

# Refining victims' self-reports on bullying: Assessing frequency, intensity, power imbalance, and goal-directedness

Tessa M. L. Kaufman  | Gijs Huitsing  | René Veenstra 

Department of Sociology and Interuniversity Center for Social Science Theory and Methodology (ICS), University of Groningen, Groningen, the Netherlands

## Correspondence

Tessa M. L. Kaufman, Department of Sociology and Interuniversity Center for Social Science Theory and Methodology (ICS), University of Groningen, Groningen, the Netherlands.

Email: t.m.l.kaufman@rug.nl

## Funding information

Nederlandse Organisatie voor Wetenschappelijk Onderzoek, Grant/Award Number: 451-17-013 and 453-14-016

## Abstract

Bullying can be differentiated from other types of peer aggression by four key characteristics: frequency, intensity, power imbalance, and goal-directedness. Existing instruments, however, usually assess the presence of these characteristics implicitly. Can self-report instruments be refined using additional questions that assess each characteristic? We examined (a) what proportion of children classified as victims by the commonly used Revised Olweus' bully/victim questionnaire (BVQ) also experienced the characteristics of bullying, and (b) the extent to which the presence of the characteristics was associated with emotional (affect, school, and classroom well-being), relational (friendship, defending), and social status (popularity, rejection) adjustment correlates among victims. Using data from 1,738 students ( $M_{\text{age}} = 10.6$ ; grades 5–8), including 138 victims according to the BVQ, the results showed that 43.1% of the children who were classified as victims by BVQ experienced all the four characteristics of bullying. Frequency ratings of victimization did not capture experiences that involved a power imbalance. Victims who reported all four key characteristics had greater emotional, relational, and social status problems than victims who did not report all characteristics. Thus, researchers who focus on victimization for diagnostic and

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. Social Development published by John Wiley & Sons Ltd

prevention purposes can enrich self-report measurements of bullying victimization by adding questions that assess the characteristics explicitly.

#### KEYWORDS

assessment, bullying, concurrent validity, Olweus, victimization

## 1 | INTRODUCTION

Bullying is widely recognized as a unique peer phenomenon (Volk, Veenstra, & Espelage, 2017). By definition, “a person is being bullied when he or she is exposed, repeatedly and over time, to negative actions on the part of one or more other persons” (Olweus, 1993). A recent concretization of this definition states that bullying can be differentiated from other types of peer aggression by four key characteristics: *frequency*, *intensity*, *power imbalance*, and *goal-directedness* (Volk, Dane, & Marini, 2014).

This distinction between victims of bullying and victims of the broader class of peer aggression is crucial in evaluating the anti-bullying interventions. Tackling victimization through bullying requires different interventions than reducing general victimization through aggression (Espelage, Low, Polanin, & Brown, 2013; Taub, 2002; Van Schoiack-Edstrom, Frey, & Beland, 2002). Bullying is embedded in the peer group and thus also requires group-focused interventions whereas general aggression may also be resolved by targeting individual skills. Thus, evaluating anti-bullying interventions calls for measures that can differentiate between victimization through bullying and through general aggression. Last, the distinction is also important to limit variability in prevalence estimates of bullying across studies (Cook, Williams, Guerra, & Kim, 2010).

However, it is unclear whether currently used measures of peer victimization are able to differentiate between victimization through bullying versus general aggression (Bauman, 2016; Furlong, Sharkey, Felix, Tanigawa, & Green, 2010; Vivolo-Kantor, Martell, Holland, & Westby, 2014). The Revised Olweus' bully/victim questionnaire (BVQ; 1996), considered the most widely used instrument (Lee & Cornell, 2009), uses a definition-first approach, meaning that it first provides children with a definition of bullying and subsequently asks about the frequency of aggressive experiences. However, children may not retain this complex, multicomponent definition in working memory and apply it when answering the questionnaire. They may fall back on their own prior assumptions about the term bullying and, therefore, report their experiences of general peer aggression as bullying (Furlong et al., 2010; Jia & Mikami, 2018). Although the BVQ explicitly measures “frequency”, this does not necessarily guarantee that it captures the other key characteristics of the definition, because higher frequency experiences may not be intense, nor happen in the context of a power imbalance or on purpose (Felix, Sharkey, Green, Furlong, & Tanigawa, 2011). Indeed, previous research has shown that BVQ responses were more likely to detect the repeated experiences than power imbalance (Green, Felix, Sharkey, Furlong, & Kras, 2013). Thus, a general concern is that the sole use of frequency ratings in a definition-first approach cannot fully discriminate between victims of bullying and victims of other types of peer aggression.

A possible way to improve the differentiation between victimization groups is to add questions to instruments such as the BVQ that address each key characteristic of the bullying definition explicitly (Bauman, 2016; Furlong et al., 2010; Jia & Mikami, 2018; Volk et al., 2014). However, it is empirically unclear whether the definition characteristics are valid indicators of victimization through bullying: whether experiencing all key characteristics indeed relates to the correlates that conceptually differentiate between victimization through bullying and through general aggression (Jia & Mikami, 2018).

In addressing these concerns, our aim was twofold. First, we aimed to examine the extent to which the method of using a definition-first approach and endorsing frequency ratings only (i.e., the BVQ) captures the experiences

that match the definition of bullying. Second, we examined whether adding to the BVQ questions that explicitly assessed the characteristics helped to differentiate victimization through bullying from victimization through general aggression. Did victims who were victimized in line with the bullying definition (compared with victims who were not) show the emotional, relational, and social status adjustment correlates that conceptually relate more strongly to victimization through bullying than to victimization through general aggression?

## 1.1 | Key characteristics of victimization through bullying

In the past, bullying was perceived as an impulsive, uncontrolled outburst of aggression toward victims (Olweus, 1978). However, nowadays most people agree that bullying is part of a complex group phenomenon (Salmivalli, 2010) that is characterized by four key characteristics. Bullying (a) *repeatedly* (b) *harms* victims in the context of a (c) *power imbalance* and predominantly involves (d) strategic, *goal-directed* behavior (Olweus, 1993; Reijntjes et al., 2013; Volk et al., 2014).

These key characteristics have been conceptually and empirically related to correlates of victimization through bullying. First, the *repetitive* character of bullying refers to frequent experiences of victimization instead of a one-time occurrence (Olweus, 1993; Volk et al., 2014). The frequency of monthly ("two or three times a month") victimization seems a valid lower cutoff point for classifying children as victims, because this distinguishes victims from non-victims in levels of higher psychosocial maladjustment (Solberg & Olweus, 2003). Frequent victimization has been related to lower social support and well-being (e.g., Fullchange & Furlong, 2016; Solberg & Olweus, 2003; Ybarra, Espelage, & Mitchell, 2014).

Second, bullying was proposed to be characterized by *intensity*, meaning that victims experience bullying as harmful and thus intense, which differentiates it from playful teasing or fighting (Olweus, 1993; Volk et al., 2014). Perceived intensity may be a powerful predictor of worse adjustment correlates (Volk et al., 2014).

