\*\*\*WORKING DRAFT\*\*\*

Last update: 5.16.2016

Structural Validity of the HBSC Bullying Measure:  
A Self-Report Rating Scale of Youth Victimization and Perpetration Behavior

Anthony J. Roberson and Tyler L. Renshaw

Louisiana State University

Author Note

Anthony J. Roberson and Tyler L. Renshaw, Department of Psychology, Louisiana State University.

Correspondence regarding this article should be addressed to Anthony J. Roberson, Department of Psychology, Louisiana State University, 236 Audubon Hall, Baton Rouge, LA 70803 USA. Email: arob125@lsu.edu

The authors are grateful to Ronald J. Iannotti for making available the Health Behavior in School-aged Children survey datasets and their codebooks for public use.

Abstract

Bullying involvement among youth has consistently been recognized as potentially leading to serious consequences for both perpetrators and victims. To help clarify the nature and scope of youths’ bullying involvement, empirically-validated assessment instruments measuring victimization and perpetration behaviors are needed for use in research and practice. The present study investigated the latent factor structure of the 22 victimization and perpetration items within the 2009–2010 Health Behavior in School-aged Children (HBSC) self-report survey. Structural validity analyses were conducted using the representative sample of U.S. youth in grades 5–10 (*N* = 11,449) obtained from the national administration of this survey. Results suggested a four-factor latent structure comprised of *traditional perpetration*, *traditional victimization*, *cyber perpetration*, and *cyber victimization* was the most theoretically and psychometrically sound measurement model. Additionally, measurement invariance analyses showed that this model functioned equitably across student race/ethnicity, sex, and grade-level, supporting the measure’s use with diverse student populations.

*Keywords*: bullying, scale development, structural equation modeling, elementary school, high school

**Structural Validity of the HBSC Bullying Measure: A Self-Report Rating Scale of Youth Victimization and Perpetration Behavior**

The issue of bullying involvement among young people has received a great amount of research and media attention over the past few decades (Olweus, 2010) and remains a salient topic today (Hymel & Swearer, 2015). Part of the draw to study this phenomenon is the well-established link between bullying involvement and numerous deleterious effects on youths’ social, academic, and psychological health (Jimerson, Nickerson, Mayer, & Furlong, 2012). Furthermore, the many negative effects associated with bullying involvement have been found across demographics, including gender, race/ethnicity, grade-level, and nationality (Nansel, Overpeck, Pilla, Ruan, Simons-Morton, & Scheidt, 2001; Nansel, Craig, Overpeck, Saluja, & Ruan, 2004; Renshaw, Roberson, & Hammons, 2015). At the same time, some longitudinal studies of prevalence rates have shown that bullying involvement has actually been decreasing steadily in the U.S. over the last several years. Though this may suggest reason for optimism, some estimated rates of bullying involvement are nonetheless intolerably high, with certain studies suggesting as many as 60.4% of youth endorsing involvement as a perpetrator and/or victim to some degree (Renshaw, Hammons, & Roberson, in press). Despite the evident public health risks, bullying involvement among students frequently goes under-addressed, or completely unaddressed, in schools (Swearer, Espalage, Vaillancourt, & Hymel, 2010).

Due to the elevated risks connected with bullying involvement, it is an important precondition to prevention and intervention that researchers and school professionals use technically adequate instruments to measure youths’ perpetration and victimization behavior. While some recent meta-analytic reviews of instruments for assessing bullying involvement show that more than 40 measures are available (Vessey, Strout, DiFazio, & Walker, 2014), these studies indicate major discrepancies in conceptual approaches to bullying measurement as well as varied evidence quality backing the validity of such measures (Vivolo-Kantor, Martell, Holland, & Westby, 2014). This variability in validity evidence makes it difficult to reach consensus about the nature and scope of youths’ bullying involvement (e.g., Renshaw et al., in press). Furthermore, these issues pose a challenge for practitioners in selecting appropriate assessment tools for screening or progress monitoring intervention and prevention efforts (Bradshaw, 2015). In order to remedy these issues, additional studies aimed at developing conceptually coherent and psychometrically robust measures of youths’ victimization and perpetration behavior are essential.

Though no single instrument has been adequately researched, Vessey et al. (2014) have suggested that the revisedOlweus Bully/Victim Questionnaire (OBVQ; Olweus, 1996) is the most technically adequate and commonly used bullying measure to date. While some other self-report bullying scales used for general prevalence estimation utilize only a scarce few items––such as the two items concerning school victimization in the Youth Risk Behavior Surveillance Survey (YRBSS; Center for Disease Control and Prevention, 2014) or the single victimization item in the National Crime Victimization Survey (NCVS; Bureau of Justice Statistics, 2015)––the OBVQ assesses both perpetration and victimization using multiple items targeting many bullying behavior classes (i.e., verbal, physical, relational, racial, cyber, and general), making it among the most conceptually broad instruments available.

Given the evidence that both perpetration and victimization are associated with negative outcomes (Nansel et al., 2004; Renshaw et al., 2015), the ability to derive scores for both perpetration and victimization behaviors are important features of a self-report bullying instrument. Assessing perpetration in conjunction with victimization also allows for the identification of students who experience both types of bullying behaviors (i.e., perpetrator–victims), whom research has linked with the most negative outcomes of all (Haynie et al., 2001; Renshaw et al., 2015). Although several existing measures use single, domain-general items to assess overall victimization and perpetration behavior—like those on the YRBSS or NCVS––research suggests that items targeting specific bullying behaviors yield higher prevalence rates, suggesting improved measurement specificity (Renshaw et al., in press). Furthermore, it seems reasonable that measures composed of items targeting specific behavior classes would contribute to greater precision in scientific research regarding bullying as well as greater usability (see Glover & Albers, 2007) and treatment utility (see Hayes, Nelson, & Jarrett, 1987) for practitioners intervening with bullying in schools.

