thereof) of the assessment processes might affect youths' endorsements of bullying involvement. All of these issues are essentially concerned with the self-report method of assessing bullying, which is, to date, the primary method for measuring involvement levels, deriving prevalence rates, and investigating concurrent outcomes. These critiques further suggest that findings resulting from different measures that purport to have the same function are likely to yield different data regarding youths' bullying

involvement (see Swearer, Siebecker, Johnson-Frerichs, & Wang, 2010).

The most widely used self-report instruments for gauging youths' bullying involvement in US schools are the Youth Risk Behavior Surveillance Survey (YRBSS; Centers for Disease Control and Prevention, 2014), the National Crime Victimization Survey (NCVS; Bureau of Justice Statistics, 2014), the Olweus Bully/Victim Questionnaire (OBVQ; Olweus, 1995), and the Health Behavior in School-Aged Children Survey (HBSC; World Health Organization, 2014). The YRBSS includes two items targeting victimization, asking youth to report how often they have been "bullied on school property" as well as cyberbullied during the past year. The NCVS, on the other hand, includes only a single-item targeting general victimization at school, which is often interpreted as representing bullying-related victimization. The OBVQ is a much more robust measure than both the YRBSS and NCVS, as it consists of several items targeting victimization and a few items targeting perpetration behavior, specifying various forms of bullying-related behavior (i.e., verbal, cyber, physical, relational, racial, and overall). Yet the HBSC's measure of bullying, which builds upon and extends the OBVQ, appears to be the most comprehensive measure to date, as it includes equal numbers of items targeting both victimization and perpetration across ten forms of bullying behavior: verbal, social exclusion, physical, social sabotage, racial, religious, sexual, computer, cell phone, and overall. Given its balanced focus on both victimization and perpetration as well as its administrative scope as a survey sponsored by the World Health Organization (2014), we suggest that the HBSC is currently the most useful measure available for assessing youths' bullying involvement in schools.

As mentioned above, some attention has been paid to critical issues surrounding item wording, prompts, and response scales of these common self-report measures, but little attention has been paid to the classification schemas used to operationalize and interpret the data resulting from such measures. So far, the majority of available research classifying students into different levels of bullying involvement has used the same four-group classification schema outlined earlier (i.e., *uninvolved*, *victim only*, *perpetrator only*, and *perpetrator–victim*; e.g., Greif Green, Felix, Sharkey, Furlong, & Kras, 2013; Nansel, Overpeck, Haynie, Ruan and Scheidt 2003), which places youth into mutually exclusive groups based on their categorical endorsement (or lack thereof) of bullying perpetration and victimization. And our review of the contemporary scholarship failed to yield any conceptual or empirical literature regarding competing classification schemas. Given that the original bullying involvement classification schema—which we will refer to hereafter as the *standard four-group*

model—dichotomizes the endorsement of bullying behaviors (i.e., completely uninvolved vs. involved at any level) and ignores the relative frequency that characterizes such behavior (i.e., lower vs. higher rates of involvement), we suggest that a logical next step for progressing research in this area is to evaluate competing classification schemas that vary according to their function of differentiating the frequency of youths' bullying involvement as victims and perpetrators. Although the present study uses data obtained via the HBSC, it is plausible that such competing bullying involvement schemas could be investigated with any survey that yields measures of both victimization and perpetration behaviors (e.g., OBVQ; Olweus, 1995), but not with measures that only target victimization or perpetration behavior (e.g., California Bullying Victimization Survey; Felix, Sharkey, Green, Furlong, & Tanigawa, 2011).

Purpose of the Present Study

Given the context sketched above, the purpose of the present study was to initiate a line of research exploring the functionality of competing bullying involvement classification schemas, which differ according to where the conceptual boundaries are drawn regarding students' relative rates of endorsing victimization and perpetration behaviors at school. Although classifications in the present study were defined by relative frequency of bullying involvement, suggesting interval or ratio data, the measures of victimization and perpetration were conceptualized as ordinal data because they were derived from single Likert-type items with limited response ranges, not from internally consistent scales of such items (see Carifio & Perla, 2008, for more on this issue). More specifically, the aim of this study was to investigate the differential prevalence rates and concurrent validity of the *standard four-group model* for classifying bullying involvement in comparison with three competing models—the *alternative four-group model*, the *nine-group model*, and the *sixteen-group model* (see the "Method" section, below, for operationalizations of these classification variables)—in relation to indicators of student mental health and school functioning. Considering that these competing classification schemas differed primarily according to the rate or frequency of bullying victimization and perpetration behaviors endorsed by students, we assumed that the prevalence rates yielded by all schemas would differ substantially, with the nine-group and sixteen-group schemas yielding fewer students at higher involvement levels than the four-group schemas. Furthermore, given that, theoretically, increased levels of bullying involvement should be associated with increased levels of student risk, we further hypothesized that compared to the standard and alternative four-group models, both the nine-group and the sixteen-group classification

schemas would demonstrate incremental validity across two broad indicators of mental health (i.e., psychological distress and well-being) and two general indicators of school functioning (i.e., academic performance and feelings toward school).