Third, *power imbalance* means that bullies have a more powerful position in the group and choose victims who have less physical or social strength (Nelson, Kendall, Burns, Schonert-Reichl, & Kane, 2019; Olweus, 1993) in order to lower the cost of their behaviors. For example, bullies minimize the loss of affection by choosing victims who are not likely to be defended by significant others (Veenstra, Lindenberg, Munniksmas, & Dijkstra, 2010), which reflects the group process of bullying. Imbalance of power differentiates bullying from general aggression, in which aggressors can be equal in power, and reflects bullying's social nature (Volk et al., 2014). Being victimized by someone with greater power may elicit depressive symptoms, because victims feel powerless to change the situation (Hunter, Boyle, & Warden, 2007), and can interfere with victims' relationships with friends and family and their schoolwork (Ybarra et al., 2014).

Finally, bullies' minimization of the loss of affection is also reflected in *goal-directedness*. This characteristic refers to the strategic nature of bullying, as opposed to accidental behavior. Whereas the intention of bullying was previously described as wanting to harm another child (Olweus, 1993), recent insights based on evolutionary and sociological theory emphasize that bullying may be aimed at obtaining or maintaining social dominance (Olthof, Goossens, Vermande, Aleva, & Van der Meulen, 2011; Van Der Ploeg, Steglich, & Veenstra, 2020; Volk et al., 2014). This can involve retaliation for previous actions by victims that were aimed at the bullies or their friends (Frey, Pearson, & Cohen, 2015), but does not necessarily need to be reactive. The group process that characterizes bullying is also reflected by this goal to achieve social dominance (Salmivalli, 2010).

Can children's experiences with all key characteristics help to discriminate between victimization through bullying and victimization through general aggression? If the BVQ appears to not capture bullying as it is defined, this is only problematic when experiencing all key characteristics would indeed improve the concurrent validity of the measure. A way to investigate this question is by examining whether experiencing all key characteristics, versus not all characteristics, contributes to correlates that conceptually differ between victimization through bullying and victimization through general aggression (Jia & Mikami, 2018).

Conceptually, victims of bullying are particularly distinct from victims of general aggression in their greater emotional adjustment problems and problems in the peer group, such as compromised social relationships and status (Hunter et al., 2007; Olweus, 1996; Solberg & Olweus, 2003). First, victims of bullying will theoretically show greater emotional maladjustment because bullying concerns structural social exclusion, which the victim cannot easily escape from, given the power differential. This type of exclusion has a detrimental effect on mental health because it interferes with the fundamental human need to belong to the group (Baumeister & Leary, 1995). Second, the exclusion that characterizes bullying is extra painful because it is person-oriented and thus not directed toward a random target (Olweus, 1993; Volk et al., 2014). Third, bullying is a social phenomenon that is embedded in the peer group, and thus relates to victims' social adjustment. Bullies target victims who already have less supportive peer relationships and lower social status in the peer group. Others do no longer want to affiliate with these victims because this could lower their own status and increase the risk of being the next target (Salmivalli, 2010), leaving victims with fewer peers who befriend or defend them. Although bullying also impairs victims' functioning in other domains such as academic functioning, physical health, and parent-child relationships, the majority of those factors may be associated with the emotional and social adjustment problems.

Some evidence already provides support for the suggestion that experiencing multiple key characteristics of the bullying definition relates to greater correlates of victimization through bullying. First, victims of repeated aggression who were less powerful than the bully had poorer mental health than those who only experienced repeated aggression (Ybarra et al., 2014). Second, compared with victims who experienced repeated aggression, victims who also were both less powerful than the bully and experienced that the bully victimized them on purpose had higher levels of depressive symptoms (Hunter et al., 2007; Malecki et al., 2015), anxiety, and lower self-esteem (Malecki et al., 2015).

An important next step in this research on the concurrent validity of the key characteristics of the bullying definition was to also (a) include effects on *school-related* affect, because bullying is strongly embedded in the school context, and (b) focus on relational and social status adjustment correlates, which particularly characterize this social phenomenon. Lastly (c), previous research has not examined all four key characteristics together and their relative importance; this is central to differentiating victimization through bullying from victimization through general aggression.

## 1.2 | Current study

In this study, we examined whether extending self-reports of victimization with explicit assessment of each key characteristic of the bullying definition can improve the differentiation between victims of bullying and victims of other types of peer aggression. The first research question (RQ1) concerned the extent to which the definition-first approach with frequency ratings, which is employed by the most commonly used measure (the Revised BVQ; 1996), captures experiences with all key characteristics of the bullying definition: *frequency*, *intensity*, *power imbalance*, and *goal-directedness* (Olweus, 1993; Volk et al., 2014). To this end, we added questions to the BVQ that explicitly assessed the key characteristics, and examined (RQ1a) how many of the children who were classified as victims by the BVQ (victimized at least "monthly"; Solberg & Olweus, 2003) also experienced all key characteristics, and how many did not. Second, we examined which key characteristics may particularly be overlooked when using frequency ratings: (RQ1b) how strongly is frequency associated with experiences of intensity, power imbalance, and goal-directedness? We were interested in both linear and quadratic effects, because a very powerful or strategic bully may only need to victimize once to reach their goals.

The second research question (RQ2) investigated whether extending the BVQ with a more narrowband approach, thus posing questions that explicitly address experiences with each key characteristic, improves discrimination between victims of bullying and victims of general aggression. We examined (RQ2a) to what extent

victims who experienced the key characteristics, versus those who did not, showed greater emotional (affect, school/classroom well-being), relational (friendships, defenders), and social status (popularity, rejection) adjustment correlates that are conceptually stronger among victims of bullying than among victims of more general peer aggression. Last (RQ2b), we examined whether the BVQ was still relevant or could be replaced by the new specific questions. We analyzed, within the sample of victims who experienced all key characteristics, differences in adjustment correlates between victims who would also have been classified as victims by the BVQ (because they reported systematic, thus monthly, victimization in the BVQ measure) and those who would not have been classified as victims by the BVQ (because they initially reported occasional victimization in the BVQ measure).

## 2 | METHODS

### 2.1 | Procedure

First, we developed questions that explicitly assessed the key characteristics of the bullying definition. The questions were based on theory and previous research and at the same time were practical to use (e.g., Malecki et al., 2015; Volk et al., 2014; Ybarra et al., 2014). Second, we obtained IRB approval by the Ethics Committee from the Department of Sociology for a pilot study in four schools on the practical use and formulation of the questions, and the duration of filling in the questionnaire. After the pilot, we adjusted the questions and administered them together with measures of adjustment correlates to students in a larger sample of schools that took part in the eleventh wave of the ongoing study on the KiVa anti-bullying program (Kärnä et al., 2011) in the Netherlands (Huitsing et al., under review). Schools were selected based on the criteria that they had distributed informed consent forms to parents and that they administered the bi-annual questionnaire at about the same time.

