Published in final edited form as:

Nurs Res. 2018 ; 67(4): 294–304. doi:10.1097/NNR.0000000000000283.

Item Generation and Content Validity of the Child-Adolescent Bullying Scale

Rachel L. Difazio, PhD, RN, PPCNP-BC, FAAN*,

Nurse Scientist, Orthopedic Center, Boston Children's Hospital, Instructor, Harvard Medical School, 300 Longwood Ave., Boston, MA

Tania D. Strout, PhD, RN, MS,

Director of Research, Associate Professor, Maine Medical Center, Department of Emergency Medicine, Tufts University School of Medicine, Portland, ME

Judith A. Vessey, PhD, MBA, RN, FAAN, and

Lelia Holden Carroll Professor in Nursing, William F. Connell School of Nursing, Boston College, Chestnut Hill, MA. Nurse Scientist, Boston Children's Hospital, Medicine Patient Services

Amanda Lulloff, MSN, RN, PCNS, CPHON

Nurse Educator, Dana-Farber Jimmy Fund Clinic, Boston, MA

Abstract

Background—Healthcare providers do not uniformly screen young patients for exposure to bullying and no screening instruments have been developed for widespread use in clinical settings.

Objectives—The objectives of this study were to: 1) generate scale items by identifying and eliciting concepts relevant to youths with potential exposure to bullying as well as to professionals who work with bullied youths and 2) assess the content validity of the new Child-Adolescent Bullying Scale (CABS) instrument.

Design—A mixed methods design was used to develop an initial pool of 52 items. The study was conducted in four phases: I) comprehensive review of the existing literature, II) concept elicitation through the conduct of focus groups with school-age youths and professionals who work with bullied youths, III) concept selection and item construction, and IV) content validation assessment of relevance, clarity and dimension of each item by a panel of 30 international bullying and measurement experts through completion of an online survey.

Results—An initial pool of 52 potential items was developed during Phase I – III of the study. During Phase IV, item- and scale-level content validity indices were calculated and were used to refine the item pool. These strategies resulted in a new, 22-item tool with scale-level content validity indices of 0.954 for clarity and 0.920 for representativeness.

Discussion—Evaluation of the CABS tool with a sample of youths drawn from healthcare settings will be necessary to assess the performance of the CABS items, further evaluate its psychometric properties, and further refine the tool.

*Corresponding Author: Phone: 617-240-8287, Rachel.Difazio@childrens.harvard.edu.

Conflicts of Interest: The authors declare no conflicts of interest.

Keywords

Adolescent; Psychometrics; Focus Groups; Bullying; Surveys and Questionnaires

With one in five American middle school students experiencing bullying annually, youth bullying is a significant public health problem (U.S. Department of Education, 2016). Bullied youths often experience a wide-range of serious psychosocial, physical and academic difficulties; these problems can persist into adulthood and contribute to poorer life outcomes (Nakamoto & Schwartz, 2010; Reijntjes, Kamphuis, Prinzie, & Telch, 2010; Reijntjes et al., 2011). To begin to address this, *Healthy People 2020* and widespread federal and state legislation specifically seek to reduce bullying behavior among youths (HealthyPeople 2020 2012; Sacco, Silbaugh, Corredor, Casey, & Doherty, 2012).

Despite its prevalence and the associated poor outcomes, bullying is not routinely assessed by healthcare providers. There is a critical need for healthcare providers ‘on the front lines’ to adopt proactive, systematic screening for youth bullying. Over 90% of middle school youths have at least one visit to a healthcare provider annually, often for concerns where bullying might be a factor (Uddin, O’Connor, & Ashman, 2016). There is no evidence that healthcare providers routinely initiate conversations about bullying resulting in millions of missed opportunities for early identification and intervention for those exposed to bullying (Spector & Kelly, 2006; Vessey, DiFazio, & Strout, 2013). Explanations for this are that federal and professional oversight bodies only offer generic advice on the need for providers to routinely address violence prevention but do not explicitly recommend assessing for bullying either as part of preventive care or episodic care for known associated conditions (U.S. Preventive Services Task Force, 2018; Committee on Injury, Violence, and Poison Prevention, 2009). A lack of attention to bullying during healthcare encounters may occur because youths do not readily identify being bullied as a healthcare problem (Vessey, DiFazio, & Strout, 2017). When seeking care for psychosocial and physical complaints resulting from bullying, both the patient and provider may not be aware of the association between the youth’s distress and bullying that occurs (Lyznicki, McCaffree, & Robinowitz, 2004).

Healthcare settings are an appropriate venue for screening for bullying victimization. Standardized screening for common pediatric conditions is an essential aspect of pediatric healthcare (Committee on Practice and Ambulatory Medicine, Bright Futures Periodicity Schedule Workgroup, 2017). Screening helps to presumptively detect an otherwise unrecognized condition, thus identifying youths who are more likely to benefit from further assessment and treatment if necessary (Wilson & Jungner, 1968). While bullying meets the guidelines for conditions appropriate for screening, healthcare providers are hampered by the lack of a psychometrically valid screening measure appropriate for use in busy healthcare settings (Wilson & Jungner, 1968). Being able to effectively screen for bullying will assist healthcare providers in intervening through toolkits such as available through the federal bullying prevention website (StopBullying.gov). Suggested interventions include education and counseling strategies about bullying, referrals to mental healthcare professionals, and tips for working with school and community personnel (Committee on

Injury, Violence, and Poison Prevention, 2009; Fekkes, Pijpers, Fredriks, Vogels, & Verloove-Vanhorick 2006; Lyszynski, 2004; U.S. Health Resources Services Administration, 2017).

In addition, the rapidly expanding science of bullying has been limited by a lack of definitional clarity and consistency, the absence of theoretical models upon which to ground bullying research, and limited use of research designs beyond description and correlation (Volk, Veenstra, & Espelage, 2017). Although a plethora of instruments exists, virtually all are designed for group assessment or evaluation of bullying prevalence, many are not psychometrically defensible and none were designed for screening individual youths for bullying exposure (Vessey, Strout, DiFazio, & Walker, 2014). Importantly, youths' voices were rarely captured during the item development process for most existing measures, a limitation that may have resulted in instruments that may not adequately reflect youths' own perspectives and experiences (Vessey et al., 2014). Collectively, the design and intent of existing instruments renders them inappropriate for widespread screening for bullying of individual youths in busy healthcare settings.