Given the context sketched above, the purpose of the present study was to investigate the structural validity of an under-researched measure that targets specific bullying victimization and perpetration behaviors: the self-report bullying items within the Health Behavior in School-Aged Children survey (HBSC), sponsored by the World Health Organization (2014). These bullying items were modeled after the OBVQ and expanded in content to target parallel classes of victimization and perpetration behaviors (e.g., physical victimization and physical perpetration). Although several studies investigating bullying involvement have been conducted using isolated items from HBSC self-report surveys (e.g., Renshaw et al., in press; Nansel et al., 2004), no study has yet to investigate the structural validity of the specific involvement items when taken together as a coherent measure. Given that an item-content evaluation suggested the HBSC self-report survey targeted at least two overarching classes of bullying behavior­—victimization and perpetration—the following hypotheses were proposed:

1. Exploratory factor analysis (EFA) of the HBSC items targeting specific bullying behaviors will yield two distinct latent factors: victimization and perpetration.
2. Confirmatory factor analysis (CFA) of this same set of items will corroborate the measurement model yielded by the EFA.
3. The measurement model will be invariant across several key student demographic factors: grade-level, sex, and race/ethnicity.

**Method**

**Participants**

The current study utilized the publicly-available 2009–2010 HBSC dataset composed of a stratified random sample of U.S. youth (Iannotti, 2013). A detailed summary of participant sex, grade-level, and race/ethnicity across the original full sample (*N* = 11,449), the full sample with missing data removed (*N* = 9,979), and two random split-half subsamples (derived for data analytic purposes, described below) is found in Table 1. Participants in the original full sample were approximately equally male and female, enrolled in grades 5–10 (grade-level representation range = 13.7–20.3%), and approximated racial/ethnic proportions of U.S. youth, with the largest identities being White (49.6%), Hispanic (17.7%), and Black/African-American (15.3%). Relative proportions of each demographic were largely maintained across all iterations of the sample (see Table 1).

**Measure**

The self-report version of the HBSC is administered to a nationally-representative sample of U.S. students in grades 5–10 every four years. The survey assesses a broad range of behaviors among youth, including drug/alcohol use, body image, attitudes about school, peer/family relationships, physical health, bullying involvement, and more. To operationalize the “bullying” construct for respondents, a prompt with a formal definition––adapted from the OBVQ—was provided at the beginning of the item set:

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.

The HBSC contains 22 items targeting specific classes of bullying behaviors––11 items for victimization (see Table 2) and 11 parallel items for perpetration (see Table 3)––covering teasing, social exclusion, physical aggression, spreading lies, various kinds of harassment, as well as multiple kinds of cyber bullying. The victimization and perpetration items were arranged in two item sub-sets and were prefaced with a similar item stem (i.e., *How often have you [been bullied/bullied another students(s)] at this school in the past couple of months in the ways listed below?*). Response options for all bullying involvement items were arranged along a five-point, relative frequency-based scale (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*).

**Data Analyses**

The structural validity of the HBSC bullying involvement items was analyzed in a multi-stage process using IBM SPSS version 23, IBM Amos version 23 (Arbuckle, 2014), and R statistical environment (R Core Team, 2015). Initial exploratory analyses involved calculation of general descriptive statistics to investigate the distributionality of the target items. Evaluation of skewness, kurtosis, and omnibus normality suggested that all items were substantially non-normally distributed (all coefficients > |2.0|, *p* < .001).

Missing data were handled in a two-step process. First, 647 cases (5.7%) from the original full sample were deleted list-wise because they had missing data across all bullying items of interest, suggesting respondents failed to complete this section of the survey. Second, of the remaining cases, 823 (7.2%) had partially missing data for the relevant variables and were also deleted list-wise. Because the number of cases with partially missing responses accounted for less than 10% of the sample, list-wise deletion methods were deemed acceptable and unlikely to significantly bias statistical estimates (Langkamp, Lehman, & Lemeshow, 2010). The remaining cases were then divided into two random split-halves––subsample one (S1; *N* = 4,989) and subsample two (S2; *N* = 4,990)––to perform EFA and structural equation modeling analyses (i.e., CFA, measurement invariance, and latent means analyses), respectively. Once an appropriate factor structure was established, subscale scores were calculated at the observed level to investigate their psychometric properties.

**Results**

**Latent Factor Structure**

**EFA (S1).** Given the non-normal distributionality of the data and the hypothesized two-factor solution, EFA was conducted using the principal axis factoring method with a direct oblimin rotation (Field, Miles, & Field, 2012). Results of the first analysis, which constrained a two-factor solution, showed strong Kaiser-Meyer-Olkin sampling adequacy (.94), no multicollinearity (matrix determinant > 0), and adequate extracted item communalities (*h*2 range = .36–.76). Table 4 summarizes the factor loading output and psychometric characteristics of this measurement model. Eigenvalues for both factors were larger than the 1.17 cutoff suggested by parallel analysis and collectively accounted for 54.65% of the variance. Internal consistency estimates were strong for both factors (α > .70). Pattern matrix factor loadings for all victimization (λ range = .63­­–.71) and perpetration (λ range = .60­–.88) items were greater than the minimum threshold of .30, with no significant cross-loadings. Overall, these EFA results suggested a two-factor measurement model was reasonable.

An empirically-derived EFA solution was also computed for comparison. Visual analysis of the scree plot and inspection of eigenvalues suggested that there were three statistically meaningful factors with eigenvalues greater than the 1.17 cutoff, accounting for 60.88% of the variance. Review of the pattern matrix showed three factors consisting of 11 items related to perpetration (λ range = .65–.84), seven items related to traditional (i.e., non-cyber) victimization (λ range = .43­­–.80), and four items pertaining to cyber victimization plus the vReligious item (λ range = .33–.82), which cross-loaded with the traditional victimization factor. A third EFA was conducted to examine the factor structure after removing vReligious item and indicated a strong three-factor model (*perpetration*, *traditional victimization*, and *cyber victimization*) with no further cross-loadings. However, because this empirically-derived model did not preserve a parallel structure for victimization and perpetration behaviors––which may create practical challenges for researchers and practitioners wanting to assess matching dimensions of bullying––the original, theoretically-specified, two-factor solution was selected as the preferred measurement model for the next phase of analysis.