Information about the study and consent forms were sent to parents prior to assessments. An active consent procedure was used in which parents were asked to indicate whether they wished to allow their child to participate in the research. Students were informed at school about the research and gave oral assent. Students did not participate when they had no permission (parents refused participation or did not return the consent form), when they did not want to participate themselves, or when they were unable to complete the questionnaire. Internet-based questionnaires were completed individually in the classroom during regular school hours with primary teachers present to answer questions and assist students when necessary. The order of questions and instruments used was randomized to avoid systematic effects of question order.

### 2.2 | Participants

Of the 2,257 students in the sample of schools in the main study, we used the data of 1,738 (77.0%) students who had permission to participate in the study. Children with and without permission did not differ in gender (as reported by the school) or peer-reported data (nominations received for popularity, rejection, defending, and bullying),  $ps > .05$ . The students in the final sample attended 26 schools (196 classrooms; 50.6% boys), in Dutch grades 5 to 8 (US-level grades 3 to 6;  $M_{\text{age}} = 10.6$ ,  $SD = 1.2$ ). Of these students, 272 (15.7% of the total sample) had been victimized at least once in recent months (*victims sample*), and thus received questions about the key characteristics. The BVQ-classified systematic victims concerned 138 children (11.1% of the total sample). Last, 1,238 children (71.2% of the total sample) did not report victimization (*non-victims*). See Supporting Information Figure A1 for an overview of the sub-samples.

## 2.3 | Measures

### 2.3.1 | Victimization and key characteristics

We measured victimization using the traditional Olweus' (1996) BVQ. The Olweus' BVQ provides children with a definition of bullying (repeatedly harassing another child and the victim has problems defending themselves, with examples that explain that it is intense and happens on purpose); this was presented to children individually in a video that was integrated in the questionnaire. Children responded to one global item ("How often have you been bullied during the past couple of months?") followed by questions about specific forms of bullying. These items distinguished five forms of bullying (7 items in total): physical, verbal (two items), relational (two items), material (taking or breaking others' property), and cyber-victimization (receiving nasty or insulting messages, calls, or pictures). Children answered on a five-point scale how often they experienced each form: 0 = *not at all*, 1 = *only once or twice*, 2 = *two or three times a month*, 3 = *about once a week*, 4 = *several times per week*.

To assess victims' experiences with the key characteristics of the bullying definition (referring to whether they were victims of bullying or general aggression), we extended the BVQ with new questions that assessed explicitly whether children experienced each key characteristic. Children were only presented with these questions when they reported being victimized at least "once or twice" on any of the BVQ items (see Supporting Information Table A1 for the full questionnaire). Children were first asked to select the names of up to three children that bullied them most often, starting with the person who bullied them the most followed by those who also bullied them. If children were bullied by a group, they could name members of the group separately in this way.

For each bully, we then presented them with seven questions about their victimization experiences (up to 21 items). One item measured the *frequency* of victimization by a specific bully (Olweus, 1996), and one item assessed intensity as the extent to which children experienced the bullying as *harmful* (Volk et al., 2014). *Power imbalance* was assessed using two items that measured whether the victims perceived the bully as being stronger (one item) or more popular (one item) than themselves. The items used were those most often used to assess power imbalance (Felix et al., 2011; Green et al., 2013; Hunter et al., 2007; Malecki et al., 2015; Ybarra et al., 2014) and which have been shown to measure the most relevant types of power imbalance in victimization experiences (Nelson et al., 2019). *Goal-directedness* has previously been conceptualized as the intention to be mean (Felix et al., 2011; Hunter et al., 2007; Malecki et al., 2015), but we adhered to the theoretical conceptualization of bullying as a strategic behavior aimed at obtaining or maintaining social dominance (Olthof et al., 2011). Therefore, we assessed it using three items that measured to what extent children were sure that the bully bullied them on purpose (one item), and (regardless of their answer to the previous item) their perception of this purpose: to gain social reputation ("to be cool") and to take revenge (one item each). Children responded to each question using a 5-point scale, representing "not experienced" to "experienced strongly"; from this we created a scale that represented the maximum score across the three items, and not the average, because bullies may have only one goal.

### 2.3.2 | Emotional adjustment correlates

We assessed *positive and negative affect* using the PANAS-C (Watson, Clark, & Tellegen, 1988), which consisted of ten adjectives that described emotions. Children indicated how often they had experienced each feeling during the previous 2 weeks. We computed the mean of the five items for both scales ( $\alpha = .82$  for positive affect and  $\alpha = .75$  for negative affect). We assessed *well-being at school* using the mean of five items concerning perceptions of the classroom and school (Kärnä et al., 2011). Students responded to items such as "I feel accepted as I am at school" (1 = *never*, 4 = *always*),  $\alpha = .82$ . We measured *classroom comfort* using four items from the Classroom Peer Context Questionnaire that focused on comfort (Boor-Klip, Segers, Hendrickx, & Cillessen, 2016), for example, "In this class, I belong to the group" (1 = *never*, 4 = *always*),  $\alpha = .84$ . Latent Factor Analysis showed that the four indicators

represented one overarching construct (CFI = .98, TLI = .99): they all showed significant and acceptable to high Geomin Rotated factor loadings (positive affect  $\lambda = 0.68$ , negative affect  $\lambda = -0.45$ , school wellbeing  $\lambda = 0.90$  and classroom comfort  $\lambda = 0.87$ ). To minimize the number of analyses we used this latent factor in the analyses of RQ2.

### 2.3.3 | Relational and social status adjustment correlates

To assess *defending*, children were first asked to read a piece explaining that some children help children who are bullied by supporting, comforting, or otherwise helping them. We then asked them to nominate the classmates who defended them: "Which classmates defend you when you are victimized?" In addition, they nominated the classmates they perceived as their best friends ("Which classmates are your best friends?" *friendship*), as most popular ("Who are the most popular students in your class?" *popularity*), and who they disliked ("Which classmates do you dislike?" *rejection*). For each student, nominations received (for defending: outgoing) for each variable were summed and divided by the number of participating classmates, resulting in proportion scores (0–1).

All measures were equally reliable across Dutch grade groups 4 ( $n = 2$ ), 5, and 6 versus 7 and 8 (difference in  $\alpha < .03$ ; US-level grades 2–4 vs. 5–6).

## 2.4 | Analyses

We conducted the analyses in *Mplus 7* and used Maximum Likelihood Estimation (MLR), which is robust to violations of non-normality. We computed intra-class correlations (ICC) of the manifest variables of victimization and adjustment correlates at the classroom level. The level of explained variance at the classroom level varied between ICC = .03 (negative affect) and ICC = .32 (friendships). We thus used a multilevel structure with the cluster command to take into account the dependent structure of the data. We controlled for children's gender and age.

### 2.4.1 | RQ1: Analysis steps 1 and 2

First (RQ1a), we examined to what extent the BVQ classifies children as victims when they have experiences that are in line with the definition of bullying. Within the sample of children who would be classified by the BVQ as victims, we examined how many of these victims did, and how many did not, (a) experience each key characteristic, and (b) experienced all characteristics together, thus in line with the definition. The cutoff criteria to determine that a characteristic was experienced (vs. not experienced) were that, at least with one of the three bullies, victimization had happened (a) at least monthly and the victim perceived: (b) it as at least a little bit intense; (c) the bully as at least a little bit stronger or more popular; (d) it as on purpose.