Recognizing the limitations of bullying assessment, the Centers for Disease Control and Prevention (CDC) collaborated with the U.S. Department of Education and the National Center for Injury Prevention and Control to craft the first Uniform Definition of bullying to promote standardization in measurement. It also suggests data elements specifically to improve surveillance and inform intervention efforts (Gladden et al., 2014). They define bullying as: "any unwanted aggressive behavior(s) by another youth or group of youths who are not siblings or current dating partners that involves an observed or perceived power imbalance and is repeated multiple times or is highly likely to be repeated; bullying may inflict harm or distress on the targeted youth including physical, psychological, social, or educational harm" (Gladden et al., 2014, pg. 7). This new definition, as well as experiences and pitfalls identified in earlier bullying instrument development efforts, were taken into account when planning our approach to construction of a new youth bullying screening tool appropriate for use in healthcare settings, the Child-Adolescent Bullying Screen (CABS).

The overarching objectives of this study were to: 1) generate scale items by identifying and eliciting concepts relevant to youths with potential exposure to bullying as well as to professionals who work with bullied youths and 2) assess the content validity of the new CABS instrument. These objectives were addressed in four phases: I) review of the existing literature, II) concept elicitation, III) concept selection and item construction, and IV) content validation (Figure 1). Here, we describe the methods and results for each of these phases. In addition, we report on how findings from each phase informed the subsequent phases to culminate in the new CABS instrument.

Methods

Across the study, a mixed methods approach was employed to data collection and analysis, with qualitative methodologies supporting the first three phases and a methodologic design guiding Phase IV. Polit and Beck (2014) define mixed methods research as studies "in which both qualitative and quantitative data are collected and analyzed to address different but

related questions” (Polit & Beck, 2014, p. 385). Burns and Grove (2009) describe methodologic designs as unique designs that specifically aid in the development or modification of measurement tools by assisting in the generation of evidence supporting the reliability and validity of the tool under study (Burns & Grove, 2009, pp. 371–394). Methodologic designs generally employ the collection and analysis of quantitative data, as in the present study. Additionally, reflecting the current state of the youth bullying science, both deductive and inductive approaches to item generation were employed (Hinkin, 1998). Institutional Review Board (IRB) approval, parental or participant informed consent, and child assent were all obtained as required throughout each phase of the study.

Phase I: Literature Review

Instruments are likely to have strong content validity when a deductive approach is used in creating a classification schema for item formation, selection and delineation (Hinkin, 1998). An initial and thorough literature review is requisite to this process. Literature reviews that use iterative search strategies and review techniques lead to a fuller understanding of the construct under examination and help to identify knowledge gaps, and inform the creation of a template to use in item development and selection (Finfgeld-Connett & Johnson, 2013; Hinkin, 1995). Integrative reviews summarize both empirical and theoretical literature and thus are highly suited in providing a comprehensive picture of the phenomenon under study (Whittemore & Knafl, 2005). Systematic reviews, although designed to answer discreet questions using explicit methods, can further contribute to construct delineation (Grant & Booth, 2009).

Method—For the reasons stated above, information gleaned from three literature reviews conducted by the research team informed the development of the CABS items (Vessey, Luloff, Hernandez Meneses, Strout, & DiFazio, 2016; Vessey et al., 2013; Vessey et al., 2014). Review strategies were guided by standardized approaches including the PRISMA guidelines (Moher, Liberati, Tetzlaff, Altman, & PRISMA Group, 2009).

Results—The first of these was an integrative review of the ‘state of the science’ of youth bullying. Most of the articles reviewed were notably atheoretical, focusing on discreet bullying behaviors and outcomes. The review resulted in the identification of youth bullying as a serious public health problem with important consequences for its victims, however it is rarely addressed by health care providers regardless of practice setting. Moreover, the overall quality of available evidence specific to detecting bullying was limited, underscoring the need for context-relevant measurement strategies for screening (Vessey et al., 2013).

The second, a systematic review of the psychometric properties of self-report youth bullying instruments was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher, Liberati, Tetzlaff, Altman, & PRISMA Group, 2009; Vessey et al., 2014). A search of four electronic databases identified 31 studies. Further review revealed that there was sufficient information available on 27 unique instruments needed for evaluating their structure, item domains, and items. Data extraction and appraisal provided objective evidence on the instruments characteristics and their psychometric quality. Conceptual and methodological quality issues identified in these

existing youth bullying measures included the lack of theoretical grounding and limited evidence supporting their reliability, validity, and responsiveness. Findings provided further support that all existing measures were primarily designed for school or group use, none were appropriate for screening in healthcare settings, and there was limited evidence supporting the psychometric soundness of the instruments.

Because research on youth bullying is highly dynamic, an additional integrative review primarily comprised of systematic reviews and meta-analyses, was conducted to identify emerging knowledge in the field (Vessey et al., 2016). Across all three reviews, the lack of a coherent theoretical framework within the youth bullying literature was evident. The bullying definition most frequently utilized in the reviewed articles was from Olweus' seminal work, "(a) person is bullied when he or she is exposed, repeatedly and over time, to negative actions on the part of one or more other persons, and he or she has difficulty defending himself or herself" (Olweus, 1993, p. 9), despite that little testing of this definition or its elements is in evidence. Virtually the entire group of peer reviewed articles was written prior to the release of the Uniform Definition of bullying (Gladden et al., 2014) and none used this as a standard upon which to base conceptual or methodological discussions.

Despite these limitations, the findings allowed for the creation of an initial template for item construction and classification. Conceptual areas that were identified included bullying roles and the defining characteristics of bullying including: power imbalance, repetition, intent, unwanted negative behavior or aggression, effects of bullying, school connectedness or disconnectedness, and help-seeking activities. An initial set of potential items also were culled from existing tools for later consideration.