Method

Participants

This study was conducted using the 2009–2010 sample of the Health Behavior in School-Aged Children Survey (HBSC), which is publicly available via the Inter-university Consortium for Political and Social Research (www.icpsr.umich.edu). This nationally representative sample of US students consists of 12,642 youth in Grades 5–10. Randomized samplings of school districts were used to collect data from balanced cross sections of the country. The present sample consisted of roughly equal numbers of males (51.4 %) and females (48.6 %), while the distribution of participants by grade was also approximately equivalent (5 = 13.6 %, 6 = 16.2 %, 7 = 19.2 %, 8 = 19.6 %, 9 = 16.4 %, and 10 = 15.1 %). Youths' self-reported racial/ethnic identities were predominately White (48.8 %), followed by Hispanic (19.8 %) and Black/African American (17.9 %), with smaller proportions of Multiracial (6.8 %), Asian (3.9 %), American Indian/Alaska Native (1.8 %), and Native Hawaiian/Pacific Islander (0.9 %). Further details regarding the sampling and surveying procedures as well as more details regarding participant demographics for this particular sample are available via the 2009–2010 HBSC codebook (Iannotti, 2013).

Measures

HBSC Self-Report

The youth self-report form of the HBSC is used as the primary instrument for a cross-national research study sponsored by the World Health Organization (2014). This study originated in the 1985–1986 school year and is based on independent surveys conducted every four years in over 40 countries, including the USA. The main goals for the HBSC are to facilitate the development of longitudinal data that represents youth's health and to deliver these findings to the research community to provide further examination of risk behaviors and attitudes related to these behaviors. Further information regarding the HBSC self-report is available via www.hbsc.org and Iannotti (2013). As mentioned above, the present study used the dataset from the 2009–2010 version of the HBSC, which consists of various self-report items regarding youths' mental health, school functioning, nutrition and physical health-related behaviors, and various other aspects of psychosocial functioning. For the purposes of this study, however, only a limited number of items related to youths' bullying involvement, mental health, and school functioning were analyzed.

Bullying Involvement

The HBSC survey includes two self-report items assessing general bullying involvement at school: one item targeting victimization and the other targeting perpetration. Prior to answering these items, youth are presented with a response prompt that consists of a definition of bullying derived from the Olweus Bully/Victim Questionnaire (Olweus, 1995):

> We say a student is BEING BULLIED when another student, or a group of students, say or do nasty or unpleasant things to him or her. It is also bullying when a student is teased repeatedly in a way her or she does not like or when they are deliberately left out of things. But it is NOT BULLYING when students of about the same strength or power argue or fight. It is also not bullying when a student is teased in a friendly and playful way.

Following, the items assessing general perpetration and victimization read: "*How often have you been bullied at school in the past couple of months?*" and "*How often have you taken part in bullying another student(s) at school in the past couple of months?*" Both of these global items were arranged along a five-point, relative frequency-based response scale (i.e., 1 = *I haven't been bullied at school the past couple of months*, 2 = *It has only happened once or twice*, 3 = *2 or 3 times a month*, 4 = *About once a week*, 5 = *Several times a week*). For the purposes of the present study, and based on preliminary analyses of differentiations among the response options indicated in the response distributions, the five original response options were collapsed to four and were renamed using the following shorthand labels: 1 = *never*, 2 = *rarely*, 3/4 = *sometimes*, and 5 = *often*. All analyses were conducted and interpreted using this revised response scale. Preliminary analyses also indicated that participants' responses to both bullying involvement items were substantially non-normally distributed (skewness and kurtosis >2). However, given that these responses were conceptualized categorically and were used to create the bullying classification schemas (described below), instead of functioning as continuous outcome variables, such non-normality was deemed to be inconsequential.

Bullying Involvement Classification Schemas

The two bullying involvement items described above were used to classify youth into four competing schemas of bullying involvement, which varied primarily as a function of the relative frequency of behavior endorsed in each response. The classification variable for the *standard four-group model* was created by categorizing youth endorsing responses of "1" (indicating no involvement) for both the victimization and perpetration items as *uninvolved*, those reporting "2" or higher (indicating some level of involvement) on only the victimization item as a *victim only*, those reporting "2" or higher on only the perpetrator items as a *perpetrator only*, and those reporting "2" or higher on both items as a *perpetrator–victim*. The *alternative four-group model* used similar decision rules, classifying youth endorsing responses of "1" or "2" into the *never/rarely involved* categories and those endorsing responses of "3/4" or "5" into the *sometimes/often involved* categories for both victimization and perpetration. The *nine-group model* also classified students endorsing responses of "1" or "2" into *never/rarely involved* categories, but it categorized those endorsing responses of "3/4" as *sometimes involved* and those endorsing responses of "5" as *often involved*. Finally, the *sixteen-group model* separated out the *never involved* (endorsing "1") and *rarely involved* (endorsing "2") categories, while maintaining the same decision rules as the nine-group model for the *sometimes involved* (endorsing "3/4") and *often involved* (endorsing "5") categories. A visual representation of the decision rules and resulting categories for each classification schema is presented in Tables 1, 2, 3, and 4, respectively.

Mental Health

Youths' mental health was assessed using the Psychological Wellbeing and Distress Scale (PWDS), which is a 10-item measure embedded within the self-report HBSC survey (Renshaw & Bolognino, 2015). The PWDS has five items measuring psychological well-being and five items measuring psychological distress. The psychological well-being items use the stem, "*Thinking about last week…,*" followed by questions regarding how often one had a desirable or positive experience (e.g., "*… have you felt full*

of energy?*"). Participants endorse these items using a five-point, relative frequency-based response scale (1 = *never*, 2 = *seldom*, 3 = *quite often*, 4 = *very often*, 5 = *always*). Two items measuring psychological distress use a similar stem and response scaling, while three items use the stem, "*In the past 6 months how often have you had the following…,*" followed by statements regarding how often one experienced a particular mental health symptom (e.g., "*… feeling low*"). Response options for this stem are also arranged along a five-point, relative frequency-based response scale (1 = *about every day*, 2 = *more than once a week*, 3 = *about every week*, 4 = *about every month*, 5 = *rarely or never*). Exploratory and confirmatory factor analyses have indicated that the two subscales of the PWDS represent two distinct mental health constructs that are strongly negatively correlated ($\phi = -.59$; Renshaw & Bolognino, 2015). For the present sample, preliminary analyses indicated that both scales were relatively normally distributed (skewness and kurtosis < |1|) and had adequate internal reliability ($\alpha > .70$).