Second (RQ1b), we examined which key characteristics are especially missed using measures that only assess frequency explicitly, such as the BVQ. Using regression analyses we examined to what extent frequency was associated (a) with experiencing the other key characteristics (intensity, power imbalance, and goal-directedness) all together, and (b) with the extent to which each individual key characteristic was experienced, using the original, continuous, key characteristic measures (0 = *not experienced* to 4 = *experienced strongly*).

### 2.4.2 | RQ2: Analysis steps 3 and 4

We examined the concurrent validity of the addition of the key characteristics to the BVQ. First (RQ2a), did victims who experienced the key characteristics (victims of bullying) and those who did not (victims of general aggression)

indeed differ on correlates that are conceptually greater among victims of bullying than victims of general aggression? We focused on emotional, relational (friendship, defending), and social status (popularity, rejection) adjustment correlates. Next, we examined the contribution of each individual characteristic by regressing the effects of each key characteristic on the adjustment correlates. We allowed intercorrelations across all outcomes and key characteristics.

Last (RQ2b), we examined whether the BVQ was still relevant or could be replaced by the new specific questions. We analyzed, within the sample of victims who experienced all key characteristics, differences in adjustment correlates between victims who would also have been classified as victims by the BVQ (because they reported systematic victimization in the BVQ measure) and those who would not have been classified as victims by the BVQ (because they initially reported occasional victimization in the BVQ measure).

To reduce the risk for false discovery rates (FDR) (regression analyses with five outcomes), we used an FDR controlling procedure when determining statistical significance (Benjamini & Hochberg, 1995). Thus,  $p$  values across the five outcomes were first ordered from smallest to largest, ranking them  $i = 1$  to  $i = 5$ . A threshold of significance (critical value) was established according to the formula: critical value  $(p_i) \frac{i}{m} Q$  ( $m$  = number of tests,  $Q$  = percentage of false discoveries  $5\% = .05$ ). This procedure resulted in the following critical values:  $p_{(1)} \leq .01$ ,  $p_{(2)} \leq .02$ ,  $p_{(3)} \leq .03$ ,  $p_{(4)} \leq .04$ ,  $p_{(5)} \leq .05$ . Each ranked  $p$  value was then compared with its corresponding critical value, starting with  $i = 5$ . The critical value that was used was that of the highest ranking  $p$  value that was below its corresponding critical value. Thus, the lowest  $p$  value was compared to  $p_{(1)} \leq .01$ , the second lowest to  $p_{(2)} \leq .02$ , and so on. We applied this method to all analyses except the RQ2b analyses, because the small sample size in those analyses ( $N = 82$ ) would increase the risk for Type 2 error; instead, we used the  $p < .05$  threshold.

### 2.4.3 | Supplementary analyses using different operationalizations

We conducted two types of supplementary analyses. First, in our main analyses, we used a dichotomized instead of continuous approach to compute the victimization groups (i.e., victims of bullying vs. victims of general aggression), to ensure that the victims of bullying experienced all, and not only some, key characteristics. The definition of bullying proposes that all four characteristics need to be present in order to classify experiences as bullying; if two people fight regularly to achieve social dominance but they are equally strong, it is not bullying (Olweus, 1993). This required a dichotomous approach because if we had used the mean across all characteristics (the continuous approach), victims could also receive a high score on "being bullied" when one characteristic was highly present (e.g., frequency), but another characteristic was absent (e.g., there was no power imbalance). However, we conducted a sensitivity analysis to determine whether the results (of RQ2a) were consistent when the effects of a continuous measure of victimization through bullying on adjustment correlates were analyzed. In this analysis, victimization through bullying represented the average score across all key characteristics.

Last, in our main analyses, the score on each individual key characteristic represented the highest (maximum) score across the three bullies, because key characteristics only need to be experienced with one of the bullies in order to have an impact. However, computing each key characteristic as the average across three bullies also seems informative to indicate the severity of victims' problems. For example, if victims experience a power imbalance with three bullies (high average for power imbalance), this may signal that they have extreme difficulty in defending themselves. Therefore, we replicated the analyses in RQ2a by computing the key characteristics as the average across the bullies.

## 3 | RESULTS

### 3.1 | Step 1: Experiences of key characteristics of the definition

We first examined the extent to which victims in the BVQ reported the key characteristics of bullying. In total, there were 138 systematic victims who reported the name of at least one bully: 135 (97.8% of the victims) selected



the name of one bully, 91 children (65.9%) selected a second bully, and 55 children (39.9%) selected a third bully. Table 1 shows how many self-reported victims reported the key characteristics: it shows whether frequency (row 1), intensity (row 2), power imbalance (rows 3–5), goal-directedness (rows 6–8), and all characteristics together (row 9) were experienced. The columns show how often the characteristics were experienced on average across three bullies (column 1), and, most centrally, among at least one of the three bullies (column 2).

Children mentioned all characteristics on average in 29.8% of the victimization experiences (column 1). Victimization was *repetitive* in, on average, 59.6% of the cases and was experienced as *intense* in 79.4% of the cases. On average, 59.8% of the victimization experiences were characterized by a *power imbalance*; the majority were characterized by a difference in *strength* (68.7) or *popularity* (82.6%). Last, on average 70.7% of the children experienced victimization as being *goal-directed*, and within this group, a majority perceived the goal to be to increase one's status, and a minority perceived the goal to be to take revenge. Thus, overall, each key characteristic was present in most victimization cases, but not in all.

Victims were somewhat more likely to experience the key characteristics in *any bully* (column 2), compared with the averages across all bullies in column 1. About two-fifths (43.1%) were victimized by any bully in line with the definition, and three-fifths were not. Further, the majority of the children experienced each individual characteristic in at least one bully, ranging from 71.0% in which a power imbalance was present in at least one bully, to 79.7% in which victimization was goal-directed for at least one bully, and 87.6% in which victimization was experienced as intense.

### 3.2 | Step 2: Regressions of intensity, power imbalance, and goal-directedness on frequency

Children who experienced all key aspects in any bully were victimized more frequently than those who were victimized but did not experience at least one key aspect,  $\beta = .25$ ,  $SE = 0.05$ ,  $p < .001$ ,  $R^2 = 0.07$ . Further, intensity ( $\beta = .39$ ,  $SE = 0.06$ ,  $p < .001$ ) and goal-directedness ( $\beta = .19$ ,  $SE = 0.05$ ,  $p = .001$ ) but not power imbalance ( $\beta = .01$ ,  $SE = 0.05$ ,  $p = .846$ ) were linearly associated with higher frequency,  $R^2 = 0.23$ . Age and gender were unrelated to frequency, and quadratic effects of the predictors on frequency were not significant ( $ps > .05$ ). Thus, high frequency ratings were also associated with intensity and, to a lesser extent, goal-directed experiences, but frequent victimization was not necessarily characterized by a power imbalance.