Phase II: Concept Elicitation

Using both the concepts identified through the literature reviews and elements of the Universal Definition (Gladden et al., 2014), it was determined that obtaining additional qualitative data was necessary for further concept elicitation as well as to obtain youths' and professionals perspectives on the definition of bullying. A focus group method was selected as participant interaction in the groups provides mutual support and produce insights that may have not surfaced during individual interviews (Stewart & Shamdasani, 2014).

Method—Two sets of focus groups were conducted, youths and professionals who work with bullied youths, to capture the fullest array of information possible. Although in-depth descriptions of the conduct of these focus groups is provided elsewhere (Strout, DiFazio, & Vessey, 2017; Vessey, DiFazio, & Strout, 2017), information specific to concept elicitation and item generation is presented here.

Middle school and junior high youths aged 12–16 years participated in four focus groups; three groups were conducted with participants from diverse communities across New England and one group was conducted with young people hospitalized with chronic health conditions. Experienced moderators used a semi-structured approach and an interview guide. The interview guide was constructed to specifically explore concepts relevant to bullying and its components as previously derived from the literature review. General questions about bullying and its sequelae and questions about the components of bullying were included.

A purposive sample of professionals who directly worked with youths also was recruited from the Northeast United States. These professionals participated in two focus groups following the same format as described above, but the groups were facilitated through use of online meeting software in order to garner participation of professionals from diverse disciplines and locales. For these groups however, additional questions were added to the interview guide specific to bullying measurement from their own experiences and perspectives.

For both groups, audiotapes were transcribed verbatim and were analyzed using conventional qualitative content analysis methods as detailed by Krippendorff (2013). An in-depth description of the conduct of these focus groups is provided elsewhere (Strout et al., 2017; Vessey et al. 2017).

Results—The demographic characteristics for youth and professional focus group participants are displayed in Table 1. Analysis of the transcripts indicates that each of the three major components of the federal bullying definition was addressed. The end result was an item pool consisting of 38 separate items. In large part, findings from the professional focus groups mirrored those of the youth focus groups. Both youths and professionals supported the concept of unwanted aggressive behavior and clearly distinguished unwanted aggression from other forms of peer aggression. Youth participants identified a range of physical, social, and psychological behaviors that supported the importance of unwanted aggression as a defining characteristic of bullying.

The concept of power imbalance was more abstract for younger youths (those just entering middle school), especially when discussing non-physical forms of bullying. In addition, youths discussed the intent of the instigator and the manner in which the behavior is received by the victim. Younger participants, however, were less able to delve inside the motivations of instigators. Both the concepts of power imbalance and bullies' motivations were more readily embraced by older youths who were cognitively more mature including their capacities for critical, abstract thinking (Steinberg, 2005). Not surprisingly, the professionals readily endorsed both the concepts of power imbalance and the motivation underlying bullying behaviors.

Both youths and professionals endorsed the concept of repetition to a lesser degree. They acknowledged the possibility that bullying could occur as a singular event, as a repeated dynamic between one bully-victim dyad, or as a repeated event between a bully and multiple victims.

At the conclusion of the analysis, concepts that were developed from the focus group dialogs were harmonized with the literature review findings to further elucidate the concept of youth bullying. These findings formed the basis for construction of the logic model displayed as Figure 2. Harmonized findings suggest that the experience of youth bullying, while unique for each individual, is primarily influenced by three factors: the *goals or intent of the instigator*, the recipient's perception of the *presence of an unwanted, aggressive behavior* perpetrated by the instigator that causes the recipient harm or distress, and the recipient's perception of the *presence of a power imbalance* between themselves and the instigator.

These three primary factors intersect, while being influenced by conditions internal and external to the recipient, to produce the recipient's subjective experience of, and response to, youth bullying. Notably, the repetition or frequency of the behavior is not central to this model, acknowledging our participants' beliefs that a single unwanted aggressive act can inflict harm to an individual and that the same act, committed several times, might cause various degrees of harm to different individuals.

Phase III: Concept Selection and Item Construction

The goal of Phase III was to generate a complete initial set of items to be incorporated into the draft of the CABS instrument based on the information gleaned in Phases I and II.

Method—Because the overarching goal was to develop a screening tool appropriate for the rapid identification of youths likely to have been bullied and suffering from its sequelae, initially each conceptual category was considered as to whether potential items that could be derived from it would meet this goal. For example, as the intent of the CABS was to identify victims, items specific to additional bullying roles—such as being the bully or bystander, were no longer considered. Similarly, items on help-seeking behaviors were beyond the scope of a screening instrument designed to detect victimization and were therefore dropped from further discussion at this point.

The CABS' mode of administration, administration instructions, recall period, and response options were also evaluated from the perspectives of measurement and developmental theories (Lerner & Castellino, 2002; Mellor & Moore, 2014). Draft items, instructions, and response options were generated, reviewed, and finalized using a consensus-based process by the investigative team. This included input on wording and design and pilot-testing by a subgroup of youths who participated in the focus groups.

Results—The above process resulted in the development of a pool of 52 potential items for content validity assessment. As prior research has demonstrated that some children, particularly those who are younger or who have poor verbal skills, have limited ability to respond to negatively worded items, all CABS items were phrased positively (Mellor & Moore, 2014). To assess youth bullying, a 5-point, Likert-type response scale was chosen with responses ranging from '*strongly agree*' to '*strongly disagree*.' While some psychometric experts have noted that up to seven response options typically improve internal reliability, research with children has demonstrated that younger children are able to use three to five-point Likert-type scales to accurately report their experiences (Chambers & Johnston, 2002). A recall period of '*in the past month*' was chosen to reflect the CABS goal of identifying young people currently exposed to bullying in order to provide early therapeutic resources.

Phase IV: Content Evaluation/Validity Testing

Evaluating a scale's content validity is an essential step in the early development of an instrument and for improving the construct validity. This provides developers the ability to select items that reflect the construct in the measure or content domain.

Method—The content evaluation phase of this study used a methodological design and included an electronic survey. A panel of national and international experts was identified by the researchers through contacts at professional meetings, review of the literature, membership in related professional organizations and existing professional contacts. The panel was comprised of measurement and content experts who had previously participated in the development of the Uniform Definition, authored relevant psychometric papers, or who had participated in the prior focus groups. Panel member composition represented pediatric healthcare, mental healthcare, elementary and middle school education, youth bullying research, and instrument development.