School Functioning

Youths' school functioning was assessed using two single items embedded within the self-report HBSC survey: one measuring academic performance and the other feelings (or attitudes) toward school. The item assessing academic performance read as follows: "*In your opinion, what does your class teacher(s) think about your school performance compared to your classmates?*" Four response options followed this question, all of which were anchored to generic qualitative descriptors of academic performance (1 = *very good*, 2 = *good*, 3 = *average*, 4 = *below average*). The item assessing feelings toward school read as follows: "*How do you feel about school at present?*" This item was also followed by a four-point response scale, which was characterized by generic attitudinal statements about school (1 = *I like it a lot*, 2 = *I like it a bit*, 3 = *I don't like it very much*, 4 = *I don't like it at all*). For the present sample, preliminary analyses indicated that both school functioning variables were relatively normally distributed (skewness and kurtosis < |1|).

Data Analyses

After creating the classification variables for the four competing schemas, prevalence rates for bullying involvement within each category were calculated using frequency analyses, which were then transformed into percentages for ease of interpretation. Trend analysis was then used to explore differences in the prevalence rates among the subgroups within each classification schema as well as across schemas. Following, the comparative

Table 1 Standard four-group classification schema of bullying involvement

Perpetration score	Victimization score	
	1	2, 3/4, or 5
1	Uninvolved	Victim only
2, 3/4, or 5	Perpetrator only	Perpetrator–victim

Table 2 Alternative four-group classification schema of bullying involvement

Perpetration score	Victimization score	
	1 or 2	3/4 or 5
1 or 2	Never/rarely involved	Sometimes/often victim only
3/4 or 5	Sometimes/often perpetrator only	Sometimes/often perpetrator–victim

Table 3 Nine-group classification schema of bullying involvement

Perpetration score	Victimization score		
	1 or 2	3/4	5
1 or 2	Never/rarely involved	Sometimes victim only	Often victim only
3/4	Sometimes perpetrator only	Sometimes perpetrator–victim	Sometimes perpetrator–often victim
5	Often perpetrator only	Often perpetrator–sometimes victim	Often perpetrator–victim

Table 4 Sixteen-group classification schema of bullying involvement

Perpetration score	Victimization score			
	1	2	3/4	5
1	Uninvolved	Rarely victim only	Sometimes victim only	Often victim only
2	Rarely perpetrator only	Rarely perpetrator–victim	Rarely perpetrator–sometimes victim	Rarely perpetrator–often victim
3/4	Sometimes perpetrator only	Sometimes perpetrator–rarely victim	Sometimes perpetrator–victim	Sometimes perpetrator–often victim
5	Often perpetrator only	Often perpetrator–rarely victim	Often perpetrator–sometimes victim	Often perpetrator–victim

concurrent validity of the four competing classification schemas was investigated by analyzing the relation between each of the classification variables and students' risk status across the two mental health and two school functioning outcomes. Prior to conducting this analysis, dichotomous risk-status variables were created for each concurrent outcome by transforming all variable scores into standardized values (z scores) and then using 1 SD cutoff points, which are commonly used in standardized testing and screening processes, for sorting students into "at-risk" and "typical" groups. Specifically, for the psychological distress variable, participants with z scores >1 SD were designated as "at-risk," while those with standardized scores ≤1 SD were deemed "typical." Conversely, for the psychological well-being, academic performance, and feelings toward school variables, participants with z scores <1 SD were designated as "at-risk," while those with scores ≥1 SD were deemed "typical." These dichotomous risk-status variables were then used as the criterion variables in a series of binary logistic regressions wherein the classification schema variables functioned as the predictor variables. To interpret the

logistic regression results, both statistical significance levels (p values) and effect sizes (odds ratios) were considered. Visual analysis was then used to explore the patterns of resulting odds ratios among the subgroups within each classification schema as well as across schemas.

Results

Prevalence Rates

Prevalence rates for the four competing bullying involvement classification schemas are presented in Table 5. Results for the *standard four-group model* show that approximately 43.2 % of students report being involved in bullying of some kind and to some degree in the recent past. That is, they indicated a response >"1" (or greater than "never") on either one or both of the victimization and perpetration items. Comparatively, the *alternative four-group model* yielded a dramatic decrease in the proportion of students endorsing substantive bullying involvement (16.4 %), assuming that there are no