**TABLE 1** Prevalence of key characteristics of victimization across three reported bullies (in systematic victims sample:  $N = 138$ )

Row	Variable	Average prevalence (across the three bullies) <sup>a</sup> (%)	Prevalence in any of the three bullies <sup>b</sup> % (n)
1	Frequency	59.6%	67.4% (93)
2	Intensity	79.4%	87.6% (120)
3	Power imbalance	59.8%	71.0% (98)
4	Stronger	41.1%	55.1% (76)
5	More popular	49.4%	55.8% (77)
6	Goal-directed	70.7%	79.7% (110)
7	For status	42.5%	53.6% (74)
8	For revenge	13.2%	18.1% (25)
9	All characteristics	29.8%	43.1% (59) <sup>b</sup>

Note: Bold indicates very important value in the article.

<sup>a</sup>Refers to how often children experienced the key characteristic, on average across the three bullies.

<sup>b</sup>Refers to experiencing the key characteristic in any of the three bullies (thus, in at least one bully).

**TABLE 2** Results of regression analyses of victimization characteristics in any of the three bullies on adjustment correlates ( $N = 273$ )

	Emotional			Relational			Defending			Popularity			Rejection					
	Emotional adjustment (affect, classroom and school well-being)						Friendship			Defending			Popularity			Rejection		
	$\beta$ (SE)	$p$		$\beta$ (SE)	$p$		$\beta$ (SE)	$p$		$\beta$ (SE)	$p$		$\beta$ (SE)	$p$		$\beta$ (SE)	$p$	
Intensity	.03 (0.06)	.599		-.01 (0.06)	.852		-.03 (0.06)	.641		-.03 (0.07)	.634		.05 (0.05)	.356				
Frequency	<b>-.29 (0.07)</b>	<.001		-.14 (0.07)	.062		-.05 (0.05)	.319		-.09 (0.06)	.121		<b>.16 (0.06)</b>	.004				
Power imbalance	<b>-.23 (0.05)</b>	<.001		<b>-.15 (0.06)</b>	.014		-.11 (0.08)	.167		<b>-.36 (0.07)</b>	<.001		.04 (0.06)	.529				
Goal-directed	<b>-.17 (0.05)</b>	.001		.08 (0.06)	.168		.02 (0.07)	.721		.02 (0.06)	.712		.07 (0.06)	.290				
Gender (ref = girl)	-.04 (0.06)	.562		-.13 (0.04)	.036		.11 (0.06)	.051		.03 (0.05)	.568		.14 (0.07)	.052				
Age	-.10 (0.05)	.049		<b>-.26 (0.06)</b>	<.001		-.04 (0.08)	.629		<b>.18 (0.06)</b>	.004		.08 (0.06)	.198				
$R^2$	.20			.13			.03			.18								

Note: Classroom was used as clustering variable. Standardized effects are shown. Bold statistics are the significant effects. B-H = Benjamini-Hochberg procedure; B-H critical values:  $P_{(1)} \leq .01$ ,  $P_{(2)} \leq .02$ ,  $P_{(3)} \leq .03$ ,  $P_{(4)} \leq .04$ ,  $P_{(5)} \leq .05$ .

### 3.3 | Steps 3 and 4: Regressions of the key characteristics on adjustment correlates

Children who were victimized but did not experience all key characteristics (*victims of general aggression*;  $N = 191$ ) had worse emotional ( $\beta = -.38$ ,  $SE = 0.03$ ,  $p < .001$ ,  $R^2 = 0.15$ ), relational (friendships:  $\beta = -.08$ ,  $SE = 0.02$ ,  $p = .001$ ,  $R^2 = 0.05$ ), and status (only rejection:  $\beta = .18$ ,  $SE = 0.03$ ,  $p < .001$ ,  $R^2 = 0.05$ ; popularity:  $\beta = -.02$ ,  $SE = 0.02$ ,  $p = .478$ ,  $R^2 = 0.03$ ) adjustment than non-victims. Second and most centrally, victims who experienced all key characteristics (*victims of bullying*;  $N = 82$ ) had greater adjustment problems than general victims on all correlates. Thus, of all groups, the victims of bullying reported the most emotional maladjustment ( $\beta = -.31$ ,  $SE = 0.06$ ,  $p < .001$ ,  $R^2 = 0.11$ ), which was a latent factor based on lower positive and higher negative affect and lower school well-being and classroom comfort. Victims of bullying had also the fewest friendships ( $\beta = -.14$ ,  $SE = 0.06$ ,  $p = .019$ ,  $R^2 = 0.10$ ), the fewest defenders ( $\beta = -.11$ ,  $SE = 0.05$ ,  $p = .028$ ,  $R^2 = 0.03$ ), were the least popular ( $\beta = -.22$ ,  $SE = 0.04$ ,  $p < .001$ ,  $R^2 = 0.08$ ), and the most rejected ( $\beta = .11$ ,  $SE = 0.05$ ,  $p = .042$ ,  $R^2 = 0.04$ ). These differences between victims of bullying and general victims were found to be similar in an additional analysis in which victimization through bullying represented the continuous mean score across all key characteristics, with negligible differences in the sizes of the beta coefficients (differences  $< .02$ ; Supporting Information Table A2, row 8). Thus, experiences of all key characteristics were indicators of emotional, relational, and social status adjustment correlates that are conceptually related to victimization of bullying.

Further regression analyses showed that each separate key characteristic, except for intensity, was associated with adjustment correlates when the other key characteristics were controlled for (Table 2). Being *frequently* victimized and a *power imbalance* were related to emotional adjustment and having fewer friendships (only power imbalance). Moreover, *goal-directedness* of victimization additionally explained children's emotional adjustment over and above frequency and power imbalance. Only being *less powerful* than the bully, but not the other key characteristics, was associated with being less popular. Last, only greater *frequency*, but not the other key characteristics, was related to greater rejection by classmates. All results were similar in supplementary analyses in which each key characteristic represented the mean score (instead of the highest score) across all bullies, with small differences in the sizes of the beta coefficients (differences  $< .05$ ; Supporting Information Table A2, rows 1–4) and with the exception of the now marginal instead of significant effect of frequency on rejection. Overall, all key characteristics except for intensity were thus in some way related to victims' adjustment correlates, particularly to school- and classroom-related well-being.