An invitation was sent to all potential participants using electronic mail. For those expressing interest in participating, a cover letter, set of instructions for completing the review, and a rating tool to evaluate each item and the measure as a whole were provided. The cover letter reviewed the objectives of the study, the risks and potential benefits of participating, and a statement regarding the voluntary nature of the study. The instructions included the Uniform Definition of youth bullying (Gladden et al., 2014) and an explanation of how to evaluate the representativeness, clarity and dimension of each of the potential CABS items. The definition was provided to help the participants conceptualize youth bullying. Experts were asked to rate the relevance and clarity of each item on a four-point ordinal scale; 1=not relevant (or clear), 2=somewhat relevant (or clear), 3=quite relevant (or clear), and 4=highly relevant (or clear) (Davis, 1992).

As youth bullying is influenced by other constructs, participants were also asked to determine which dimension the item represented from the list provided on the survey. The dimensions included those in the Uniform Definition and others identified in the literature search and through the focus groups with youths and professionals. These dimensions were: (1) power imbalance, (2) behavior repeated over time, (3) unwanted, aggressive behavior, (4) school connectedness, (5) effects of bullying, (6) the bully or instigator's intent, (7) disconnectedness at school, and (8) unable to classify. Participants were also provided with the opportunity to include comments and suggestions on each item as well as on the survey as a whole. In addition, they were encouraged to provide recommendations for adding, deleting and modifying items on the survey and describing any barriers or challenges they had encountered during their own instrument development work, when appropriate. As the survey was conducted using the Qualtrics™ platform, all responses were completed electronically. Figure 3 provides a sample survey item as presented to the expert panels.

Statistical Analysis: Statistical evaluation of data generated from the responses of the experts included computation of item- and scale-level content validity indices (CVI), the factorial validity index (FVI), and modified kappa as an assessment of interrater agreement. Data were analyzed using SPSS for Windows version 24.0 (SPSS, Inc., Chicago, IL) and MedCalc, version 9.6.4.0 (Frank Schoonjans, Belgium).

The item-level CVI, defined as the proportion of experts who rate the content as valid (relevance ratings of 3 or 4), was calculated for each item (Polit, Beck & Owen, 2007). Items were rated as being *excellent* when the I-CVI value was greater than 0.78 (Fleiss, 1981). Fleiss' kappa statistic (κ), an index of agreement among experts regarding the

relevance of an item, was calculated to correct for chance agreement (Fleiss, 1981; Polit et al., 2007). Items considered *excellent* ($\kappa = 0.74$, I-CVI = 0.78) were retained for future full psychometric evaluation (Fleiss, 1981). Items on the threshold ($\kappa = 0.74$, I-CVI = 0.75) were individually assessed. For complete scale validation, all I-CVI values were averaged to compute a scale-level content validity index (S-CVI), for which a value greater than 0.90 is considered *excellent* (Polit et al., 2007).

The factorial validity index (FVI) is a measure of the degree to which survey respondents identify individual items as belonging to their intended dimensions. This provides a preliminary indication of the factorial validity of the measure. For each item, the number of respondents who correctly associated the item and dimension was divided by the total number of respondents. As for the scale-level CVI, the average was calculated across the items to determine FVI for the measure as a whole (Rubio, Berg-Weger, Tebb, Lee, & Rauch, 2003).

Results—Of the 46 experts invited participants, 30 (65.2%) agreed to participate; 12/22(55 %) of the invited content experts and 18/24(75%) of the invited measurement experts. Respondents were primarily female (n = 21, 70%), non-Hispanic (n = 30, 100%), and white (n = 26, 87%). Experts worked as professors (n=16, 53.3%); were middle school professionals (n=5, 16.6%); were violence prevention organization professionals (n=4, 13.3%), public health professionals (n=1, 3.3%), healthcare professionals (n=3, 10%) and Parent Teacher Association leaders (n=1, 3.3%). Experts participated from diverse locations including Australia, Italy, the United Kingdom, Norway, Canada and the United States.

Content validity indices for clarity and representativeness are summarized in Table 2. All 52 draft items received clarity and representativeness feedback from at least 25 of 30 participating experts. Three draft items (#1, “*I am close to people at my school;*” #32, “*I wish my classmates would leave me alone;*” #38, “*I enjoy teasing or goofing around with my friends and classmates*”) received item-level ratings for clarity below the 0.78 threshold used in the study. All other items were considered to be clear (CVI = 0.78) by the expert panel. Twenty-two items received ratings between 0.23 and 0.77 for representativeness, below the 0.78 threshold. In sum, three items received below-threshold ratings for both clarity and representativeness and expert qualitative comments did not support item revision; these items were excluded from further consideration (#1, ‘*I am close to people at school;*’ #32, ‘*I wish my classmates would leave me alone;*’ and #38, “*I enjoy teasing and goofing around with my friends or classmates*’). Twenty-seven items with acceptable clarity but low representativeness were judged by the panel as being representative of a related concept such as school connectedness or bullying sequelae. For example, participants identified items #3 “*I feel like I fit in with other kids at my school,*” #10 “*I feel happy at school,*” and #7 “*I have at least one friend who looks out for me at school,*” as representing school connectedness. These 27 items were excluded from further analysis.

Two pairs of items were identified as having overlapping content and these pairs were merged into two, rather than four, items. For example, item #50 “*I have nightmares about being bullied,*” and #51 “*I have upsetting flashbacks about being bullied,*” were combined to create a single item, “*I have distressing memories of being bullied.*” Eighteen other items

were revised using experts' suggestions for improving clarity, readability, and representativeness, ultimately yielding 20 items with scale-level CVIs of 0.954 for clarity and 0.920 for representativeness. Experts had difficulty correctly matching individual items with their intended domains; FVIs ranged from 0.12 – 1.00 for individual items. Consistent with feedback from the panel, two new items were added to assess for the presence of power imbalance and to assess for repetition of bullying behavior. These additions resulted in a 22-item tool for further psychometric evaluation.

