ABSTRACT. This article examines the conceptual basis for and methods used to assess school bullying, including the core bullying behavior elements of repetition, intentionality, and power differential and instruments needed to foster comparability across studies and to improve the precision of intervention capacity. Common bully selfreport procedures (Olweus Bully/Victim Questionnaire, Olweus, 2004; Reynolds Bully Victimization Scale, Reynolds, 2003; The Bully Surveys, Swearer, 2001) are examined for the thoroughness with which they assess these core elements that distinguish bullying from other forms of peer victimization. It is concluded that bullying assessment can be enhanced by systematically including all core bully behaviors, more thoroughly examining sources of power differential between bullies and victims, and giving more attention to the dynamic nature of the bullying process.

Bullying is recognized as a significant problem affecting youth in American schools (Colvin, Tolin, Beard, Hagan, & Sprague, 1998; Espelage & Swearer, 2003; Nansel et al., 2001). It has been linked with school and community violence (e.g., Nansel, Overpeck, Haynie, Ruan, & Scheidt, 2003) and bullying victimization has been identified as a common factor among school shooters (Newman, Fox, Harding, Mehta, & Roth, 2004). Researchers have hypothesized that bullying is a component of a cycle of violence that begins with bullying at the elementary and junior high school levels, and then can progress to dating violence, harassment, and assault at the high school level (Colvin et al,,1998; Stein, 2003). In adulthood, aggression may display itself in the form of child abuse, hate crimes, or other domestic violence, which is passed on to the next generation of children who react by acting out as bullies at school (Colvin et al., 1998). There is consensus among researchers that chronic bully victimization is associated with a range of negative academic, social, and emotional outcomes (e.g., Hawker & Boulton, 2000; Nickerson, Brock, Change, & OMalley, 2005); hence, it demands the attention of educators.

The quality of understanding of youth experiences with bullying and peer victimization hinges on the ability to effectively assess these constructs. Since the inception of bullying research, various assessments have been proposed to measure this form of peer victimization. Despite the availability of measures, an assessment originally developed by Dan Olweus in 1978 (and more recently modified) has remained the most popular method for measuring bullying (Pellegrini, 2001). For example, the World Health Organization recently administered a version of Olweus measure to over 15,000 American youth (Nansel et al., 2001). Although the Olweus Bully/Victim Questionnaire (original and revised versions) has been widely used and modified in a large number of studies, other measures have been proposed that also show promise (e.g., Reynolds, 2003; Swearer, 2001).

Bullying is a specific type of peer victimization and, as such, requires special attention to measurement. Because of the repetitive nature of bullying and the power differential that makes it nearly impossible for victims to defend themselves, bullying is an especially harmful form of victimization. As such, bullying has a special status as a form of victimization, necessitating that it is clearly distinguished from other types of peer victimization.

This paper presents key theoretical issues in the effective assessment of bullying. In particular, it focuses on how selfreport assessments measure victimization.1 Although there are many psychometric issues related to measuring bullying, our focus is solely on theoretical topics. More specifically, this paper considers important issues in developing and selecting assessments that include the following: the purpose(s) of the assessment, the definition of bullying presented to students, and how the assessment conceptualizes bullying.

PURPOSE OF ASSESSMENT

Assessments of bullying have various functions and the intended implications of the assessment can be used to select the appropriate measurement tool. Most often, bullying assessments are used to gather prevalence rates and determine national trends. At a macrolevel, this information impacts the development of laws and policies on the state and federal scale (Limber & Small, 2003). At a local level, schools may use prevalence data for school safety planning or to facilitate buyin from school staff, students, and parents when implementing interventions. When using an assessment to gain a general sense for rates of bullying, it is important that the assessment is brief and can capture many forms of victimization.

At an individual level, assessments of bullying can be used to understand the experiences of students. A comprehensive picture of the victims experiences can inform secondary prevention and interventions efforts. Measures designed to assess individual experiences will most likely be longer and more comprehensive. They may ask questions along the lines of a functional behavioral assessment to identify factors that maintain the bullying process and develop plans to stop its reoccurrence (Colvin et al., 1998). The purpose of the assessment guides the selection of appropriate methods and questions to best address the predetermined goal.

DEFINITIONS IN ASSESSMENT

Perhaps the most important theoretical issue to consider when evaluating assessments of bullying is how the measure defines the unique aspects of bullying as a subset of all peer victimization (see Nickerson et al., 2005). There is some variability in accepted definitions of bullying and there are certainly discrepancies in the extent to which measures are faithful to these definitions. The most popular definition of bullying was developed by Olweus (1978) and defines bullying as peer aggression that incorporates three components: (1) repetition over time, (2) intentionally designed to harm the victim, and (3) involving an imbalance of power between the person doing the bullying and the person being bullied. This definition is accepted by most researchers in the field (e.g., Espelage & Swearer, 2003), with some investigators suggesting that there are additional components: (4) bullying is unprovoked by the victim, and (5) bullying occurs within social groups when other peers are present (cited by Griffin & Gross, 2004). Issues related to how well assessments include each of the definitional components will be considered next.

Measures of bullying often begin by providing a definition of bully

ing to students, and then asking a series of questions about whether the student has experienced bullying, as described. For example, the Revised Olweus Bully/Victim Questionnaire provides the following definition of bullying that is designed to incorporate the three major components of the term:

We say a student is being bullied when another student or several

other students

say mean and hurtful things or make fun of him or her or call him or her mean and hurtful names;

completely ignore or exclude him or her from their group of friends or leave him or her out of things on purpose;

hit, kick, push, shove around, or threaten him or her;

tell lies or spread false rumors about him or her or send mean notes and try to make other students dislike him or her; and

do other hurtful things like that.

These things may take place frequently, and it is difficult for the student being bullied to defend himself or herself. It is also bullying when a student is teased repeatedly in a mean and harmful way.

But we dont call it bullying when the teasing is done in a friendly and playful way. Also it is not bullying when two students of about the same strength or power argue or fight. (Solberg & Olweus, 2003, p. 246)

This description is designed to foster a common understanding of the term bullying among youth that incorporates all components of the definition of bullying. In addition, it offers examples to explain the types of behaviors that the investigator would like students to attend to, while providing enough ambiguity for students to consider other hurtful behaviors that are not explicitly listed in the definition.

There are, however, several limitations to using a definition such as the one provided above and asking youth to rate the frequency with which they have experienced bullying. Some of these limitations are not specific to bullying assessments. In their review of recommendations for assessing child victimization, Hamby and Finkelhor (2000) advocate using short sentences, easy vocabulary, and simple grammar. They refer to a study that found that 80% of 9 to 11yearolds can comprehend questions comprising 1 to 9 words, but this percentage of comprehension dropped to 40% of children when the questions were more than 20 words long.