Last, we examined whether the BVQ measure was still needed in addition to the new specific questions. Among the victims of bullying who experienced all key characteristics based on the new questions ( $N = 82$ ), most (72%) were systematically victimized according to their BVQ response ("BVQ victims") whereas 28% were not: they reported occasional victimization in the BVQ ("non-BVQ victims"). The BVQ victims did not differ from non-BVQ victims in their emotional adjustment ( $\beta = -.10$ ,  $SE = 0.12$ ,  $p = .415$ ,  $R^2 = 0.06$ ), but had partly worse relational adjustment (only fewer friendships:  $\beta = -.21$ ,  $SE = 0.10$ ,  $p = .044$ ,  $R^2 = 0.22$ ; for defending:  $\beta = -.02$ ,  $SE = 0.11$ ,  $p = .840$ ,  $R^2 = 0.01$ ) and partly lower social status (only rejection:  $\beta = .23$ ,  $SE = 0.10$ ,  $p = .023$ ,  $R^2 = 0.11$ ; popularity:  $\beta = .06$ ,  $SE = 0.10$ ,  $p = .581$ ,  $R^2 = 0.03$ ) than non-BVQ victims. Thus, the BVQ classification partly helped to identify victims of bullying who had worse adjustment problems that are conceptually related to bullying.

## 4 | DISCUSSION

This research addressed the question of to what extent the widely used BVQ instrument for the assessment of bullying can be extended to make a more stringent differentiation between victims of bullying (according to its definition) and other types of peer aggression. This distinction is considered essential to evaluate anti-bullying interventions and for prevalence assessments of this unique group phenomenon (Jia & Mikami, 2018). Our research showed that less than half of the children who reported being victimized according to the BVQ also experienced all

key characteristics of the bullying definition (*frequency, intensity, power imbalance, and goal-directedness*), despite being provided with a definition. More than half of the self-reported victims did not experience all characteristics of bullying, and, therefore, seemed to be victims of another type of peer aggression. Particularly power imbalance, and to a lesser extent goal-directedness, were experienced least often by children who were victimized according to the BVQ. Frequency ratings only do not seem to capture experiences that match the full definition of bullying.

Does adding explicit questions about these key characteristics help to differentiate victimization through bullying from victimization through general aggression (concurrent validity)? In support of earlier research focusing on one or some of the key characteristics of bullying (Green et al., 2013; Hunter et al., 2007; Malecki et al., 2015), children who mentioned the key characteristics had greater emotional, relational, and status adjustment problems that are conceptually related to victimization; not only as compared with non-victimized children but also as compared with children who experienced more general peer victimization. Overall, our findings imply that measures that use the definition-first approach, such as the BVQ, could gain more precision in identifying victims of bullying specifically according to their definition. This could be done by adding to the measure a relatively minor number of questions that explicitly assess experiences with each key characteristic.

Children who did not stick to the definition of bullying when answering the BVQ may have relied on their own definitions instead of adhering to the definition provided. Children's definitions often deviate from the formal conceptualizations of bullying (Vaillancourt et al., 2008), which can be problematic because their experiences require different interventions.

Zooming in on the key characteristics, many children who reported victimization did not experience a power imbalance (29.0%) or goal-directedness (20.3%). This is in line with previous research that focused exclusively on power imbalance, and not goal-directedness; suggesting that the BVQ was less sensitive to power imbalances (Green et al., 2013). Our results make sense because a higher frequency of victimization, the only characteristic explicitly assessed in the BVQ, was not or to a small extent related to power imbalance and goal-directedness. Previous research has shown that many children did not spontaneously mention power imbalance (74.0%) or goal-directedness ("intentionality"; 98.3%) as essential characteristics of bullying (Vaillancourt et al., 2008), which might explain why not all of them included it in their answers, despite being given a definition of bullying. Nevertheless, a power differential is acknowledged as a central characteristic of bullying because it embodies its social nature (Olweus, 1993; Volk et al., 2014). In our study, victims who reported a power imbalance also had poorer emotional, relational, and social status adjustment. Considering goal-directedness is relevant because this characteristic defines the strategic nature of bullying (Olthof et al., 2011; Veenstra et al., 2007; Volk et al., 2014) and was also associated with emotional maladjustment at school in our study.

Notably, the intensity associated with victimization did not predict any adjustment correlates. Perhaps the measure had too little variation: the vast majority of the children (87.6%) experienced intensity. Alternatively, the children's recollections of intenseness may not be valid, because they may have suppressed or positively reappraised the experiences to decrease their emotional impact (Gross, 1998). Indeed, victimization was associated with poorer adjustment regardless of whether children reported it as intense, and omitting the intensity item did not affect the results. That most children experienced victimization as intense is a sign that intense experiences are incorporated in children's definition of bullying.

Our findings highlight the relevance of the use of the definition to discriminate between victimization through bullying and through general aggression, by showing that experiences that were in line with the definition were not only associated with mental health (Hunter et al., 2007; Malecki et al., 2015), but also with social adjustment. Compared with victims who did not report all the key characteristics of bullying, bullied children experienced greater emotional problems in general and related to school, had fewer friends and defenders, and were less popular and more rejected. Power imbalance played an important role in the associations with social adjustment; this characteristic was most strongly related to fewer friendships and lower popularity. Frequency was the only correlate of rejection which could, although tentatively, support the expectation that bullies aim to send a strong signal of dominance, but do so in a way that minimizes the costs to other aspects of their reputation and pick

targets who are already rejected. Overall, the findings support conceptualizations of bullying as a unique social phenomenon that interferes with functioning in the peer group (Salmivalli, 2010).

#### 4.1 | Implications and suggestions for future research

Our findings provide empirical support for suggestions (Bauman, 2016; Furlong et al., 2010; Jia & Mikami, 2018; Volk et al., 2014) that assessment of victimization through bullying may require an extension of current instruments, by explicitly addressing experiences with the key characteristics of the bullying definition.

More specificity in the assessment of bullying seems especially relevant for prevalence estimation and intervention evaluation. First, it could reduce the variability in victimization through bullying estimates across studies (Cook et al., 2010). Without addressing the key characteristics of bullying, about three out of five children may be incorrectly diagnosed as victims of bullying according to the definition. Moreover, differentiation between bullying and general aggression is relevant when deciding about or evaluating interventions. Strategies to tackle physical fighting and other forms of aggression seem to be unsuccessful in preventing bullying perpetration (Espelage et al., 2013; Van Schoiack-Edstrom et al., 2002) whereas some bullying prevention programs are not effective at preventing violence and aggression (Ferguson, Miguel, Kilburn, & Sanchez, 2007). Therefore, tackling bullying calls for unique strategies. For example, the power imbalance of bullying calls for teaching children to defend themselves by finding support, and the social status goals that often characterize bullying might be addressed by providing bullies with alternative prosocial strategies to achieve these goals (Ellis, Volk, Gonzalez, & Embry, 2015).

Refining the assessment of victimization through bullying for such research goals may be simple: the addition of seven additional questions to the BVQ per bully achieved 56.9% more precision in the differentiation between victims of bullying and victims of general aggression. In developing the questions, we balanced accuracy with practical use. For example, asking for all possible types of power imbalance involved in bullying would lead to a lengthy questionnaire, while only asking a general question about "whether the other was more powerful" would be too abstract. We therefore focused on two types of power imbalance (in strength and popularity) that had been used previously (Felix et al., 2011; Malecki et al., 2015), had the highest loadings on physical and social factors distinguishing power imbalance (Nelson et al., 2019), and were often experienced by our participants. The questions are also practical to use if researchers aim to omit peer nominations; it could be investigated if researchers can skip the question about the names of the bullies and replace these names in the items by "the bully who victimized you most frequently", and "second-" and "third-most frequently". Despite these theoretical and practical strengths, the additional questions should not replace the BVQ but should be used to complement it. We showed that the BVQ functioned as a "gatekeeper", by helping to identify victims with worse adjustment problems that are conceptually related to victimization.