Table 5 Prevalence rates for classification schemas of bullying involvement

Classification schema/subgroup	Prevalence rate (%)	Classification schema/subgroup	Prevalence rate (%)
Standard four-group		*Alternative four-group*	
Uninvolved	56.8	Never/rarely involved	83.6
Victim only (VO)	15.5	Sometimes/often VO	8.9
Perpetrator only (PO)	15.5	Sometimes/often PO	5.3
Perpetrator–victim (P–V)	12.2	Sometimes/often P–V	2.1
Nine-group		*Sixteen-group*	
Never/rarely involved	83.6	Uninvolved	56.8
Sometimes VO	5.5	Rarely VO	9.6
Often VO	3.4	Sometimes VO	3.4
Sometimes PO	3.9	Often VO	2.5
Often PO	1.4	Rarely PO	11.7
Sometimes P–V	1.2	Sometimes PO	2.6
Sometimes P–often V	0.4	Often PO	1.1
Often P–sometimes V	0.2	Rarely P–V	5.5
Often P–V	0.4	Rarely P–sometimes V	2.0
		Rarely P–often V	0.9
		Sometimes P–rarely V	1.3
		Sometimes P–V	1.2
		Sometimes P–often V	0.4
		Often P–rarely V	0.3
		Often P–sometimes V	0.2
		Often P–V	0.4

meaningful differences between being "never" and "rarely" involved across victimization and perpetration domains. The *nine-group model*, on the other hand, showed the same proportion of students reporting involvement with bullying as was found in the *alternative four-group model* (16.4 %), as the decision rules for this subcategory were held constant, and also indicated generally smaller proportions of students endorsing involvement at the "often" levels compared to the "sometimes" levels. Finally, the *sixteen-group model* further subdivided students' endorsements across four response categories for both victimization and perpetration, yielding the same number of uninvolved students as the *standard four-group model* yet varying levels and patterns of involvement across the remaining categories. Although there are some exceptions, the general trend of findings observed in Table 5 indicates that higher prevalence rates of bullying involvement are observed for categories characterized by lower levels of frequency of endorsement.

Concurrent Validity

Odds ratios (OR) resulting from the logistic regressions for the *standard four-group model* are presented in Table 6

and show predominantly statistically significant yet small effect sizes (excepting the highest and lowest OR) across all levels of involvement (compared to the uninvolved referent group) and all four concurrent outcomes, ranging from 1.29 (a negligible effect) to 2.67 (a medium effect). Findings for the logistic regressions for the *alternative four-group model* are presented in Table 7 and indicate statistically significant and slightly stronger OR effect sizes overall, ranging from 1.42 (a negligible effect) to 2.95 (a medium effect). OR trends for the logistic regressions for the *nine-group model* are presented in Table 8 and again show generally statistically significant and larger effect sizes overall than for both the *standard and alternative four-group models*, ranging from 1.00 (no effect compared to the referent group) to 4.58 (a large effect). Finally, results for the logistic regressions for the *sixteen-group model* are presented in Table 9 and indicate a trend of statistically significant and highest-overall OR effect sizes across concurrent outcomes in comparison with the three other classification schemas, ranging from 1.12 (a negligible effect) to 5.39 (a large effect). Although there are some exceptions, the general trend observed among the patterns of resulting OR within and between the four competing schemas indicates that higher levels of bullying

Table 6 Standard four-group classification schema: concurrent validity with student risk indicators

Bullying involvement classification	Odds ratio [95 % CI] for student risk			
	Academic performance	Attitude toward school	Psychological well-being	Psychological distress
Uninvolved	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)
Victim only	1.29 [1.14, 1.46][a]	1.45 [1.28, 1.65][a]	2.16 [1.88, 2.48][a]	1.69 [1.44, 1.98][a]
Perpetrator only	2.08 [1.85, 2.33][a]	2.27 [2.01, 2.56][a]	1.51 [1.30, 1.76][a]	1.54 [1.31, 1.81][a]
Perpetrator–victim	1.71 [1.50, 1.94][a]	1.88 [1.64, 2.15][a]	2.67 [2.31, 3.09][a]	2.18 [1.85, 2.57][a]

Effect size (OR) interpretation: ≥ 1.44 = small effect, ≥ 2.47 = medium effect, ≥ 4.25 = large effect

[a] Significant at the $p \leq .001$ level

[b] $p \leq .01$ level

[c] $p \leq .05$ level

Table 7 Alternative four-group classification schema: concurrent validity with student risk indicators

Bullying involvement classification	Odds ratio [95 % CI] for student risk			
	Academic performance	Attitude toward school	Psychological well-being	Psychological distress
Never/rarely involved	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)
Sometimes/often VO	1.50 [1.30, 1.73][a]	1.89 [1.64, 2.19][a]	2.81 [2.41, 3.26][a]	1.97 [1.65, 2.35][a]
Sometimes/often PO	2.30 [1.94, 2.73][a]	2.81 [2.36, 3.34][a]	1.87 [1.53, 2.30][a]	1.67 [1.33, 2.10][a]
Sometimes/often PV	1.42 [1.08, 1.88][c]	2.38 [1.82, 3.12][a]	2.95 [2.21, 3.92][a]	2.20 [1.59, 3.05][a]

Effect size (OR) interpretation: ≥ 1.44 = small effect, ≥ 2.47 = medium effect, ≥ 4.25 = large effect

VO victim only, *PO* perpetrator only, *PV* perpetrator–victim

[a] Significant at the $p \leq .001$ level

[b] $p \leq .01$ level

[c] $p \leq .05$ level

involvement are associated with greater odds of being "at-risk" of mental health problems and poorer school functioning.