**CFA (S2).** To corroborate the two-factor measurement model indicated by EFA, CFA was conducted with S2 by regressing each of the 22 observed bullying items onto their respective latent factors of perpetration and victimization, which were covaried. Data–model fit was assessed using fit statistics­ recommended by Kline (2015)—overall model chi-square (χ2), comparative fit index (CFI), root-mean-square-error-of-approximation (RMSEA), and standardized-root-mean-square-residual (SRMR). The following criterion values were used to indicate at least adequate data–model fit: χ2 with an associated *p* > .05 (Kline, 2015), CFI ≥ .85 (Little, 2013), RMSEA with 90% confidence interval ≤ .08 (Kenny, 2014), and SRMR ≤ .08 (Hu & Bentler, 1999). Given the distributionality of the items in S2 showed extreme non-normality (similar to S1), a maximum likelihood (ML) estimation method with bootstrapped standard errors was deemed the most appropriate analytic approach (Kline, 2015; Kahle, n.d.). Bollen-Stine bootstrapped probability values were calculated from 2000 random samples of the data and reported alongside the traditional ML χ2. Non-significant (*p* > .05) Bollen-Stine estimates suggest adequate overall model fit.

Findings showed the two-factor measurement model was characterized by robust latent factor loadings (≥ .30) for perpetration (λ range = .54–.89) and victimization (λ range = .49–.84), as well as a large covariance between the factors (ϕ= .58). However, model fit indices indicated poor data–model fit: ML χ2 (208) = 13157.72, *p* < .001, Bollen-Stine bootstrap χ2 *p* < .001, CFI = .837, RMSEA = .112 [90% CI = .110­–.113], SRMR = .066. Modification indices were then used to explore how data–model fit may change after adding covariances between several item error terms. However, this approach did not result in an adequately parsimonious and parallel structure, therefore we proceeded to testing alternate measurement models.

To follow-up on the empirically-derived three-factor measurement model suggested by EFA, alternative models were tested with victimization, perpetration, and cyber victimization as factors with vReligious included––ML χ2 (206) = 8803.04, *p* < .001, Bollen-Stine bootstrap χ2 *p* < .001, CFI = .892, RMSEA = .091 [90% CI = .090­–.093], SRMR = .053––and excluded––ML χ2 (186) = 8134.73, *p* < .001, Bollen-Stine bootstrap χ2 *p* < .001, CFI = .896, RMSEA = .093 [90% CI = .091–.094], SRMR = .052. These models showed mixed data–model fit both before and after error covariance modifications were introduced. Additionally, these models again posed practical difficulties given the non-parallelism of the factor structure.

Finally, we tested a four–factor “compromise” model that separated traditional and cyber bullying while preserving the victimization/perpetration parallelism: *traditional victimization* (7 items), *traditional perpetration* (7 items), *cyber victimization* (4 items), and *cyber perpetration* (4 items). Each latent factor was covaried with all others. Results showed strong factor loadings (λ > .60) for all items, moderate-to-large covariances between each latent factor (ϕ range = .40–.87), and strong data–model fit: ML χ2 (203) = 5760.42, *p* < .001; B­ollen-Stine bootstrap χ2 *p* < .001, CFI = .930, RMSEA = .074 [90% CI = .072­–.076], SRMR = .043 (see Figure 1). Additionally, each latent factor had strong internal reliability (*H* range = .86–.95) and the items accounted for moderate-to-large amounts of variance for their respective factor (average variance extracted range = .47–.81). This four-factor structure was ultimately retained as the preferred measurement model for the 22 bullying involvement items, as it was the most conceptually sound and statistically robust of all the models tested.

**Measurement invariance and latent means analysis (S2).** Multigroup CFA was then conducted with the same subsample (S2) and four-factor structure to investigate the model’s measurement invariance across three demographic factors: grade-level, sex, and race/ethnicity. Due to insufficient subgroup sample sizes (*n* < 100) in S2, students identifying as Native American or Pacific Islander were not included in these analyses. Results from the measurement invariance analyses (see Table 5) showed that the data had good data–model fit at configural, metric, and metric-and-scalar levels across student sex. Comparatively, RMSEA for grade-level and race/ethnicity were equally strong but CFI were weaker yet still minimally acceptable for grade-level and race/ethnicity. Moreover, the ∆CFI between levels of invariance were < .01 for all demographics (see Table 5), suggesting the measurement model was sufficiently invariant across groups. Given the finding of multigroup invariance, subsequent analysis of latent means was deemed appropriate. Results from these tests yielded several statistically-significant (*p* < .05) estimated-standardized-mean-differences (ESMD), although the magnitude of effects appeared to be negligible (ESMD < |0.2|) across all comparisons for grade-level (see Table 6), sex (see Table 7), and race/ethnicity (see Table 8).

**Observed subscale descriptive statistics (S2).** In addition to the latent-level analyses (described above), basic descriptive statistics were also calculated at the observed level (see Table 9). Results indicated that scores derived from the four scales of the HBSC bullying measure were characterized by significant non-normality (> |2.0|), strong internal consistency (α range = .86–.94), and moderate-to-large inter-item correlations (*r* range = .47–.81).