Discussion

This study describes the four-phase process for item generation and content validity testing of a new measure, the CABS, designed to screen for self-reported exposure to bullying behavior in school age youths interfacing with the healthcare system.

Our review of literature identified a plethora of studies discussing the antecedents and consequences of bullying victimization with a smaller cadre of studies reporting on instruments designed to measure youth bullying; most were devoid of any theoretical framework and methodologic rigor was highly varied. The lack of definitional and measurement clarity present in this body of literature has been noted to be a challenge to moving the science of youth bullying forward, both by limiting the validity of measures and studies employing them and by making comparisons across studies difficult (Casper, Meter, & Card, 2015; Green, Felix, Sharkey, Furlong, & Kras, 2013; Volk et al., 2017). Therefore, a critical component of the instrument development process used here was to clearly define the concept that is the target of the new measure, in this case, youth bullying.

While critically important to developing meaningful measures, defining bullying is notably challenging (Gladden, et al., 2014; Thornberg, 2015; Volk, et al., 2017). The most widely used youth bullying definition was proposed by Dan Olweus: "it is aggressive behavior or intentional 'harm doing,' which is carried out repeatedly and over time in an interpersonal relationship characterized by an imbalance of power" (1993). Despite its widespread use, empiric evidence supporting the validity of the Olweus definition or supporting its congruity with existing research is not readily available (Greif & Furlong, 2006). In addition, evidence on youths' own views of youth bullying have not been routinely considered nor do they generally support use of all components of the original Olweus definition. This creates a mismatch between what youths are experiencing and what researchers and clinicians are measuring (Green et al., 2013; Hellstrom, Persson, & Hagquist, 2015; Vaillancourt, et al., 2008).

To prevent this issue, robust instrument development processes require that qualitative approaches be incorporated into early development phases as these methods are best suited to capture the voices of all populations of interest. Focus group data from youths of varying backgrounds and professionals from diverse disciplines captured their distinct but harmonic views of the youth bullying experience that was essential to item development, using the Uniform Definition as the touchstone for this process (Gladstone et al., 2014).

The results of the literature reviews and focus groups indicate that using primary concepts—aggression, power imbalance, and repetition—articulated in the Uniform Definition are requisite but insufficient. These three concepts have been central in seminal research and legislative definitions of bullying (Olweus, 1993; Sacco et al., 2012) and often form the basis of school-based bullying prevention programs. Youths and professionals who work with them, however, hold different views of power imbalance and repetition and other concepts such as intent that are important to the populations of interest frequently have gone unaddressed in instrument development.

These findings underscore the need to craft items designed to capture this concept with care. For example, with theoretical support for bullying as a goal-oriented behavior used by those who bully to achieve concrete outcomes, we included items reflecting youths' perceptions of bullies' goals in the CABS (Veenstra, Lindenberg, Munniksmma, & Dijkstra, 2010; Volk, Camilleri, Dane & Marini, 2012). In another example youths reported that their schools' definitions of bullying required a specific behavior(s) be repeated a certain number of times before it could be classified as bullying. This provided our team with a particularly poignant illustration of how bullying definitions and measures translated into local school policy can potentially result in missed cases and underreporting based upon the 'repetition' criteria. With this in mind, the CABS instrument asks only one frequency item and does not require a particular frequency to classify an individual as having been exposed to bullying.

Prior to content validity testing, one concern addressed by the researchers was in identifying the expert panel. It became clear that three groups of experts--professionals who directly work with youths of diverse backgrounds, researchers with expertise in bullying content, and methodologists--were all needed for the content evaluation of the initial CABS items. Collectively, they addressed the strengths and weaknesses of the potential items' content and instrument design and execution. The result is the first tool specifically designed for use in healthcare screening. It is grounded in our logic model of the youth bullying experience and items reflect the components of the model. The logic model is also consistent with the Uniform Definition of bullying, but does not exclude other closely related definitions. In sum, the results of this study provide support for the content validity of the new CABS tool.

Conclusions

Using a systematic approach, a new measure of youth bullying is being developed specifically for use in healthcare settings to help identify middle school youths affected by peer bullying. The CABS is grounded in the three components of the current Uniform Definition of bullying: power imbalance, unwanted aggressive behavior, and repetition (Gladden et al., 2014). Such measures are needed to help establish uniformity across bullying surveillance activities needed, in part, to help advance the development of meaningful interventions. Additional research is warranted to evaluate the psychometric properties and underlying structure of the newly developed CABS instrument in a large sample of youths presenting to the healthcare setting. Further refinement of the tool, for example by reducing the number of items, may be necessary to produce a measure that is both accurate and appropriate for use in the busy healthcare setting.