More generally, our findings highlight the importance for bullying researchers in general of differentiating clearly between bullying and more general aggression in their communication with other researchers and practitioners, and of explicating that bullying is a separate and more harmful form of aggression that has specific characteristics.

Future research with even larger and diverse sample sizes is needed to test (a) whether children's self-reported experiences of intensity can be associated with maladjustment correlates, (b) whether latent profiles can be distinguished based on experiences of key characteristics, and whether the findings differ (c) across age or grade groups, (d) between offline and online victimization (Corcoran, Guckin, & Prentice, 2015), and (e) depending on the number of bullies. Moreover, future research can examine the predictive validity of assessing each key characteristic.

Lastly, future research could extend the current study by differentiating the direct and indirect victimization. Low frequency acts of indirect victimization may be perceived as more harmful/intense, and have a greater negative impact, than low frequency experiences of direct victimization. For example, embarrassing, reputation-damaging information that is spread through rumors may be harmful even if it is only done once, because damaging

one's reputation may persist. It would also be relevant to investigate whether the frequency of direct/indirect forms of victimization is related to different types of power imbalance. For example, indirect forms may be more strongly related to popularity differences while direct forms may be more strongly associated with differences in strength.

## 4.2 | Limitations and strengths

Our results need to be interpreted with some limitations in mind. First, the participating children were in schools that participated in an anti-bullying program, and thus were familiar with the differences between bullying and other forms of aggression: this is part of the curriculum. Our findings might have overestimated the proportion of children who adhered to the bullying definition. Second, we used the victims' perspective to assess the definition's key characteristics, but victims might not always be the most reliable informants. They can misinterpret bullies' goals, adjust their recollections of victimization to suppress negative emotions (Gross, 1998), or underreport their problems because of social desirability. Although victims' perceptions determine the impact of victimization (Olweus, 2013), we recognize that a multi-informant perspective might also be valuable. Last, we could not examine the temporal order of associations between the key characteristics and adjustment. However, we were interested in correlates of victimization through bullying (concurrent validity), regardless of whether those are precursors or consequences of victimization.

This study was the first to assess to what extent the most commonly used method to assess victimization can discriminate between victimization through bullying and other types of peer aggression, in accordance with its definition. We relied on information from a large sample of 1,738 children, including 138 systematic victims of bullying; we focused on their experiences with specific bullies to prevent them from reporting the presence of different key characteristics across bullying occasions. Using these data, we showed that less than half of the children who reported victimization in the widely used BVQ also experienced all key characteristics of the bullying definition, and that victims who endorsed all key characteristics had poorer emotional, relational, and social status adjustment, which conceptually related to victimization through bullying as compared with victimization through general aggression. Providing children with a multicomponent definition, and hoping that they will take this into account when responding to questions about how frequently this happened to them (ignoring the power imbalance, intensity, and goal-directedness) may not be the most valid approach to differentiating victimization through bullying from victimization through general aggression.

## ACKNOWLEDGMENTS

This research is funded by the Dutch Science Foundation (NWO VICI 453-14-016; NWO VENI 451-17-013).

## CONFLICT OF INTEREST

We declare that we do not have any potential sources of conflict of interest.

## DATA AVAILABILITY STATEMENT

The data used in this study are available upon reasonable request (contact the corresponding author), but are not publicly available due to privacy restrictions.