**Discussion**

Although several instruments assessing youths’ bullying involvement currently exit, to date, none appear as conceptually broad or as balanced across victimization and perpetration domains as the measure included within the HBSC self-report survey. The present study was the first to investigate the structural validity of the 22 items that make up this measure, using the publicly-available, representative sample of U.S. youth in grades 5–10 from the 2009–2010 iteration of the survey. We hypothesized that results from factor analyses would support a two-factor measurement model that was invariant across key demographic factors. Although EFA findings supported the tenability of a two-factor measurement model, CFA findings indicated that this two-factor structure yielded poor data–model fit. Results from other, theoretically-derived CFA models indicated that a four-factor measurement model characterized by *traditional perpetration* (e.g., calling others mean or hurtful names), *traditional victimization* (e.g., getting hit, kicked, or pushed by others), *cyber perpetration* (e.g., bullying others via a computer), and *cyber victimization* (e.g., getting bullied via text messages or cellphone calls) was the most conceptually sound and yielded the best data–model fit. Thus, this four-factor structure was selected as the preferred measurement model. The hypothesis that the preferred structure would function equitably across student sex, race/ethnicity, and grade-level was largely supported from measurement invariance analyses. Findings from follow-up latent means analyses indicated negligible ESMD between demographic subgroups across all factors, suggesting no practically meaningful differences across groups.

Given that the four-factor measurement model for the instrument was sufficiently invariant across a representative sample of U.S. students, no adjustments to scoring or changes to administration procedures appear necessary when using the measure to assess bullying involvement among students from diverse demographic backgrounds. Furthermore, because this study utilized such a large and representative dataset, the observed subscale means and standard deviations presented herein (see Table 9) may be used as norms for future research or practice, such as universal screening for identifying students at-risk for bullying involvement or for progress monitoring prevention or intervention efforts at the level of classrooms, grade-levels, or schools.

Though this study capitalized on a considerably large, publicly-available dataset with a demographically diverse sample of youth, more research is called for to further validate the HBSC self-report bullying measure. In order to potentially create a more parsimonious instrument, future studies could employ item response theory procedures to evaluate differential item functioning that may lead to reducing the total number of items while retaining strong technical adequacy. Future research concerning the measure’s concurrent and predictive validity with other physical and psychological health indicators is also encouraged. This would allow for a greater understanding of which classes of bullying behaviors represented by this measurement model—traditional and cyber victimization and perpetration, respectively—are more or less salient indicators of risk. Relatedly, this measure could also be co-validated with pre-existing bullying instruments to determine which has stronger predictive power in relation to valued student outcomes. Finally, the measure could be applied in future treatment utility studies (see Hayes et al., 1987) that investigate how well scores derived from the instrument inform intervention.

References

Arbuckle, J. L. (2014). Amos (Version 23.0) [Computer Program]. Chicago, IL: IBM SPSS.

Bureau of Justice Statistics. (2015). *Indicators of school crime and safety*. Washington, DC: Bureau of Justice Statistics. Retrieved from www.bjs.gov/index.cfm?ty=pbse&sid=8

Bradshaw, C. P. (2015). Translating research to practice in bullying prevention. *American Psychologist*, *70*, 322–332. doi:10.1037/a0039114

Centers for Disease Control and Prevention. (2014). *Youth Risk Behavior Surveillance System* *(YRBSS).* Retrieved from http://www.cdc.gov/mmwr/pdf/ss/ss6304.pdf

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. London, England: SAGE.

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

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. *The Journal of Early Adolescence*, *21*, 29–49. doi:10.1177/0272431601021001002

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55. doi:10.1080/10705519909540118

Hymel, S., & Swearer, S. M. (2015). Four decades of research on school bullying: An Introduction. *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.

Kahle, C. P. (n.d.). *Handling non-normal data using AMOS*. Retrieved from https://stat.utexas.edu/software-faqs/amos

Kenny, D. A. (2014). *Measuring model fit in structural equation modeling*. Retrieved from www.davidak- enny.net/cm/fit.htm

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. New York, NY: Guilford.

Langkamp, D. L., Lehman, A., & Lemeshow, S. (2010). Techniques for handling missing data in secondary analyses of large surveys. *Academic Pediatrics*, *10*, 205–210. doi:10.1016/j.acap.2010.01.005.

Little, P. T. D. (2013). *Longitudinal structural equation modeling*. New York, NY: Guilford.

Nansel, T. R., Overpeck, M., Pilla, R. S., Ruan, W. J., Simons-Morton, B., & Scheidt, P. (2001). Bullying behaviors among US youth: Prevalence and association with psychosocial adjustment. *The Journal of the American Medical Association*, *285*, 2094–2100. doi:10.1001/jama.285.16.2094.

Nansel, T. R., Craig, W., Overpeck, M. D., Saluja, G., & Ruan, W. J. (2004). Cross-national consistency in the relationship between bullying behaviors and psychosocial adjustment. *Archives of Pediatrics & Adolescent Medicine*, *158*, 730–736. doi:10.1001/archpedi.158.8.730

Olweus, D. (1996). *The Revised Olweus Bully/Victim Questionnaire*. Bergen, Norway: University of Bergen, Research Center for Health Promotion.

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.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Renshaw, T. L., Hammons, K. N., Roberson, A. J. (in press). General vs. specific methods for classifying U.S. students’ bullying involvement: Investigating classification agreement, prevalence rates, and concurrent validity. *School Psychology Review*.

Renshaw, T. L., Roberson, A. J., & Hammons, K. N. (2015). The functionality of four bullying involvement classification schemas: Prevalence rates and associations with mental health and school outcomes. *School Mental Health.* Advance online publication. doi:10.1007/s12310-015-9171-y

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

Vessey, J., Strout, T. D., DiFazio, R. L., & Walker, A. (2014). Measuring the youth bullying experience: A systematic review of the psychometric properties of available instruments. *Journal of School Health*, *84*, 819­–843. doi:10.1111/josh.12210

Vivolo-Kantor, A. M., Martell, B. N., Holland, K. M., & Westby, R. (2014). A systematic review and content analysis of bullying and cyber-bullying measurement strategies. *Aggression and Violent Behavior*, *19*, 423–434. doi:10.1016/j.avb.2014.06.008

