DOI: 10.1111/cdev.13866

EMPIRICAL ARTICLE

CHILD DEVELOPMENT

# Does defending victimized peers put youth at risk of being victimized?

**Sarah T. Malamut**[1,2] | **Jessica Trach**[1] | **Claire F. Garandeau**[1] | **Christina Salmivalli**[1]

[1]INVEST Research Flagship Center, University of Turku, Turku, Finland

[2]Behavioural Science Institute, Radboud University, Nijmegen, Netherlands

**Correspondence**
Sarah T. Malamut, INVEST Research Flagship Center, University of Turku, Turku, Finland.
Email: sarah.malamut@utu.fi or
Email: sarah.malamut@ru.nl

**Funding information**
Academy of Finland, Grant/Award Number: 320162; Eunice Kennedy Shriver National Institute of Child Health and Human Development, Grant/Award Number: F32HD100054

**Abstract**

Defending peers who have been bullied is often thought to put defenders at risk of becoming victimized themselves. The study investigated the concurrent and prospective associations between defending and (peer- and self-reported) victimization, and examined popularity and classroom norms as potential moderators. Participants included 4085 Finnish youth (43.9% boys; $M_{age}$ = 14.56, $SD$ = .75; 97% born in Finland). Concurrently, defending was positively associated with self-reported victimization in classrooms with high bullying-popularity norms ($b$ = .28, $SE$ = .16). Defending was negatively associated with peer-reported victimization in classrooms with high defending-popularity norms ($b$ = −.07, $SE$ = .03). Defending was not significantly associated with future victimization, suggesting that it is generally not a risk factor for victimization.

## DOES DEFENDING VICTIMIZED PEERS PUT YOUTH AT RISK OF BEING VICTIMIZED?

To reduce school bullying, which is a serious problem that impacts youth's mental and physical health (e.g., Christina et al., 2021; Schoeler et al., 2018), many intervention efforts focus on empowering peer bystanders to defend victimized peers. Defending refers to behavior that supports a victimized peer, either directly by attempting to stop the bullying or indirectly by offering emotional support to the victim to aid their recovery (Ma et al., 2019; Trach et al., 2010). Several studies indicate that bystanders' defending of victimized peers may be associated with reduced levels of bullying in the peer group, as well as lower concurrent levels of negative adjustment for victimized youth (e.g., Ma & Chen, 2019; Salmivalli et al., 2011; Williford et al., 2012). Research has also shown that defending can lead to increased popularity (van der Ploeg et al., 2017), an indicator of how well-known, socially central, and emulated youth are among peers (Cillessen & Rose, 2005). Yet, few adolescents consistently defend others (e.g., Ma et al., 2019), presumably out of fear of becoming victimized themselves (Strindberg et al., 2020). Defending is widely described as a 'risky behavior' in research on the topic and has been found to be concurrently associated with higher rates of victimization (e.g., Lambe et al., 2019). However, a recent meta-analysis found that the effect size for this association was small and was only significant when youth self-reported their defending behavior (Ma et al., 2019). Moreover, only two empirical investigations on the prospective links between defending and risks of becoming victimized have been conducted; one found

The data necessary to reproduce the analyses presented here are not publicly available. Analytic code and materials are available from the first author. The analyses presented here were not preregistered.

This research was supported by the INVEST Research Flagship, as well as by the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health under Award Number F32HD100054. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors declare they have no conflict of interest.

that defending was a risk factor for future victimization (by the bully of the defended peer; Huitsing et al., 2014) and the other found that defending was associated with decreased victimization (Meter & Card, 2015).

Given the strong emphasis of many anti-bullying intervention efforts on encouraging youth to defend victimized peers (e.g., Salmivalli et al., 2021), it is essential to know whether defending one's peers puts youth at increased risk for being victimized themselves. To this end, the current study will clarify the concurrent association between defending and victimization, as well as extend the previous literature by examining the longitudinal effects of defending on victimization. Whether or not defending is related to concurrent and future victimization may depend on youth's individual characteristics as well as the norms of the classroom. Thus, we will consider the potential moderating effects of both individual (i.e., popularity level) and contextual factors (i.e., descriptive norms and popularity norms of bullying and defending). Moreover, given that correlates of victimization differ based on informant (e.g., Bouman et al., 2012), we will examine these effects for both peer- and self-reported victimization. In the current study, we examine these questions in early adolescence, which is a developmental period characterized by heightened concern for peer status and sensitivity to social norms (e.g., Laursen & Veenstra, 2021).

