So you want to study bullying? Recommendations to enhance the validity, transparency, and compatibility of bullying research

Anthony A. Volk, René Veenstra, Dorothy L. Espelage

Department of Child and Youth Studies, Brock University, 1812 Sir Isaac Brock Way, St. Catharines, ON L2S3A1, Canada
Department of Sociology, University of Groningen, Grote Rozenstraat 31, 9712 TG Groningen, the Netherlands
Department of Psychology, University of Florida, 945 Center Drive, Gainesville, FL 32611-2250, United States

ABSTRACT

Bullying is a serious problem that affects millions of individuals worldwide each year. In response to this, thousands of research articles have been published on bullying. Unfortunately, much of bullying research remains largely atheoretical in its approach to defining bullying as a unique form of aggression. Another key problem in bullying research is the proliferation of heterogeneity of bullying measures whose validity is sometimes questionable. Combined, these two problems have made progress difficult as comparisons between studies and results are impeded by a lack of commonality. As a solution to these problems a discussion of the issues surrounding defining and measuring bullying is offered. This paper aims to promote thoughts and insights about the critical issues and concepts facing those who seek to define and measure bullying for research, intervention, or policy work. Although suggestions for best practices are offered, the overriding goal is to promote all practices that enhance the validity, transparency, and compatibility of bullying research. The time seems right for a general call to action for researchers to individually produce data that are both theoretically and empirically more communicable to the broader bullying community.

1. Introduction

Bullying is a serious topic that has generated steadily increased attention from governments and the general public (Hymel & Swearer, 2015). It affects millions of people each year across virtually all known cultures and social domains, including schools, family homes, recreational activities, work, and prisons (Monks et al., 2009; Volk, Camilleri, Dane, & Marini, 2012; Volk & Lagzdins, 2009). As a result, bullying research has increased exponentially in the last decades. PsycINFO lists 27 peer-reviewed journal articles in response to a search of bully* in the first 80 years of the 20th century, 29 in the 1980s, 275 in the 1990s, and 1898 in the first decade of the new millennium. This exponential trend has continued, with over 5000 peer-reviewed articles on bullying being published in the preceding six years (PsycINFO, February 2017). By any standard then, bullying is an important, global, phenomenon that has generated strong and growing interest amongst researchers.

In large part, this is because of growing evidence that there are significant long-term adverse outcomes for victims (e.g., Copeland, Wolke, Angold, & Costello, 2013; Copeland et al., 2014; Kretschmer et al., 2017; Wolke, Copeland, Angold, & Costello, 2013) and to a lesser extent bullies (Bender & Lösel, 2011; Kretschmer et al., 2017; Ttofi, Farrington, Lösel, & Loeber, 2011). In response to this evidence, researchers have attempted to identify the causes of bullying to develop intervention programs that target the risk and protective factors associated with bullying involvement (Iudici & Faccio, 2014; Olweus, 1993; Rigby, 2010; Ttofi & Farrington, 2011). Thus, bullying remains a challenging behavior to study adequately. This is due in part to the very nature of bullying as a behavior that aims to avoid detection and sanctions from authority figures, meaning that it is often deliberately hidden or misrepresented (e.g., Craig & Pepler, 1998). Beyond this concealment, bullying is a complex behavior that often requires an intimate understanding of the social dynamics of youth, peer groups, and schools to properly observe and interpret the behaviors and outcomes (Rodkin, Espelage, & Hanish, 2015; Salmivalli, 2010; Sutton, Smith, & Swettenham, 1999).

There are many challenges to conducting research on bullying, including many different ways of defining (e.g., Nansel & Overpeck, 2003; Vaillancourt et al., 2008) and measuring bullying (e.g., Casper, Meter, & Card, 2015; Hamburger, Basile, & Vivolo, 2011). This heterogeneity in methods and definitions is not new (Arora, 1996). In fact, over two decades have passed but still bullying research is in need of more clarity regarding defining and measuring bullying (Bradshaw,
2. Defining bullying

To begin with, bullying is a difficult behavior to define (Gladden, Vivolo-Kantor, Hamburger, & Lumpkin, 2014; Thornberg, 2015). The modal definition of bullying was provided by Dan Olweus in 1993 (pp.8–9): “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.” This definition highlights three main features: intentional aggression, repetition, and an imbalance of power. Although cited thousands of times, this definition is not without problems. On the applied level, it is not clear that participants use the core elements of this definition when answering questions about bullying (e.g., they omit power, harm, repetition, and/or intentionality; Bazelon, 2013; Green et al., 2013; Helleström, Persson, & Hagquist, 2015; Oldenburg, Bosman, & Veenstra, 2016; Vaillancourt et al., 2008). A lack of agreement with participants represents a challenge to the validity of any definition (DeLara, 2012). It also appears that researchers do not always assess all three aspects of this definition, leading to data that relates more to general aggression (any harmful behavior) than to bullying (Bradshaw, 2015; Finkelhor, Turner, & Hamby, 2012; Hawley et al., 2011). There are also theoretical concerns with this definition that were born out of practical observations without any serious theoretical examination of potential limits on its validity and/or compatibility with existing bodies of research (Greif & Furlong, 2006; Olweus, 1993; see below for a discussion of some of these theoretical issues). These problems have led researchers to explore numerous alternative definitions of bullying (Arora, 1996; Corcoran, Guckin, & Prentice, 2015; Cowie, Naylor, Rivers, Smith, & Pereira, 2002; Gladden et al., 2014; Monks et al., 2009; Sercombe & Donnelly, 2013).