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

Table 1

*2009–2010 HBSC Demographic Proportions for Sex, Grade-level, and Race/Ethnicity across all Samples*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Full Sample | | Full Sample after Removing Cases | | Subsample 1 | | Subsample 2 | |
|  | *N* | *%* | *N* | % | *N* | % | *N* | % |
| *Total Sample* | 11449 | 100 | 9979 | 100 | 4989 | 100 | 4990 | 100 |
| *Sex* |  |  |  |  |  |  |  |  |
| Male | 5883 | 51.4 | 5048 | 50.6 | 2518 | 50.5 | 2530 | 50.7 |
| Female | 5562 | 48.6 | 4930 | 49.4 | 2470 | 49.5 | 2460 | 49.3 |
| Missing Data | 4 | 0.0 | 1 | 0.0 | 1 | 0.0 | 0 | 0.0 |
| *Grade-level* |  |  |  |  |  |  |  |  |
| 5th | 1574 | 13.7 | 1207 | 12.1 | 617 | 12.4 | 590 | 11.8 |
| 6th | 1781 | 15.6 | 1524 | 15.3 | 750 | 15.0 | 774 | 15.5 |
| 7th | 2269 | 19.8 | 1963 | 19.7 | 966 | 19.4 | 997 | 20.0 |
| 8th | 2327 | 20.3 | 2061 | 20.7 | 1010 | 20.2 | 1051 | 21.1 |
| 9th | 1817 | 15.9 | 1662 | 16.7 | 855 | 17.1 | 807 | 16.2 |
| 10th | 1681 | 14.7 | 1562 | 15.7 | 791 | 15.9 | 771 | 15.5 |
| *Race/Ethnicity* |  |  |  |  |  |  |  |  |
| American Indian/ Alaskan Native | 206 | 1.8 | 176 | 1.8 | 83 | 1.7 | 93 | 1.9 |
| Asian | 448 | 3.9 | 406 | 4.1 | 194 | 3.9 | 212 | 4.2 |
| Black/African American | 1841 | 16.1 | 1522 | 15.3 | 771 | 15.5 | 751 | 15.1 |
| Hispanic | 2097 | 18.3 | 1768 | 17.7 | 861 | 17.3 | 907 | 18.2 |
| Native Hawaiian/Other Pacific Islander | 105 | 0.9 | 89 | 0.9 | 48 | 1.0 | 41 | 0.8 |
| White | 5485 | 47.9 | 4953 | 49.6 | 2512 | 50.5 | 2441 | 48.9 |
| Two or More | 757 | 6.6 | 679 | 6.8 | 334 | 6.7 | 345 | 6.9 |
| Missing Data | 510 | 4.5 | 386 | 3.9 | 186 | 3.7 | 200 | 4.0 |

Table 2

*2009–2010 HBSC Items Targeting Bullying Victimization*

|  |  |
| --- | --- |
| Code | Item |
| vVerbal | “I was called mean names, was made fun of, or teased in a hurtful way.” |
| vExclusion | “Other students left me out of things on purpose, excluded me from their group of friends, or completely ignored me.” |
| vPhysical | “I was hit, kicked, pushed, shoved around, or locked indoors.” |
| vRelational | “Other students told lies or spread false rumors about me and tried to make others dislike me.” |
| vRacial | “I was bullied with mean names and comments about my race or color.” |
| vReligious | “I was bullied with mean names and comments about my religion.” |
| vSexual | “Other students made sexual jokes, comments or gestures to me.” |
| vComp | “I was bullied using a computer or e-mail messages or pictures.” |
| vCell | “I was bullied using a cell phone.” |
| vCompOut | “I was bullied outside of school using a computer or e-mail messages or pictures.” |
| vCellOut | “I was bullied outside of school using a cell phone.” |

Table 3

*2009–2010 HBSC Items Targeting Bullying Perpetration*

|  |  |
| --- | --- |
| Code | Item |
| pVerbal | “I called another student(s) mean names, made fun of, or teased him or her in a hurtful way.” |
| pExclusion | “I kept another student(s) out things on purpose, excluded him or her from my group of friends or completely ignored him or her.” |
| pPhysical | “I hit, kicked, pushed, shoved around, or locked another student(s) indoors.” |
| pRelational | “I spread false rumors about another student(s) and tried to make others dislike him or her.” |
| pRacial | “I bullied another student(s) with mean names or comments about his or her race or color.” |
| pReligious | “I bullied another student(s) with mean names or comments about his or her religion.” |
| pSexual | “I made sexual jokes, comments or gestures to another student(s).” |
| pComp | “I bullied another student(s) using a computer, e-mail messages or pictures.” |
| pCell | “I bullied another student(s) using a cell phone.” |
| pCompOut | “I bullied others outside of school using a computer, e-mail messages or pictures.” |
| pCellOut | “I bullied others outside of school using a cell phone.” |

Table 4

*EFA Pattern Matrix Results for the Two-Factor Measurement Model*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | | Factor Loadings (λ) | | | |
| Item Code | | Perpetration (ξ1) | | Victimization (ξ2) | |
| pCell | **.879** | | -.012 | |
| pComp | **.867** | | -.011 | |
| pCompOut | **.866** | | -.012 | |
| pReligious | **.834** | | -.008 | |
| pCellOut | **.825** | | .004 | |
| pRacial | **.796** | | -.033 | |
| pRelational | **.774** | | .054 | |
| pSexual | **.734** | | .025 | |
| pPhysical | **.719** | | .016 | |
| pExclusion | **.670** | | .021 | |
| pVerbal | **.597** | | .016 | |
| vRelational | -.092 | | **.707** | |
| vExclusion | -.089 | | **.690** | |
| vComp | .128 | | **.683** | |
| vCell | .140 | | **.678** | |
| vCompOut | .151 | | **.677** | |
| vVerbal | -.131 | | **.674** | |
| vPhysical | .002 | | **.673** | |
| vReligious | .059 | | **.664** | |
| vCellOut | .155 | | **.658** | |
| vRacial | .025 | | **.643** | |
| vSexual | -.022 | | **.625** | |
| Eigenvalues | | 8.78 | | 3.25 | |
| % variance | | 39.89 | | 14.76 | |
| α | | .89 | | .94 | |