## Does defending lead to victimization?

A common narrative found in the bullying literature to explain the relatively small proportion of individuals who defend victimized peers is that defending is a socially risky behavior, with potential negative consequences for those who try to help the victim (e.g., Pozzoli et al., 2012). This may be particularly relevant in adolescence, when defending behaviors tend to decline (Trach et al., 2010). Youth who bully may seek revenge for being challenged or confronted, and thus defenders may become targets themselves. Furthermore, youth who bully may also want to demonstrate to the peer group that attempts to thwart their aggression will be punished, as well as discourage future defending as it could result in a loss of social power for them. Indeed, it is often argued that youth may be reluctant to defend their victimized peers because they are fearful of becoming the next target if they get involved (e.g., the *retaliation hypothesis*; Huitsing et al., 2014). In support of this assumption, qualitative studies have found that participants often state fear of retaliation as one reason that other youth might not intervene to help a peer being bullied (Spadafora et al., 2020; Strindberg et al., 2020; Thornberg et al., 2012). However, it is also conceivable that standing up for victimized peers sends a message to actual and potential perpetrators that one is not afraid to challenge their power. As victimization has been found to be associated with social withdrawal or a lack of assertiveness (Wei & Chen, 2008), which can signal vulnerability, active intervention in favor of the victimized peer might also protect youth who defend from being the targets of peer abuse in the future.

Despite the evidence that youth are apprehensive to defend out of fear of being victimized, the research findings are mixed with respect to *actual* experiences of victimization following defending actions. Two recent reviews of the defending literature concluded that victimization was positively correlated with current defending among early adolescents (Lambe et al., 2019; Ma et al., 2019). It is important to note, however, that this effect was small and depended on the informant of defending behavior, with stronger associations for self-reported defending (Ma et al., 2019). On the other hand, limited research has examined whether defending leads to victimization over time and has not resulted in a consistent pattern of findings. Using social network analyses, Huitsing et al. (2014) found that defenders (aged 8–11) were at risk of becoming victimized by the bullies of the victims they defended. In contrast, Meter and Card (2015) found that defending was associated with a decrease in peer-reported victimization over time in a sample of youth aged 11–15 years old. Thus, it is still unclear whether defending puts youth at risk for victimization or whether it might actually help to protect them from peer harassment. Furthermore, no study has yet investigated possible moderators of this association. Both the concurrent and prospective associations between defending and victimization may depend on youth's individual characteristics and/or the social context, as well as the informant of victimization.

## The role of defenders' popularity status

Not all defenders may be similarly at risk to experience victimization themselves. For example, youth with high levels of popularity (i.e., prestige and visibility in the peer group: Cillessen & Marks, 2011) may have enough social resources and social power to defend their peers without experiencing negative repercussions. Although there is growing evidence that high-popular youth can also be targets of aggression (e.g., Andrews et al., 2016; Dawes & Malamut, 2018; Malamut et al., 2020), the social benefits associated with popularity may still buffer youth against potential retaliation after defending (e.g., Garandeau et al., 2022; Peets et al., 2015).

Popular adolescents also tend to have better social skills (Andreou, 2006) and higher social intelligence (Meijs et al., 2010) than their peers, which might allow them to defend more effectively, that is, in a way that dissuades counter attacks. Thus, even though popular youth are not immune to victimization, they may still be better equipped to defend their peers with minimal consequences compared to less popular youth. Consistent with this idea, Malamut, Trach, et al. (2021) found that defending was only a risk factor for elevated depressive

symptoms over time for youth who were already vulnerable (i.e., those low in popularity and high in victimization). Therefore, defending does indeed appear to be riskier and more costly for low-popular youth with less social support.

## The role of classroom norms

In early adolescence in particular, an increasing number of young people begin to prioritize their status among peers (LaFontana & Cillessen, 2010) and become motivated to conform to the norms of the peer group. Youth have been found to behave in ways consistent with the social norms of the peer group, as well as in ways that are perceived to be associated with high status (Brechwald & Prinstein, 2011). Furthermore, consistent with the "social misfit" model (Wright et al., 1986), those who deviate from the norms of the peer group are typically more at risk of being victimized (e.g., Sentse et al., 2007). Thus, the degree to which defending is (in)congruent with the norms and values of the peer group is likely related to whether defending is associated with victimization—especially for adolescents. Consistent with the influence-compatibility model (Laursen & Veenstra, 2021), conformity peaks in early adolescence, and adolescents tend to behave in ways consistent with the norms of their peer group. Individuals who are insufficiently compatible with their peer group face potential social consequences (Laursen & Veenstra, 2021). Therefore, peer group norms regarding bullying and defending (measured in this study at the classroom level) may impact the extent to which defending is a risky behavior in early adolescence.