Discussion

Interpretation of Results

The purpose of the present study was to initiate a line of research exploring the functionality of competing classification schemas for categorizing students' bullying involvement, which differ according to where the conceptual boundaries are drawn regarding endorsements of relative rates of victimization and perpetration behaviors at school. More specifically, the aim of this study was to investigate the differential prevalence rates and concurrent validity of the *standard four-group model* for classifying bullying involvement in comparison with three other models—an *alternative four-group model*, a *nine-group model*, and the *sixteen-group model*—in relation to two mental health indicators (i.e., psychological distress and

well-being) and two school functioning outcomes (i.e., academic performance and feelings about school). We hypothesized that the prevalence rates yielded by all schemas would differ substantially, with the nine-group and sixteen-group schemas yielding fewer students at higher involvement levels than the four-group schemas. Furthermore, we hypothesized that compared to the standard and alternative four-group models, both the nine-group and the sixteen-group classification schemas would demonstrate incremental validity across all concurrent indicators of mental health and school functioning. Generally speaking, the results from the primary analyses supported our hypotheses, suggesting that there is value added in extending the current four-group bullying involvement classification schema to account for greater frequencies of victimization and perpetration behaviors.

As hypothesized, prevalence rates of bullying involvement were substantively attenuated when the standard four-group dichotomous schema was changed to the alternative four-group schema (see Table 5). The prevalence rate of the *perpetrator only* category, for example, dropped from 15.5 to 5.3 %. This trend of attenuated prevalence rates

Table 8 Nine-group classification schema: concurrent validity with student risk indicators

Bullying involvement classification	Odds ratio [95 % CI] for student risk			
	Academic performance	Attitude toward school	Psychological well-being	Psychological distress
Never/rarely involved	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)
Sometimes VO	1.26 [1.05, 1.51][c]	1.47 [1.22, 1.77][a]	2.23 [1.84, 2.71][a]	1.91 [1.54, 2.38][a]
Often VO	1.94 [1.56, 2.41][a]	2.73 [2.20, 3.38][a]	3.89 [3.12, 4.85][a]	2.06 [1.58, 2.69][a]
Sometimes PO	2.11 [1.73, 2.58][a]	2.52 [2.06, 3.08][a]	1.89 [1.47, 3.25][a]	1.60 [1.22, 2.09][a]
Often PO	2.93 [2.11, 4.05][a]	3.78 [2.73, 5.22][a]	1.92 [1.31, 2.81][a]	1.88 [1.24, 2.86][b]
Sometimes PV	1.00 [0.67, 1.49]	2.39 [1.67, 3.42][a]	2.56 [1.74, 3.77][a]	2.32 [1.58, 2.69][a]
Sometimes P–Often V	1.86 [1.00, 3.45][c]	1.51 [0.77, 2.96]	2.92 [1.51, 5.66][a]	1.64 [0.73, 3.71][a]
Often P–Sometimes V	1.26 [0.48, 3.33]	2.20 [0.87, 5.60]	2.91 [1.10, 7.66][c]	1.54 [0.45, 5.92]
Often PV	3.02 [1.62, 5.63][a]	3.78 [2.03, 7.03][a]	4.58 [2.40, 8.75][a]	2.83 [1.38, 5.83][b]

Effect size (OR) interpretation: ≥1.44 = small effect, ≥2.47 = medium effect, ≥4.25 = large effect

VO victim only, PO perpetrator only, P perpetrator, V victim, PV perpetrator–victim

[a] Significant at the p ≤ .001 level

[b] p ≤ .01 level

[c] p ≤ .05 level

Table 9 Sixteen-group classification schema: concurrent validity with student risk indicators

Bullying involvement classification	Odds ratio [95 % CI] for student risk			
	Academic performance	Attitude toward school	Psychological well-being	Psychological distress
Uninvolved	1.00 (referent)	1.00 (referent)	1.00 (referent)	1.00 (referent)
Rarely VO	1.12 [0.96, 1.30]	1.12 [0.95, 1.32]	1.58 [1.32, 1.90][a]	1.34 [1.10, 1.65][b]
Sometimes VO	1.33 [1.06, 1.68][c]	1.66 [1.30, 2.11][a]	2.55 [1.98, 3.27][a]	2.26 [1.72, 2.98][a]
Often VO	2.03 [1.57, 2.62][a]	2.81 [2.18, 3.63][a]	4.42 [3.41, 5.74][a]	2.33 [1.70, 3.19][a]
Rarely PO	1.88 [1.65, 2.13][a]	1.92 [1.68, 2.20][a]	1.43 [1.21, 1.70][a]	1.45 [1.21, 1.75][a]
Sometimes PO	2.64 [2.07, 3.37][a]	3.11 [2.43, 3.97][a]	1.75 [1.28, 2.38][a]	1.65 [1.17, 2.32][b]
Often PO	3.30 [2.29, 4.75][a]	5.29 [3.67, 7.62][a]	1.85 [1.18, 2.92][b]	2.26 [1.42, 3.59][a]
Rarely PV	1.47 [1.22, 1.76][a]	1.29 [1.05, 1.58][c]	1.89 [1.52, 2.35][a]	1.99 [1.57, 2.50][a]
Rarely P–sometimes V	1.65 [1.24, 2.20][a]	1.72 [1.27, 2.32][a]	2.77 [2.03, 3.77][a]	2.15 [1.51, 3.07][a]
Rarely P–often V	2.84 [1.92, 4.22][a]	4.08 [2.74, 6.05][a]	5.03 [3.34, 7.57][a]	2.57 [1.57, 4.19][a]
Sometimes P–rarely V	2.04 [1.45, 2.88][a]	2.50 [1.77, 3.54][a]	3.21 [2.22, 4.65][a]	2.28 [1.48, 3.50][a]
Sometimes PV	1.15 [0.77, 1.71]	2.74 [1.91, 3.92][a]	3.01 [2.04, 4.45][a]	2.70 [1.76, 4.14][a]
Sometimes P–often V	2.13 [1.15, 3.96][c]	1.73 [0.88, 3.39]	3.44 [1.78, 6.78][a]	1.91 [0.84, 4.31]
Often P–rarely V	3.59 [1.75, 7.37][a]	1.85 [0.85, 4.06]	4.29 [2.03, 9.06][a]	1.91 [0.73, 5.00]
Often P–sometimes V	1.45 [0.55, 3.82]	2.52 [0.99, 6.42]	3.42 [1.30, 9.03][a]	1.79 [0.52, 6.15]
Often PV	3.47 [1.86, 6.47][a]	4.33 [2.32, 8.07][a]	5.39 [2.82, 10.31][c]	3.29 [1.59, 6.78][a]