Table 5

*Measurement Invariance Results for the Four-Factor Measurement Model Across Demographics*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Demographic / Invariance Level | ML χ2 | B-S χ2 (SE) | *df* | RMSEA [90% CI] | CFI | ΔCFI |
| *Race/Ethnicity* |  |  |  |  |  |  |
| Configural | 16266.71 | 7174.80 (18.32) | 1231 | .043 [.042, .043] | .863 | –– |
| Metric | 16373.36 | 7266.78 (18.43) | 1249 | .042 [.042, .043] | .863 | .000 |
| Full Metric and Scalar | 16431.84 | 7289.19 (18.45) | 1271 | .042 [.041, .043] | .862 | .001 |
| *Grade-level* |  |  |  |  |  |  |
| Configural | 12664.12 | 7037.35 (17.16) | 1268 | .042 [.042, .043] | .869 | –– |
| Metric | 13228.61 | 7464.29 (17.85) | 1340 | .042 [.042, .043] | .863 | .006 |
| Full Metric and Scalar | 13551.73 | 7574.91 (17.96) | 1450 | .041 [.040, .042] | .861 | .002 |
| *Sex* |  |  |  |  |  |  |
| Configural | 7200.37 | 2249.00 (6.66) | 406 | .058 [.057, .059] | .916 | –– |
| Metric | 7353.10 | 2371.62 (7.01) | 424 | .057 [.056, .058] | .914 | .002 |
| Full Metric and Scalar | 7593.11 | 2393.62 (7.03) | 446 | .057 [.057, .058] | .911 | .003 |

*Note.* ML = Maximum Likelihood; B-S = Bollen-Stine bootstrap; SE = standard error; RMSEA [90% CI] = root mean square error approximation with 90% confidence interval; CFI = comparative fit index. All Maximum Likelihood and Bollen-Stine bootstrap χ2 values were significant at the *p* < .001 level.

Table 6

*Latent Means Analysis Results for Grade-level Comparisons*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Comparison / Factor | ESMD | SE | CR | *p* |
| *5th – 6th Grade* |  |  |  |  |
| Traditional Victimization | 0.008 | 0.047 | 0.18 | .86 |
| Traditional Perpetration | -0.005 | 0.029 | -0.18 | .86 |
| Cyber Victimization | 0.025 | 0.028 | 0.89 | .37 |
| Cyber Perpetration | -0.015 | 0.026 | -0.57 | .57 |
| *5th – 7th Grade* |  |  |  |  |
| Traditional Victimization | 0.020 | 0.045 | 0.45 | .65 |
| Traditional Perpetration | 0.038 | 0.027 | 1.38 | .17 |
| Cyber Victimization | 0.060 | 0.028 | 2.14 | .03 |
| Cyber Perpetration | 0.000 | 0.024 | 0.01 | .99 |
| *5th – 8th Grade* |  |  |  |  |
| Traditional Victimization | -0.076 | 0.041 | -1.84 | .07 |
| Traditional Perpetration | 0.030 | 0.027 | 1.11 | .27 |
| Cyber Victimization | 0.031 | 0.026 | 1.18 | .24 |
| Cyber Perpetration | 0.009 | 0.025 | 0.37 | .71 |
| *5th – 9th Grade* |  |  |  |  |
| Traditional Victimization | -0.124 | 0.042 | -2.92 | < .01 |
| Traditional Perpetration | 0.020 | 0.028 | 0.71 | .48 |
| Cyber Victimization | 0.027 | 0.027 | 1.00 | .32 |
| Cyber Perpetration | -0.016 | 0.025 | -0.65 | .52 |
| *5th – 10th Grade* |  |  |  |  |
| Traditional Victimization | -0.153 | 0.042 | -3.61 | < .01 |
| Traditional Perpetration | 0.049 | 0.03 | 1.64 | .10 |
| Cyber Victimization | 0.037 | 0.029 | 1.29 | .20 |
| Cyber Perpetration | 0.018 | 0.028 | 0.64 | .52 |
| *6th – 7th Grade* |  |  |  |  |
| Traditional Victimization | 0.012 | 0.041 | 0.28 | .78 |
| Traditional Perpetration | 0.043 | 0.025 | 1.70 | .09 |
| Cyber Victimization | 0.036 | 0.026 | 1.36 | .18 |
| Cyber Perpetration | 0.015 | 0.022 | 0.68 | .50 |
| *6th – 8th Grade* |  |  |  |  |
| Traditional Victimization | -0.085 | 0.038 | -2.24 | .03 |
| Traditional Perpetration | 0.035 | 0.025 | 1.42 | .16 |
| Cyber Victimization | 0.006 | 0.024 | 0.26 | .80 |
| Cyber Perpetration | 0.024 | 0.023 | 1.05 | .29 |
| *6th – 9th Grade* |  |  |  |  |
| Traditional Victimization | -0.132 | 0.039 | -3.41 | < .01 |
| Traditional Perpetration | 0.025 | 0.026 | 0.96 | .34 |
| Cyber Victimization | 0.002 | 0.025 | 0.09 | .93 |
| Cyber Perpetration | -0.001 | 0.023 | -0.05 | .96 |
| *6th – 10th Grade* |  |  |  |  |
| Traditional Victimization | -0.161 | 0.039 | -4.17 | < .01 |
| Traditional Perpetration | 0.054 | 0.028 | 1.94 | .05 |
| Cyber Victimization | 0.013 | 0.027 | 0.47 | .64 |
| Cyber Perpetration | 0.033 | 0.026 | 1.26 | .21 |
| *7th – 8th Grade* |  |  |  |  |
| Traditional Victimization | -0.096 | 0.035 | -2.78 | .01 |
| Traditional Perpetration | -0.008 | 0.023 | -0.32 | .75 |
| Cyber Victimization | -0.029 | 0.024 | -1.20 | .23 |
| Cyber Perpetration | 0.009 | 0.021 | 0.44 | .66 |
| *7th – 9th Grade* |  |  |  |  |
| Traditional Victimization | -0.144 | 0.036 | -4.02 | < .01 |
| Traditional Perpetration | -0.018 | 0.024 | -0.74 | .46 |
| Cyber Victimization | -0.033 | 0.025 | -1.31 | .19 |
| Cyber Perpetration | -0.016 | 0.021 | -0.79 | .43 |
| *7th – 10th Grade* |  |  |  |  |
| Traditional Victimization | -0.173 | 0.036 | -4.86 | < .01 |
| Traditional Perpetration | 0.011 | 0.026 | 0.42 | .68 |
| Cyber Victimization | -0.023 | 0.027 | -0.84 | .40 |
| Cyber Perpetration | 0.018 | 0.024 | 0.73 | .46 |
| *8th – 9th Grade* |  |  |  |  |
| Traditional Victimization | -0.048 | 0.032 | -1.50 | .13 |
| Traditional Perpetration | -0.010 | 0.024 | -0.43 | .67 |
| Cyber Victimization | -0.004 | 0.023 | -0.17 | .87 |
| Cyber Perpetration | -0.025 | 0.021 | -1.19 | .24 |
| *8th – 10th Grade* |  |  |  |  |
| Traditional Victimization | -0.077 | 0.032 | -2.43 | .02 |
| Traditional Perpetration | 0.019 | 0.026 | 0.71 | .48 |
| Cyber Victimization | 0.007 | 0.025 | 0.26 | .80 |
| Cyber Perpetration | 0.009 | 0.025 | 0.35 | .73 |
| *9th – 10th Grade* |  |  |  |  |
| Traditional Victimization | -0.029 | 0.033 | -0.88 | .38 |
| Traditional Perpetration | 0.029 | 0.027 | 1.07 | .28 |
| Cyber Victimization | 0.010 | 0.026 | 0.39 | .69 |
| Cyber Perpetration | 0.034 | 0.025 | 1.38 | .17 |