## ORCID

Tessa M. L. Kaufman  <https://orcid.org/0000-0003-0191-953X>

Gijs Huitsing  <https://orcid.org/0000-0001-7826-4178>

René Veenstra  <https://orcid.org/0000-0001-6686-6307>

## REFERENCES

- Bauman, S. (2016). Do we need more measures of bullying? *Journal of Adolescent Health, 59*, 487–488. <https://doi.org/10.1016/j.jadohealth.2016.08.021>
- Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*, 497–529.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, 57*, 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Boor-Klip, H. J., Segers, E., Hendrickx, M. M. H. G., & Cillessen, A. H. N. (2016). Development and psychometric properties of the classroom peer context questionnaire. *Social Development, 25*, 370–389. <https://doi.org/10.1111/sode.12137>
- Cook, C. R., Williams, K. R., Guerra, N. G., & Kim, T. E. (2010). Variability in the prevalence of bullying and victimization: A cross-national and methodological analysis. In S. R. Jimerson, S. M. Swearer, & D. L. Espelage (Eds.), *Handbook of bullying in schools: An international perspective* (pp. 347–362). New York, NY: Routledge/Taylor & Francis Group.
- Corcoran, L., Guckin, C., & Prentice, G. (2015). Cyberbullying or cyber aggression? A review of existing definitions of cyber-based peer-to-peer aggression. *Societies, 5*, 245–255. <https://doi.org/10.3390/soc5020245>
- Ellis, B. J., Volk, A. A., Gonzalez, J., & Embry, D. D. (2015). The meaningful roles intervention: An evolutionary approach to reducing bullying and increasing prosocial behavior. *Journal of Research on Adolescence, 26*, 622–637. <https://doi.org/10.1111/jora.12243>
- Espelage, D. L., Low, S., Polanin, J. R., & Brown, E. C. (2013). The impact of a middle school program to reduce aggression, victimization, and sexual violence. *Journal of Adolescent Health, 53*, 180–186. <https://doi.org/10.1016/j.jadohealth.2013.02.021>
- Felix, E. D., Sharkey, J. D., Green, J. G., Furlong, M. J., & Tanigawa, D. (2011). Getting precise and pragmatic about the assessment of bullying: The development of the California Bullying Victimization Scale. *Aggressive Behavior, 37*, 234–247. <https://doi.org/10.1002/ab.20389>
- Ferguson, C. J., Miguel, C. S., Kilburn, J. C., & Sanchez, P. (2007). The effectiveness of school-based anti-bullying programs: A meta-analytic review. *Criminal Justice Review, 32*, 401–414. <https://doi.org/10.1177/0734016807311712>
- Frey, K. S., Pearson, C. R., & Cohen, D. (2015). Revenge is seductive, if not sweet: Why friends matter for prevention efforts. *Journal of Applied Developmental Psychology, 37*, 25–35. <https://doi.org/10.1016/j.appdev.2014.08.002>
- Fullchange, A., & Furlong, M. J. (2016). An exploration of effects of bullying victimization from a complete mental health perspective. *Sage Open, 6*(1), 215824401562359. <https://doi.org/10.1177/2158244015623593>
- Furlong, M. J., Sharkey, J. D., Felix, E. D., Tanigawa, D., & Green, J. G. (2010). Bullying assessment: A call for increased precision of self-reported procedures. In S. R. Jimerson, S. M. Swearer, & D. L. Espelage (Eds.), *Handbook of bullying in schools: An international perspective* (pp. 329–345). New York, NY: Routledge.
- Green, J. G., Felix, E. D., Sharkey, J. D., Furlong, M. J., & Kras, J. E. (2013). Identifying bully victims: Definitional versus behavioral approaches. *Psychological Assessment, 15*, 651–657. <https://doi.org/10.1016/j.drugalcdp.2008.02.002>
- Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology, 2*, 271–299. <https://doi.org/10.1037/1089-2680.2.3.271>
- Huitsing, G., Lodder, G. M. A., Browne, W. J., Oldenburg, B., Van der Ploeg, R., & Veenstra, R. (Under review). A large-scale replication of the effectiveness of the KiVa anti-bullying program: A randomized controlled trial in the Netherlands.
- Hunter, S. C., Boyle, J. M. E., & Warden, D. (2007). Perceptions and correlates of peer-victimization and bullying. *British Journal of Educational Psychology, 77*, 797–810. <https://doi.org/10.1348/000709906X171046>
- Jia, M., & Mikami, A. (2018). Issues in the assessment of bullying: Implications for conceptualizations and future directions. *Aggression and Violent Behavior, 41*, 108–118. <https://doi.org/10.1016/j.avb.2018.05.004>
- Kärnä, A., Voeten, M., Little, T. D., Poskiparta, E., Kaljonen, A., & Salmivalli, C. (2011). A large-scale evaluation of the KiVa anti-bullying program: Grades 4–6. *Child Development, 82*, 311–330. <https://doi.org/10.1111/j.1467-8624.2010.01557.x>
- Lee, T., & Cornell, D. (2009). Concurrent validity of the Olweus bully/victim questionnaire. *Journal of School Violence, 9*, 56–73. <https://doi.org/10.1080/15388220903185613>
- Malecki, C. K., Demaray, M. K., Coyle, S., Geosling, R., Yu, S., Wheaton College, R., & Becker, L. D. (2015). Frequency, power differential, and intentionality and the relationship to anxiety, depression, and self-esteem for victims of bullying. *Child & Youth Care Forum, 115*–131. <https://doi.org/10.1007/s10566-014-9273-y>
- Nelson, H. J., Kendall, G. E., Burns, S. K., Schonert-Reichl, K. A., & Kane, R. T. (2019). Measuring 8 to 12 year old children's self-report of power imbalance in relation to bullying: Development of the Scale of Perceived Power Imbalance. *BMC Public Health, 19*, 1–12. <https://doi.org/10.1186/s12889-019-7375-z>
- Olthof, T., Goossens, F. A., Vermande, M. M., Aleva, E. A., & Van der Meulen, M. (2011). Bullying as strategic behavior: Relations with desired and acquired dominance in the peer group. *Journal of School Psychology, 49*, 339–359. <https://doi.org/10.1016/j.jsp.2011.03.003>
- Olweus, D. (1978). *Aggression in the schools: Bullies and whipping boys*. Washington, DC: Hemisphere (Wiley).
- Olweus, D. (1993). *Bullying at school: What we know and what we can do*. Malden, MA: Blackwell Publishing.

- Olweus, D. (1996). *The revised Olweus bully/victim questionnaire*. Bergen, Norway: Research Center for Health Promotion (HEMIL Center), University of Bergen.
- Olweus, D. (2013). School bullying: Development and some important challenges. *Annual Review of Clinical Psychology*, 9, 751–780. <https://doi.org/10.1146/annurev-clinpsy-050212-185516>
- Reijntjes, A., Vermande, M. M., Goossens, F. A., Olthof, T., Van de Schoot, R., Aleva, E. A., & Van der Meulen, M. (2013). Developmental trajectories of bullying and social dominance in youth. *Child Abuse and Neglect*, 37, 224–234. <https://doi.org/10.1016/j.chiabu.2012.12.004>
- Salmivalli, C. (2010). Bullying and the peer group: A review. *Aggression and Violent Behavior*, 15, 112–120. <https://doi.org/10.1016/j.avb.2009.08.007>
- Solberg, M. E., & Olweus, D. (2003). Prevalence estimation of school bullying with the Olweus bully/victim questionnaire. *Aggressive Behavior*, 29, 239–268. <https://doi.org/10.1002/ab.10047>
- Taub, J. (2002). Evaluation of the second step violence prevention program at a rural elementary school. *School Psychology Review*, 31, 186–200.
- Vaillancourt, T., McDougall, P., Hymel, S., Krygsman, A., Miller, J., Stiver, K., & Davis, C. (2008). Bullying: Are researchers and children/youth talking about the same thing? *International Journal of Behavioral Development*, 32, 486–495. <https://doi.org/10.1177/0165025408095553>
- Van der Ploeg, R., Steglich, C., & Veenstra, R. (2020). The way bullying works: How new ties facilitate the mutual reinforcement of status and bullying in elementary schools. *Social Networks*, 60, 71–82. <https://doi.org/10.1016/j.socnet.2018.12.006>
- Van Schoiack-Edstrom, L., Frey, K. S., & Beland, K. (2002). Changing adolescents' attitudes about relational and physical aggression: An early evaluation of a school-based intervention. *School Psychology Review*, 31, 201–215.
- Veenstra, R., Lindenberg, S., Munniksmas, A., & Dijkstra, J. (2010). The complex relation between bullying, victimization, acceptance, and rejection: Giving special attention to status, affection, and sex differences. *Child Development*, 81, 480–486. <https://doi.org/10.1111/j.1467-8624.2009.01411.x>
- Veenstra, R., Lindenberg, S., Zijlstra, H., De Winter, A., Verhulst, F. C., & Ormel, J. (2007). The dyadic nature of bullying and victimization: Testing a dual-perspective theory. *Child Development*, 78, 1843–1854. <https://doi.org/10.1111/j.1467-8624.2007.01102.x>
- 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. <https://doi.org/10.1016/j.avb.2014.06.008>
- Volk, A. A., Dane, A. V., & Marini, Z. A. (2014). What is bullying? A theoretical redefinition. *Developmental Review*, 34, 327–343. <https://doi.org/10.1016/j.dr.2014.09.001>
- Volk, A. A., Veenstra, R., & Espelage, D. L. (2017). So you want to study bullying? Recommendations to enhance the validity, transparency, and compatibility of bullying research. *Aggression and Violent Behavior*, 36, 34–43. <https://doi.org/10.1016/j.avb.2017.07.003>
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>
- Ybarra, M. L., Espelage, D. L., & Mitchell, K. J. (2014). Differentiating youth who are bullied from other victims of peer-aggression: The importance of differential power and repetition. *Journal of Adolescent Health*, 55, 293–300. <https://doi.org/10.1016/j.jadohealth.2014.02.009>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Kaufman TML, Huising G, Veenstra R. Refining victims' self-reports on bullying: Assessing frequency, intensity, power imbalance, and goal-directedness. *Social Development*. 2020;00: 1–16. <https://doi.org/10.1111/sode.12441>