Effect size (OR) interpretation: ≥1.44 = small effect, ≥2.47 = medium effect, ≥4.25 = large effect

VO victim only, PO perpetrator only, P perpetrator, V victim, PV perpetrator–victim

[a] Significant at the p ≤ .001 level

[b] p ≤ .01 level

[c] p ≤ .05 level

continued when the relative frequencies of endorsement were further separated into the nine-group (3.9 % for "sometimes" and 1.4 % for "often" *perpetrator only*) and sixteen-group schemas (2.6 % for "sometimes" and 1.1 % for "often" *perpetrator* only). Moreover, a broad trend was indicated when comparing the odds ratios of the classification schemas, showing that schemas containing more categories yielded relatively stronger effect sizes in

comparison with the referent group than schemas with fewer categories. This trend is attributed to the differentiation of classifications based on their relative frequency of bullying involvement in the expanded schemas, suggesting that greater levels of bullying involvement are associated with greater likelihoods of concurrent risk across mental health and school functioning indicators. The crucial categorization element these findings highlight is that when the "sometimes" and "often" responses are separated into distinct categories (compared to when they are combined into a single category), then much larger odds ratios are indicated. For instance, a comparison of the OR for the *perpetrator–victim* category across the four competing classification schemas shows steady increases in risk potential in comparison with the referent group as the schemas become progressively refined according to frequency of involvement, resulting in the largest observed effect sizes at the greatest levels of involvement (see Tables 6, 7, 8, 9).

Taken together, these findings indicate that progressively higher levels of bullying involvement, at both the victimization and perpetration levels, predict significantly poorer concurrent mental health and school functioning. Although this study was conceptualized from the categorical analytic tradition, the results suggest that it may be useful to classify bullying involvement using continuous data and associated norm-referenced cutoff scores, which function to place individuals with higher and lower levels of victimization and perpetration into *SD*-derived categories common to psychological assessment (e.g., *lower extreme*, *below average*, *low average*, *high average*, *above average*, and *upper extreme*). Although we intended to explore the viability of this continuous analytic approach post hoc, preliminary findings indicated that classifications derived using a continuous approach were identical to those derived using the original categorical approach. Specifically, post hoc analyses indicated that each possible score for both single-item measures was observed to represent a different *SD*-derived category for both victimization and perpetration (i.e., $1 = low\ average$, $2 = high\ average$, $3/4 = above\ average$, $5 = upper\ extreme$), and thus classifications derived from using a continuous approach were redundant with those derived from the categorical approach—indicating that additional analyses would only duplicate the findings obtained previously. We suggest that this finding is primarily an artifact of using single items with limited response ranges as the sole measure for both victimization and perpetration behaviors. Future research is therefore warranted to investigate the viability of a continuous analytic approach using more robust measures that consist of multiple self-report items instead of single items, which would allow for a range of scores (as opposed to a single score) to represent each *SD*-derived category. For example, the nine-item victimization and perpetration scales inherent within the HBSC may be useful for this particular purpose. Thus, we have initiated efforts to validate the utility of these scales for this purpose as well as to explore their incremental validity compared to the two single-item measures used in the present study.

Implications for Practice

Findings from the present study have potential implications for school mental health practice. Given that the overarching aim of the school mental health practitioner is to identify and provide services to students who are at the greatest risk of adverse outcomes, these results indicate that simply refining the classification schema used for screening bullying involvement (as a function of the relative frequency of response endorsements) may help identify students who are experiencing more mental health problems and poorer school functioning compared to the standard four-group schema, which is currently the most common approach used in practice. Beyond identifying students with greater needs for mental health services, findings from the present study also have implications for the feasibility of screening for bullying involvement in schools and how the resulting data might be used to inform mental health services within multitiered systems of support (see Stoiber, 2014). To make these implications more concrete, we can extrapolate the present findings to a hypothetical 500-student middle school, which is a moderate-sized school population. If a school mental health practitioner working in a school this size conducted a universal screening for bullying involvement, the results from this study suggest that the *alternative four-group model* is a more practically useful schema than the *standard four-group model*, as it identifies less students as substantially involved in bullying and thus warranting follow-up support services, and the students it does identify have generally greater risk levels than those identified by the standard model (see Tables 6, 7). From a practical point of view, most schools simply have neither the staff nor infrastructure to follow up with nearly half of the school population (231 students) which would be identified as involved in bullying in a 500-student school; however, identifying 82 students for follow-up services is markedly more reasonable.