*Note.* ESMD = estimated standardized mean difference; CR = critical ratio.

Table 7

*Latent Means Analysis Results for Sex Comparison*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Comparison / Factor | ESMD | SE | CR | *p* |
| *Males – Females* |  |  |  |  |
| Traditional Victimization | 0.003 | 0.022 | 0.12 | .91 |
| Traditional Perpetration | -0.075 | 0.015 | -4.99 | < .01 |
| Cyber Victimization | 0.007 | 0.015 | 0.45 | .65 |
| Cyber Perpetration | -0.053 | 0.013 | -3.98 | < .01 |

*Note.* ESMD = estimated standardized mean difference; CR = critical ratio.

Table 8

*Latent Means Analysis Results for Race/Ethnicity Comparisons*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Comparison / Factor | ESMD | SE | CR | *p* |
| *White – Black* |  |  |  |  |
| Traditional Victimization | 0.081 | 0.035 | 2.32 | .02 |
| Traditional Perpetration | 0.123 | 0.026 | 4.65 | < .01 |
| Cyber Victimization | 0.061 | 0.025 | 2.40 | .02 |
| Cyber Perpetration | 0.099 | 0.022 | 4.49 | < .01 |
| *White – Hispanic* |  |  |  |  |
| Traditional Victimization | 0.042 | 0.028 | 1.47 | .14 |
| Traditional Perpetration | 0.075 | 0.019 | 3.87 | < .01 |
| Cyber Victimization | 0.029 | 0.019 | 1.52 | .13 |
| Cyber Perpetration | 0.058 | 0.015 | 3.76 | < .01 |
| *White – Asian* |  |  |  |  |
| Traditional Victimization | 0.034 | 0.052 | 0.65 | .51 |
| Traditional Perpetration | -0.004 | 0.036 | -0.12 | .91 |
| Cyber Victimization | -0.010 | 0.035 | -0.29 | .77 |
| Cyber Perpetration | -0.002 | 0.028 | -0.08 | .94 |
| *White – Two or More* |  |  |  |  |
| Traditional Victimization | 0.000 | 0.021 | 0.00 | > .99 |
| Traditional Perpetration | 0.000 | 0.014 | 0.00 | > .99 |
| Cyber Victimization | 0.000 | 0.014 | 0.00 | > .99 |
| Cyber Perpetration | 0.000 | 0.011 | 0.00 | > .99 |
| *Black – Hispanic* |  |  |  |  |
| Traditional Victimization | -0.039 | 0.040 | -0.99 | .32 |
| Traditional Perpetration | -0.048 | 0.029 | -1.64 | .10 |
| Cyber Victimization | -0.032 | 0.029 | -1.13 | .26 |
| Cyber Perpetration | -0.042 | 0.024 | -1.70 | .09 |
| *Black – Asian* |  |  |  |  |
| Traditional Victimization | -0.047 | 0.059 | -0.79 | .43 |
| Traditional Perpetration | -0.127 | 0.042 | -3.03 | < .01 |
| Cyber Victimization | -0.072 | 0.041 | -1.74 | .08 |
| Cyber Perpetration | -0.102 | 0.034 | -2.98 | .00 |
| *Black – Two or More* |  |  |  |  |
| Traditional Victimization | -0.081 | 0.035 | -2.32 | .02 |
| Traditional Perpetration | -0.123 | 0.026 | -4.65 | < .01 |
| Cyber Victimization | -0.061 | 0.025 | -2.40 | .02 |
| Cyber Perpetration | -0.099 | 0.022 | -4.49 | < .01 |
| *Hispanic – Asian* |  |  |  |  |
| Traditional Victimization | -0.008 | 0.056 | -0.14 | .89 |
| Traditional Perpetration | -0.079 | 0.038 | -2.09 | .04 |
| Cyber Victimization | -0.039 | 0.037 | -1.05 | .29 |
| Cyber Perpetration | -0.060 | 0.030 | -1.99 | .05 |
| *Hispanic – Two or More* |  |  |  |  |
| Traditional Victimization | -0.042 | 0.028 | -1.47 | .14 |
| Traditional Perpetration | -0.075 | 0.019 | -3.87 | < .01 |
| Cyber Victimization | -0.029 | 0.019 | -1.52 | .13 |
| Cyber Perpetration | -0.058 | 0.015 | -3.76 | < .01 |
| *Asian – Two or More* |  |  |  |  |
| Traditional Victimization | -0.034 | 0.052 | -0.65 | .51 |
| Traditional Perpetration | 0.004 | 0.036 | 0.12 | .91 |
| Cyber Victimization | 0.010 | 0.035 | 0.29 | .77 |
| Cyber Perpetration | 0.002 | 0.028 | 0.08 | .94 |