Acknowledgments

Source of Funding: Research reported in this publication was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number R21HD083988. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Burns, N., Grove, SK. The concepts of measurement. In: Burns, N., Grove, SK., editors. The practice of nursing research: Appraisal, synthesis, and generation of evidence. 6. St. Louis, MO: Elsevier Saunders; 2009. p. 371-394.
- Casper DM, Meter DJ, Card NA. Addressing measurement issues related to bullying involvement. *School Psychology Review*. 2015; 44:353–371. DOI: 10.17105/spr-15-0036.1
- Chambers CT, Johnston C. Developmental differences in children's use of rating scales. *Journal of Pediatric Psychology*. 2002; 27:27–36. DOI: 10.1093/jpepsy/27.1.27 [PubMed: 11726677]
- Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *American Journal of Mental Deficiency*. 1981; 86:127–137. [PubMed: 7315877]
- Committee on Injury, Violence, and Poison Prevention. Policy Statement—Role of the Pediatrician in Youth Violence Prevention. *Pediatrics*. 2009; 124:393–402. [PubMed: 19520726]
- Committee on Practice and Ambulatory Medicine, Bright Futures Periodicity Schedule Workgroup. 2017 Pediatric preventive health care. *Pediatrics*. 2017; 139:e20170254.doi: 10.1542/peds.2013-4096 [PubMed: 28213605]
- Davis LL. Instrument review: getting the most from a panel of experts. *Applied Nursing Research*. 1992; 5:104–107. DOI: 10.1016/S0897-1897(05)80008-4
- Fekkes M, Pijpers FI, Fredriks AM, Vogels T, Verloove-Vanhorick SP. Do bullied children get ill, or do ill children get bullied? A prospective cohort study on the relationship between bullying and health-related symptoms. *Pediatrics*. 2006; 117:1568–1574. DOI: 10.1542/peds.2005-0187 [PubMed: 16651310]
- Finfeld-Connett D, Johnson ED. Literature search strategies for conducting knowledge - building and theory - generating qualitative systematic reviews. *Journal of Advanced Nursing*. 2013; 69:194–204. DOI: 10.1177/1054773812468755 [PubMed: 22591030]
- Fleiss, J. Statistical methods for rates and proportions. 2. New York, NY: John Wiley; 1981.
- Gladden, RM., Vivolo-Kantor, AM., Hamburger, ME., Lumpkin, CD. Bullying surveillance among youths: Uniform definitions for public health and recommended data elements, version 1.0. Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention, U.S. Department of Education; 2014.
- Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*. 2009; 26:91–108. DOI: 10.1111/j.1471-1842.2009.00848.x [PubMed: 19490148]
- Green JG, Felix ED, Sharkey JD, Furlong MJ, Kras JE. Identifying bully victims: Definitional versus behavioral approaches. *Psychological Assessment*. 2013; 25:651–657. DOI: 10.1037/a0031248 [PubMed: 23244644]
- Greif JL, Furlong MJ. The assessment of school bullying: Using theory to inform practice. *Journal of School Violence*. 2006; 5:33–50. DOI: 10.1300/J202v05n03_04
- Healthy People 2020. Violence and injury prevention. 2012. Aug 29. Retrieved from <http://www.healthypeople.gov/2020/topicsobjectives2020/objectiveslist.aspx?topicId=24>
- Hellstrom L, Persson L, Hagquist C. Understanding and defining bullying – adolescents' own views. *Archives of Public Health*. 2015; 73:4.doi: 10.1186/2049-3258-73-4 [PubMed: 25973194]
- Hinkin TR. A review of scale development practices in the study of organizations. *Journal of Management*. 1995; 21(5):967–988.
- Hinkin TR. A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*. 1998; 2:104–121. DOI: 10.1177/109442819800100106

- Krippendorff, K. Content analysis: An introduction to its methodology. 3. Thousand Oaks, CA: Sage; 2013.
- Lerner RM, Castellino DR. Contemporary developmental theory and adolescence: Developmental systems and applied developmental science. *Journal of Adolescent Health*. 2002; 31:122–135. [org/10.1016/S1054-139X\(02\)00495-0](https://doi.org/10.1016/S1054-139X(02)00495-0). [PubMed: 12470909]
- Lyznicki JM, McCaffree MA, Robinowitz CB. Childhood bullying: implications for physicians. *American Family Physician*. 2004; 70(9):1723–1728. [PubMed: 15554490]
- Mellor D, Moore KA. The use of Likert scales with children. *Journal of Pediatric Psychology*. 2014; 39:369–379. DOI: 10.1093/jpepsy/jst079 [PubMed: 24163438]
- Moher D, Liberati A, Tetzlaff J, Altman DG. PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine*. 2009; 6:e1000097.doi: 10.1371/journal.pmed.1000097 [PubMed: 19621072]
- Nakamoto J, Schwartz D. Is peer victimization associated with academic achievement? A meta-analytic review. *Social Development*. 2010; 19:221–242. DOI: 10.1037/a0017199
- Olweus, D. *Bullying at school: What we know and what we can do*. Oxford, UK: Blackwell; 1993.
- Polit, DF., Beck, CT. *Appraising evidence for nursing practice*. 8. New Delhi, India: Wolters Kluwer; 2014. *Essentials of nursing research*.
- Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing and Health*. 2007; 30:459–467. DOI: 10.1002/nur.20199 [PubMed: 17654487]
- Reijntjes A, Kamphuis JH, Prinzie P, Boelen PA, Van der Schoot M, Telch MJ. Prospective linkages between peer victimization and externalizing problems in children: A meta-analysis. *Aggressive Behavior*. 2011; 37:215–222. DOI: 10.1002/ab.20374 [PubMed: 21433031]
- Reijntjes A, Kamphuis JH, Prinzie P, Telch MJ. Peer victimization and internalizing problems in children: A meta-analysis of longitudinal studies. *Child Abuse & Neglect*. 2010; 34:244–252. DOI: 10.1016/j.chiabu.2009.07.009 [PubMed: 20304490]
- Rubio DM, Berg-Weger M, Tebb SS, Lee ES, Rauch S. Objectifying content validity: conducting a content validity study in social work research. *Social Work Research*. 2003; 27(2):94–104.
- Sacco, D., Silbaugh, KB., Corredor, F., Casey, J., Doherty, D. An overview of state anti-bullying legislation and other related laws. Berkman Center Research Publication No. 2013–4. 2012. Feb 23. [doi.org/10.2139/ssrn.2197961](https://ssrn.com/abstract=2197961) Retrieved from: <https://ssrn.com/abstract=2197961>
- Spector ND, Kelly SF. Pediatrician's role in screening and treatment: bullying, prediabetes, oral health. *Current Opinion in Pediatrics*. 2006; 18:661–70. DOI: 10.1097/MOP.0b013e328010a2a7 [PubMed: 17099367]
- Steinberg L. Cognitive and affective development in adolescence. *Trends in Cognitive Sciences*. 2005; 9:69–74. DOI: 10.1016/j.tics.2004.12.005 [PubMed: 15668099]
- Stewart, DW., Shamdasani, PN. *Focus groups: Theory and practice*. Newbury Park, CA: Sage; 2014.
- Strout TD, DiFazio RL, Vessey JA. Technology-enhanced focus groups as a component of instrument development: experiences from the CABS study. *Nurse Researcher*. 2017; 25:16–23. DOI: 10.7748/nr.2017.e1458 [PubMed: 28639521]
- Thornberg R. The social dynamics of school bullying: The necessary dialogue between the blind men around the elephant and the possible meeting point at the social ecological square. *Confero: Essays on Education, Philosophy, and Politics*. 2015; 3(2):161–203. DOI: 10.3384/confero.20014562.150624
- Ttofi MM, Farrington DP, Lösel F. School bullying as a predictor of violence later in life: A systematic review and meta-analysis of prospective longitudinal studies. *Aggression and Violent Behavior*. 2012; 17:405–418. DOI: 10.1016/j.avb.2012.05.002
- Uddin, SG., O'Connor, KS., Ashman, JJ. NCHS data brief, no 248. Hyattsville, MD: National Center for Health Statistics; 2016. Physician office visits by children for well and problem-focused care: United States, 2012. Available from: <https://www.cdc.gov/nchs/data/databriefs/db248.pdf>
- U.S. Department of Education. Student reports of bullying: Results from the 2015 School Crime Supplement to the National Crime Victimization Survey. 2016. Dec. Retrieved from: <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2017015>