Furthermore, when screening for bullying involvement and associated risk at the schoolwide level, the *nine-group model* identifies the same number of students as substantially involved in bullying as the *alternative four-group model* (see Table 5), yet it also shows incremental validity for identifying students with higher frequencies of involvement that are at greater risk of poorer mental health and school functioning outcomes. The *sixteen-group model*, on the other hand, identifies far more students than

would be practical to follow up with, similar to the *standard four-group model*. If the goal of the school mental health practitioner is to hone in on the students that are at greatest risk and to do so in the most feasible manner, the results of the present study suggest that the nine-group schema may be the most functional for this purpose. Both of the four-group schemas can be considered to be too simplistic and blunt to identify students with the greatest risk, while the sixteen-group schema has greater incremental validity related to risk identification but is too impractical from a follow-up service delivery perspective. In the context of the hypothetical 500-student middle school, we can extrapolate that use of the nine-group schema would identify 29 students endorsing *often* at some level of bullying or victimization and who would consequently be at the highest risk of poor mental health and school functioning outcomes. Its seems plausible that a school mental health practitioner could provide a follow-up assessment as a second gate of screening (see Glover & Albers, 2007)—potentially a more robust self-report assessment of bullying involvement—with this number of students and, if needs be, provide targeted tier-two support services (e.g., social–emotional skill building groups) as a targeted intervention.

Limitations and Future Directions

Although the findings from the present study are encouraging, they should be considered within the context of a few methodological limitations. First, given that the items assessing bullying involvement included in the HBSC consisted of five response options and the present study collapsed these to four response options, this investigation did not explore the comparative concurrent validity of a full 25-group schema (i.e., separating responses of "3" and "4" into distinct categories for another layer of analyses), which may have yielded more nuanced findings. We chose not to explore this maximized schema because the sixteen-group model already appeared to offer more information than is practically useful in school mental health practice, yet we recognize that this is a subjective decision grounded in feasibility considerations and not necessarily empirical decision rules. We should also mention that our preliminary consideration of the wording of the items (see the "Method" section) was key in deciding to combine response options, as discriminating between the relative frequency represented by responses of "3" and "4" did not appear necessary. Future research might replicate the current study using the full response options and a 25-group schema for the purposes of investigating more nuanced differences in bullying involvement categories and relations between mental health and school functioning outcomes. However, considering that findings suggest that it may be useful to classify bullying involvement using continuous data and associated norm-referenced cutoff scores, we suggest that a more fruitful direction for future research is to validate multiple-item scales that might be analyzed as continuous data and used to create competing classification schemas. Considered in this light, we suggest that the limitation of the categorical approach assumed in the present study—and the practical similarity observed between classifications derived via categorical and continuous approaches post hoc—is best understood as a methodological artifact that can be remedied in future research using more robust measures of bullying involvement.

Given that the present study used a nationally representative sample of youth from the HBSC, the typical limitations of sample demographics and the generalizability of findings to youth in schools do not apply. However, considering that data used from the HBSC were wholly derived from youths' self-reports, it is noteworthy that the findings may be biased by common method variance (i.e., the variance attributed to the measurement method rather than to the constructs represented by the measures; Podsakoff, MacKenzie, Lee, & Podsakoff 2003). To prevent against this potential bias in future studies, we recommend that follow-up investigations use variables derived from other measurement methods, including informant-report (teacher or parent) behavior rating scales of students' mental health and performance-based indicators of youths' well-being (e.g., report card grades, standardized test scores, and school discipline data). In addition, we suggest that productive next steps for progressing this line of research may include using more high-information models for data analysis that investigate the interaction effects between victimization classifications and perpetration classifications (as opposed to combining both metrics within a single variable), using continuous (as opposed to categorical) mental health and school functioning outcome variables, and testing to see whether similar trends in findings are observed when using other, non-HBSC measures of bullying involvement. Ultimately, however, we emphasize that future research in this line of inquiry should be carefully guided by concerns for incremental validity and treatment utility (see Hayes, Nelson, & Jarrett, 1987), as the intention of investigating competing classification schemas is to inform preferences for assessment methods that, in turn, inform interventions targeted to decrease youths' involvement in bullying and improve their mental health and school functioning.

References

Arseneault, L., Bowes, L., & Shakoor, S. (2010). Bullying victimization in youth and mental health problems: 'Much ado about nothing?'. *Psychological Medicine, 40*, 717–729. doi:10.1017/S0033291709991383.

Bureau of Justice Statistics. (2014). *Indicators of school crime and safety*. Washington, DC: Bureau of Justice Statistics. Retrieved from www.bjs.gov/index.cfm?ty=pbse&sid=8

Carifio, L., & Perla, R. (2008). Resolving the 50 year debate around using and misusing Likert scales. *Medical Education, 42*, 1150–1152. doi:10.1111/j.1365-2923.2008.03172.x.

Centers for Disease Control and Prevention. (2014). *Youth Risk Behavior Surveillance System (YRBSS)*. Retrieved from www.cdc.gov/HealthyYouth/yrbs/index.htm

Cook, C. R., Williams, K. 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*, 65–83. doi:10.1037/a0020149.

Cornell, D., & Cole, J. (2012). Assessment of bullying. In S. Jimerson, A. Nickerson, M. Mayer, & M. Furlong (Eds.), *The Handbook of School Violence and School Safety: International Research and Practice* (2nd ed., pp. 289–303). London: Routledge.

Farmer, T. W., Petrin, R. A., Robertson, D. L., Fraser, M. W., Hall, C. M., Day, S. H., & Dadisman, K. (2010). Peer relations of bullies, bully-victims, and victims: The two social worlds of bullying in second-grade classrooms. *Elementary School Journal, 110*, 364–392. doi:10.1086/648983.