*Note.* ESMD = estimated standardized mean difference; CR = critical ratio.

Table 9

*Observed Descriptive Statistics of the HBSC Bullying Measure*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Subscale | Items | Min, Max | *M* | *SD* | *IQR* | *g*1 | *g*2 | *K*2 | *r* | α |
| Traditional Victimization | 7 | 7, 35 | 9.59 | 4.65 | 3 | 2.71 | 8.23 | 2824.46 | .47 | .86 |
| Traditional Perpetration | 7 | 7, 35 | 8.39 | 3.64 | 1 | 4.41 | 23.01 | 4477.01 | .57 | .90 |
| Cyber Victimization | 4 | 4, 20 | 4.56 | 2.16 | 0 | 4.93 | 26.46 | 4847.81 | .74 | .92 |
| Cyber Perpetration | 4 | 4, 20 | 4.42 | 1.95 | 0 | 5.94 | 38.23 | 5557.37 | .81 | .94 |

*Note.* Min, Max = Minimum and maximum observed scale scores, *IQR* = interquartile range, *g*1 = skewness, *g*2 = kurtosis, *K*2 = omnibus normality, *r* = average inter-item correlation, α = Cronbach’s alpha. All *g*1, *g*2, and *K*2 estimates were significant at the *p* < .0001 level.

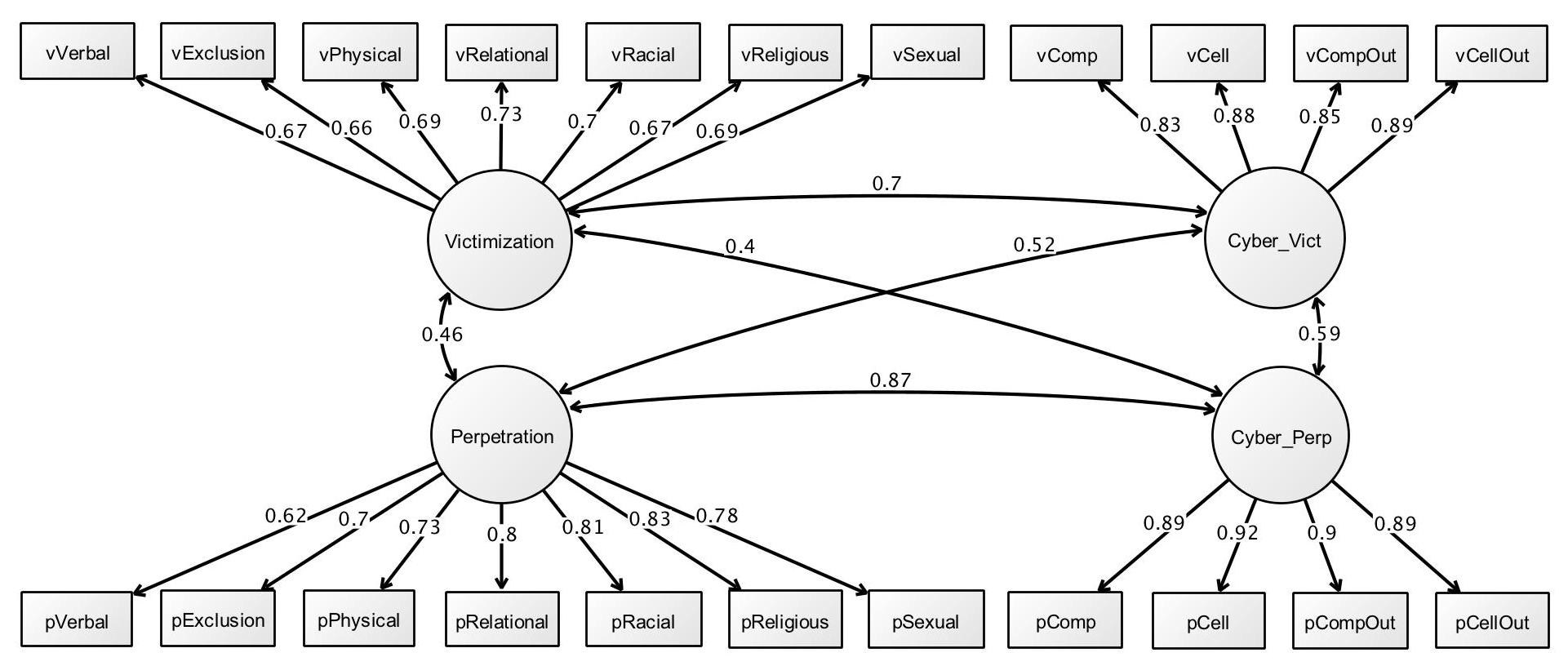
****

Figure 1. *Full structural measurement model of the four-factor HBSC Bullying Measure.*

*Note*. Victimization = *Traditional Victimization*; Perpetration = *Traditional Perpetration*; Cyber\_Vict = *Cyber Victimization*; Cyber\_Perp = *Cyber Perpetration*. Error terms were included for each observed variable during analysis but do not appear in the figure for visual simplicity.