- U.S. Health Resources Services Administration, Maternal Child Health. An evidence-informed and evidence-based Bullying Prevention Capacity Assessment and Change Package. Washington, D.C: Author; 2017. Aug. Assessing intervention capacity and implementing change. Retrieved from: https://mchb.hrsa.gov/sites/default/files/mchb/MaternalChildHealthInitiatives/MCHB_ChangePkg_8-24-17_sxf.pdf
- U.S. Preventive Services Task Force. Published Recommendations. 2018. Retrieved from: <https://www.uspreventiveservicestaskforce.org/BrowseRec/Index?age=Pediatric,Adolescent>
- Vaillancourt T, McDougall P, Hymel S, Krygsman A, Miller J, Stiver K, Davis C. Bullying: Are researchers and children/youth talking about the same thing? *International Journal of Behavioral Development*. 2008; 32:486–495. DOI: 10.1177/0165025408095553
- Veenstra R, Lindenberg S, Munniksma A, Dijkstra JK. The complex relation between bullying, victimization, acceptance, and rejection: Giving special attention to status, affection, and sex differences. *Child Development*. 2010; 81:480–486. DOI: 10.1111/j.1467-8624.2009.01411.x [PubMed: 20438454]
- Vessey JA, DiFazio RL, Strout TD. Increasing meaning in measurement: A Rasch analysis of the child-adolescent teasing scale. *Nursing Research*. 2012; 61:159–170. DOI: 10.1097/NNR.0b013e3182504a88 [PubMed: 22551990]
- Vessey JA, DiFazio RL, Strout TD. Youth bullying: A review of the science and call to action. *Nursing Outlook*. 2013; 61:347–345. DOI: 10.1016/j.outlook.2013.04.011
- Vessey JA, DiFazio RL, Strout TD. ‘I didn’t even know you cared about that stuff.’ Youths’ perceptions of health care provider roles in addressing bullying. *Journal of Pediatric Healthcare*. 2017; Epub ahead of print. doi: 10.1016/j.pedhc.2017.01.004
- Vessey JA, Strout TD, DiFazio RL, Walker A. Measuring the youth bullying experience: A systematic review of the psychometric properties of available instruments. *Journal of School Health*. 2014; 84:819–843. DOI: 10.1111/josh.12210 [PubMed: 25388599]
- Vessey JA, Lulloff A, Hernandez-Meneses L, Strout T, DiFazio RL. Implications of youth bullying on mental health. *Revista Horizonte De Enfermeria*. 2016; 27(1):9–23.
- Volk AA, Camilleri J, Dane AV, Marini ZA. Is adolescent bullying an evolutionary adaptation? *Aggressive Behavior*. 2012; 38:222–238. DOI: 10.1002/ab.21418 [PubMed: 22331629]
- Volk AA, Dane AV, Marini ZA. What is bullying? A theoretical re-definition. *Developmental Review*. 2014; 34:327–343. DOI: 10.1016/j.dr.2014.09.001
- Volk AA, Veenstra R, Espelage DL. So you want to study bullying? Recommendations to enhance the validity, transparency, and compatibility of bullying research. *Aggression and Violent Behavior*. 2017; 36:34–43. DOI: 10.1016/j.avb.2017.07.003
- Wilson, JMG., Jungner, G. Principles and practice of screening for disease. Geneva, Switzerland: World Health Organization; 1968. Retrieved from: http://whqlibdoc.who.int/php/WHO_PHP_34.pdf
- Whittemore R, Knafl K. The integrative review: updated methodology. *Journal of Advanced Nursing*. 2005; 52:546–553. DOI: 10.1111/j.1365-2648.2005.0362 [PubMed: 16268861]

Figure 1.
4-Phase development of the Child-Adolescent Bullying Scale (CABS).

Figure 2.

Logic model for the development of the CABS. We propose that the experience of youth bullying, while unique per individual, is primarily influenced by several factors including the *goals and intent of the instigator*, the recipients perception of the *presence of an unwanted aggressive behavior* perpetrated by the instigator, and the recipients perception of the *presence of a power imbalance* between themselves and the instigator – these factors intersect, along with conditions internal and external to the recipient, to produce the recipient's experience and response to the bullying experience.

Bullying Concept and Definition

For this study, **bullying is defined as**, “Any unwanted aggressive behavior(s) by another youth or group of youths who are not siblings or current dating partners that involves an observed or perceived power imbalance and is repeated multiple times or is highly likely to be repeated.” We will ask you to **use this bullying definition** as you complete the survey.

In addition, please conceptualize youth bullying as including three important components as you complete this survey:

- Bullying involves **unwanted aggressive behaviors**
- Bullying involves an **imbalance of power or strength** that may be real or perceived
- Bullying involves a pattern of behavior **repeated over time**.

Survey Instructions

As you move through this survey, you will be presented with each potential screening tool item **three times**:

- Once to consider its **representativeness**, or how well the item characterizes youth bullying or a related concept,
- Once to consider **clarity**, or how easy to understand or not confusing the item is, and
- Once to consider which **bullying-related domain** the item may represent. Domains include: power imbalance, behavior repeated over time, unwanted negative behavior, school connectedness, effects of bullying, bully's intent, and disconnectedness at school.