Recently, Volk, Dane, and Marini (2014) defined bullying as “aggressive, goal-directed behavior that harms another individual within the context of a power imbalance.” This new definition addresses three theoretical issues. First, it removes the generally difficult to measure criterion of intentionality (Bauman, Underwood, & Card, 2013; Sercombe & Donnelly, 2013) and replaces it with concrete goals that are both easier to measure and predict important measurable outcomes (Ellis, Volk, Gonzalez, & Embry, 2016; Marini & Volk, 2017). Evolutionary theory suggests that individuals bully for resources, reproductive opportunities, or a social reputation that facilitates obtaining the first two goals (Volk et al., 2012, 2014). The latter goal is closely related to, but potentially distinct from, the dominance and popularity goals that bullying is often associated with. A focus on goals also helps tease apart the proximate goals of bullying (e.g., getting peers to laugh, looking cool, showing off, cutting-off reflected failure, basking in reflected glory) from the ultimate goals of bullying (e.g., getting the best field to play in at recess or having more sexual partners). It also addresses the apparent paradox of why a bully would pick on a much weaker individual in order to gain social status. The victim represents the means through which to send a signal about the perpetrators’ willingness to use aggression to obtain or maintain high social status (Volk et al., 2014). This is supported by evidence showing that while the number of victims per bully remains generally stable over time, bullies appear to continually select new victims with which they can display their dominance (Van der Ploeg, Steglich, & Veenstra, 2017).

A goal-oriented focus highlights that whereas bullying is often proactive (Sijtsma, Veenstra, Lindenberg, & Salmivalli, 2009), proactive/reactive labels may not necessarily be the most suitable for describing bullying. If one’s goal is to obtain a reputation for being violent in the face of social threats, then it may well serve that goal to proactively bully weaker individuals to maintain the credibility of one’s threatening reputation. However, that same reputation can also be served by planned reactive aggression toward perceived or potential challenges to one’s reputation (Frey, Pearson, & Cohen, 2015). The critical distinction is that bullies strategically use aggression in the context of a power imbalance to obtain their desired goals (Reijntjes et al., 2013; Veenstra, Lindenberg, Munnikisma, & Dijkstra, 2010). Overall then, a focus on goals instead of general intentionality helps researchers to avoid inconsistencies surrounding proactive/reactive aggression, to identify important outcomes of bullying, and to circumvent messy issues surrounding the measurement of intentionality.

Second, Volk et al.’s definition (2014) recognizes that harm is a perception of the victim that relates to both the frequency and intensity of the bullying behavior. A single heinous act of aggression may suffice to permanently alter the psychology and behavior of a victim (e.g., Parke, 2012), whereas repeated acts of very low intensity (e.g., several hundred brief prank calls) may equally cause significant harm. Thus, a victim’s experience of harm is the product of frequency by intensity by individual resiliency (Volk et al., 2014). In support of this, recent findings indicate that repetition, severity, and the presence of a power imbalance amplify the perceived harm of an action (Van der Ploeg, Steglich, Salmivalli, & Veenstra, 2015; Van Noorden, Bukowski, Haselager, Lansu, & Gillessen, 2016; Ybarra, Espelage, & Mitchell, 2014). This also means that the same act performed at the same frequency may cause different harm to different individuals based on numerous different individual and environmental ecological factors (e.g., Barboza et al., 2009; Flaspohler, Elstrond, Vanderzee, Sink, & Birchmeier, 2009; Hong & Espelage, 2012; van der Ploeg et al., 2015; Vitoroulis & Vaillancourt, 2014). Olweus (1993) alluded to single-incident harmful episodes of bullying, but chose to use frequency as a convenient filter for trivial harm. The aforementioned theory and data, however, suggest that frequency is not a necessary criterion for bullying. What’s more, we do not recommend a single definition of what is specifically harmful. Rather, we argue that victim harm is a complex ecological outcome that traditional definitions of bullying often fail to address by virtue of their inability to identify the goals of bullying behaviors. We suggest that a lot of research underestimates or misses important consequences outside of mental and physical health such as access to resources or mating partners.

Finally, Volk et al.’s (2014) definition situates bullying within the broader literature on aggression. Bullying is a specific form of aggression that involves a power imbalance. A power imbalance is perhaps the feature that most clearly differentiates bullying from other forms of aggression. It is well established that a physical power imbalance greatly alters the odds of success of physical competitions (Khosla, 1968). Research shows that a power imbalance leads to more severe outcomes for victims of bullying (Hunter, Boyle, & Warden, 2007; Ybarra et al., 2014). More powerful bullies seem to be less susceptible