Felix, E., Sharkey, J., Green, J., Furlong, M. J., & Tanigawa, D. (2011). Getting precise and pragmatic about the assessment of bullying: The development of the California Bullying Victimization Scale. *Aggressive Behavior, 37*, 234–247. doi:10.1002/ab.20389.

Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research, and Evaluation, 17*(3), 1–13.

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

Greif Green, J., Felix, E. D., Sharkey, J. D., Furlong, M. J., & Kras, J. E. (2013). Identifying bully victims: Definitional versus behavioral approaches. *Psychological Assessment, 25*, 651–657. doi:10.1037/a0031248.

Hayes, S. C., Nelson, R. O., & Jarrett, R. B. (1987). The treatment utility of assessment: A functional approach to evaluating assessment quality. *American Psychologist, 42*, 963–974. doi:10.1037/0003-066X.42.11.963.

Haynie, D. L., Nansel, T., Eitel, P., Crump, A. D., Saylor, K., Yu, K., & Simons-Morton, B. (2001). Bullies, victims and bully/victims: Distinct groups of at-risk youth. *Journal of Early Adolescence, 21*, 29–49. doi:10.1177/0272431601021001002.

Hymel, S., & Swearer, S. M. (2015). Four decades of research on school bullying. *American Psychologist, 70*, 293–299. doi:10.1037/a0038928.

Iannotti, R. J. (2013). *Health behavior in school-aged children (HBSC), 2009–2010 (ICPSR34792-v1)*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research. Retrieved from http://doi.org/10.3886/ICPSR34792.v1

Jimerson, S. R., Nickerson, A. B., Mayer, M. J., & Furlong, M. J. (Eds.). (2012). *Handbook of school violence and school safety: International research and practice* (2nd ed.). New York, NY: Routledge.

Jimerson, S. R., Swearer, S. M., & Espelage, D. L. (Eds.). (2010). *Handbook of bullying in schools: An international perspective*. New York, NY: Routledge.

Kochenderfer, B. J., & Ladd, G. W. (1996). Peer victimization: Cause of consequence of school maladjustment? *Child Development, 67*, 1305–1317. doi:10.1111/j.1467-8624.1996.tb01797.x.

Nansel, T. R., Overpeck, M., Haynie, H. L., Ruan, W. J., & Scheidt, P. C. (2003). Relationship between bullying and violence among U.S. youth. *Archives of Pediatrics and Adolescent Medicine, 157*, 348–353. doi:10.1001/archpedi.157.4.348.

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education, 15*, 625–632. doi:10.1007/s10459-010-9222-y.

Olweus, D. (1995). Bullying or peer abuse at school: Facts and interventions. *Current Directions in Psychological Science, 4*, 196–200. doi:10.1111/1467-8721.ep10772640.

Olweus, D. (2010). Understanding and researching bullying: Critical issues. In S. R. Jimerson, S. M. Swearer, & D. L. Espelage (Eds.), *Handbook of bullying in schools: An international perspective* (pp. 9–34). New York, NY: Routledge.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*, 879–903. doi:10.1037/0021-9010.88.5.879.

Renshaw, T. L., & Bolognino, S. J. (2015). *Psychometric properties of the Psychological Wellbeing and Distress Screener: A brief measure of youths' bidimensional mental health*. Unpublished manuscript.

Sharkey, J. D., Dowdy, E., Twyford, J., & Furlong, M. J. (2012). An overview of measurement issues in school violence and school safety research. In S. R. Jimerson, A. B. Nickerson, M. J. Mayer, & M. J. Furlong (Eds.), *The Handbook of School Violence and School Safety: International Research and Practice* (2nd ed., pp. 259–273). London: Routledge.

Stoiber, K. C. (2014). A comprehensive framework for multitiered systems of support in school psychology. In P. Harrison & A. Thomas (Eds.), *Best practices in school psychology: Data-based and collaborative decision making* (pp. 41–70). Bethesda, MD: National Association of School Psychologists.

Swearer, S. M., Espelage, D. L., Vaillancourt, T., & Hymel, S. (2010a). What can be done about school bullying? Linking research to educational practice. *Educational Researcher, 39*, 38–47. doi:10.3102/0013189X09357622.

Swearer, S. M., Siebecker, A. B., Frerichs, L. J., & Wang, C. (2010b). Assessment of bullying/victimization: The problem of comparability across studies and across methodologies. In S. R. Jimerson, S. M. Swearer, & D. L. Espelage (Eds.), *Handbook of bullying in schools: An international perspective* (pp. 305–327). New York, NY: Routledge.

Vaillancourt, T., Trinh, V., Mcdougal, P., Duku, E., Cunningham, L., Cunningham, C., et al. (2010). Optimizing population screening of bullying in school-aged children. *Journal of School Violence, 9*, 233–250. doi:10.1080/15388220.2010.483182.

Veenstra, R., Lindenberg, S., Oldehinkel, A. J., De Winter, A. F., & Ormel, J. (2005). Bullying and victimization in elementary schools: A comparison of bullies, victims, and bully/victims, and uninvolved preadolescents. *Developmental Psychology, 41*, 672–682. doi:10.1037/0012-1649.41.4.672.

World Health Organization. (2014). *Health behavior in school-aged children: World Health Organization collaborative cross-national survey*. Retrieved from http://www.who.int/en/