Please **respond to each of these questions** and then provide any **comments or suggestions** you might have for revision of the item in the space below the item.

Click for bullying definition

Instructions: For the item below, please choose the response that best describes the **degree to which you feel this item represents youth bullying**.

	The item is not representative	The item needs major revision to be representative	The item needs minor revision to be representative	The item is representative
Kids try to make me feel bad on purpose.	☐	☐	☐	☐

Instructions: For the item below, please choose the response that best describes the **how clear you feel this item is**.

	The item is not clear	The item needs major revision to be clear	The item needs minor revision to be clear	The item is clear
Kids try to make me feel bad on purpose.	☐	☐	☐	☐

Instructions: For the item below, please choose the **bullying-related domain** that you feel this item best represents.

Item: Kids try to make me feel bad on purpose.

☐	Power imbalance
☐	Behavior repeated over time
☐	Unwanted negative behavior
☐	School connectedness
☐	Effects of bullying
☐	Bully's intent
☐	Disconnectedness at school
☐	Unable to classify

Instructions: For the item below, please provide any additional **comments, suggestions, or revisions**.

Item: Kids try to make me feel bad on purpose.

Figure 3.
Sample Child-Adolescent Bullying Scale (CABS) expert questionnaire items.

Table 1**Demographic Characteristics of Focus Group Participants**

Youths (N = 24)		
Age (Mean±SD)	12.23±1.65	
Gender	N	%
Male	6	25%
Female	18	75%
Race		
White	16	66.7%
African American or Black	3	12.5%
Asian or Pacific Islander	4	16.7%
More than one race	1	4.1%
Ethnicity		
Non-Hispanic	19	79.2%
Hispanic	5	20.8%
Grade in School		
5 th Grade	2	8.3%
6 th Grade	4	16.7%
7 th Grade	9	37.5%
8 th Grade	4	16.7%
9 th Grade	2	8.3%
10 th Grade	2	8.3%
11 th Grade	1	4.2%
Median Income of Participants' Communities		
\$20,000 – \$50,000	7	29.1%
\$50,999 – \$90,000	8	33.3%
\$90,999 – \$120,000	8	33.3%
\$120,999–\$150,000	1	4.3%
Professionals (N = 9)		
Gender	N	%
Male	3	33.3%
Female	6	66.6%
Race		
White	8	88.8%
Asian or Pacific Islander	1	11.1%
Ethnicity		
Non-Hispanic	9	100%
Profession		
School Principal (Elementary and Middle)	2	22.2%
Social Worker	2	22.2%
School Guidance Counselor	1	11.1%
Nurse Practitioner	2	22.2%

Youths (N = 24)		
Nurse	1	11.1%
Middle School Teacher	1	11.1%

Table 2
Summary of Content Validity Indices for Clarity and Representativeness for Potential CABS Items

Item Number	Item	Clarity		Representativeness	
		I-CVI	Experts Agreeing	I-CVI	Experts Agreeing
1	close at school	0.67	20/30	0.23	7/30
2	get along	0.86	25/29	0.28	8/29
3	part of school	0.93	26/28	0.29	8/28
4	fit at school	0.93	25/27	0.26	7/27
5	comfortable at school	0.96	27/28	0.36	10/28
6	friend at lunch	1.00	29/29	0.41	12/29
7	best friend	0.97	28/29	0.34	10/29
8	friend looks out	1.00	28/28	0.39	11/28
9	they like me	0.93	26/28	0.39	11/28
10	feel happy	0.90	26/29	0.38	11/29
11	school is safe	1.00	28/28	0.55	16/29
12	kids make feel bad	0.93	26/28	0.89	25/28
13	kids are mean	0.93	26/28	0.89	25/28
14	make fun of	0.93	26/28	0.96	27/28
15	turn others	0.96	26/27	0.96	26/27
16	tried to get in trouble	0.85	23/27	0.70	19/27
17	you bully	0.86	24/28	0.89	24/27
18	see others bullied	0.93	26/28	0.89	25/28
19	feel bullied	0.88	24/27	0.93	25/27
20	damaged property	0.96	26/27	0.89	24/27
21	threatened physically	0.96	27/28	0.93	26/28
22	fight with mates	0.86	24/28	0.57	16/28
23	physically unsafe	0.93	25/27	0.81	22/27
24	unsafe at school	0.93	26/28	0.86	24/28
25	unsafe outside	0.93	26/28	0.89	25/28
26	unsafe after school	0.86	24/28	0.89	25/28
27	emotionally unsafe	0.88	23/26	0.85	22/26

Item Number	Item	Clarity		Representativeness	
		I-CVI	Experts Agreeing	I-CVI	Experts Agreeing
28	not go to school	1.00	26/26	0.88	23/26
29	not go to activities	0.96	25/26	0.88	23/26
30	different school	0.96	25/26	0.81	21/26
31	avoid school	1.00	26/26	0.85	22/26
32	leave me alone	0.73	19/26	0.50	13/26
33	loser	0.85	23/27	0.52	14/27
34	lonely at school	0.88	23/26	0.54	14/26
35	kids ignore	0.96	25/26	0.92	24/26
36	mean online	1.00	27/27	0.96	26/27
37	unsafe online	0.92	24/26	0.85	22/26
38	enjoy goofing	0.77	20/26	0.35	9/26
39	not bothered teasing	0.85	22/26	0.42	11/26
40	bothered teasing	1.00	26/26	0.92	24/26
41	teasing at school	1.00	25/25	0.88	22/25
42	paying attention	0.96	25/26	0.81	21/26
43	worry about	0.96	25/26	0.81	21/26
44	talk behind back	0.92	24/26	0.88	23/26
45	parents worry	0.88	23/26	0.65	17/26
46	teachers fair	0.85	22/26	0.35	9/26
47	care what they think	0.81	21/26	0.27	7/26
48	exclude me	0.96	25/26	0.88	23/26
49	I am different	0.92	24/26	0.92	24/26
50	nightmares	0.92	24/26	0.77	20/26
51	flashback	0.96	25/26	0.81	21/26
52	distressing memories	0.96	25/26	0.81	21/26

Note. I-CVI = item-level content validity index.