Two types of norms (descriptive norms and popularity norms) can be considered for both bullying and defending. *Descriptive norms* describe the average level of a behavior in a classroom, whereas *popularity norms* refer to the extent to which a behavior is associated with popularity (generally operationalized as the within-classroom correlation between popularity and the behavior of interest). In classrooms where bullying is more common (i.e., higher bullying descriptive norms), bullying is more normative and therefore may be more tolerated, which could imply that any action that counteracts the bullying is more likely to be punished by peers. A similar process would likely occur in classrooms where defending is *less* common (i.e., lower defending descriptive norms). In both cases, in concordance with the social misfit hypothesis, defending one's peers may be seen as a violation of the norms of the peer group, and subsequently may lead to becoming a target. In addition to bullying and defending descriptive norms in the entire classroom, the behaviors of popular students can be particularly influential. In classrooms where popular students engage in bullying (i.e., higher bullying popularity norms), bullying is typically more accepted (e.g., Dijkstra et al., 2008). Given the influence that popular youth have in the peer group (Logis et al., 2013), it has been speculated that youth need more courage to defend victimized peers in classrooms where bullies enjoy higher status (Pouwels et al., 2018). On the other hand, in classrooms where popular youth are more likely to defend victimized youth (i.e., higher defending popularity norms), defending may be less risky. Nevertheless, the possible moderating effect of classroom norms on the associations of defending with victimization is not yet known.

## Peer- versus self- reported victimization

In addition to the individual-level and classroom-level moderators that could impact the link between defending and victimization, we also will examine whether the same pattern of effects is present for both peer- and self-reported victimization. Past research has consistently found distinct profiles of adjustment when considering different informants of victimization (e.g., Bouman et al., 2012; Gromann et al., 2013). Peer-reported victimization represents youth's reputation in the peer group, relies on visibility and awareness from peers, and tends to be associated with interpersonal difficulties (e.g., Scholte et al., 2013). If classmates recognize that an adolescent has defended victimized peers, they may also be aware of potential retaliation from the perpetrator (i.e., there may be a positive association between defending and peer-reported victimization). Alternately, peers may not always be aware of the victimization experiences of youth who defend, especially as defending peers may signal a level of assertiveness that prevents defenders from having a reputation for being victimized (resulting in a negative or non-significant association between defending and peer-reported victimization). In contrast, self-reported victimization represents youth's perceptions of their experiences and tends to be more strongly related to intrapersonal difficulties (Scholte et al., 2013). As individuals are necessarily aware of whether they have defended victimized peers and whether they have been victimized themselves, there could be a positive association between defending and self-reported victimization, but not peer-reported victimization. Thus, it is critical to examine the associations between victimization and both peer- and self-reported victimization.

## The current study

Past studies have found a positive concurrent relation between defending and victimization and have speculated that defending may put youth at risk for future victimization (e.g., Pozzoli et al., 2012). However, only limited research has examined the longitudinal effect of defending on victimization (Huitsing et al., 2014; Meter & Card, 2015) and has resulted in mixed findings.

Furthermore, previous studies have not accounted for individual- and classroom-level characteristics that may moderate this association. To address these gaps in the literature, the current study examined the concurrent and prospective links between defending and victimization and five possible moderators (i.e., individual popularity status and four classroom norms: bullying descriptive norms, defending descriptive norms, bullying popularity norms, and defending popularity norms).

Due to the inconsistency in previous findings, we cannot hypothesize that the main longitudinal effect of defending on victimization will be either positive or negative. Rather, we anticipate that this longitudinal association will depend on other factors. As defending victimized peers is a potentially risky behavior, it may require a certain level of status or social resources to defend without incurring negative social consequences (e.g., Garandeau et al., 2022; Peets et al., 2015). Thus, we anticipated that defending would be positively associated with victimization (concurrently and prospectively) at low levels of popularity. Furthermore, given the importance of adherence to social norms in adolescence (e.g., Brechwald & Prinstein, 2011), the link between defending and victimization likely depends on the norms of the classroom. We expected that defending would be more likely to be positively associated with concurrent and prospective victimization in classrooms with high average levels of bullying (bullying descriptive norms) and classrooms in which bullies were more popular (bullying popularity norms). On the other hand, we expected defending to be negatively associated with concurrent and prospective victimization in classrooms with high average levels of defending (defending descriptive norms) and classrooms in which defenders were more popular (defending popularity norms). All hypotheses regarding the moderators were confirmatory; however, we refrained from making a directional hypothesis regarding the main effect of defending on victimization. In addition, we separately tested for the effect of defending on peer- and self-reported victimization, due to known differences in the correlates of victimization between informants (e.g., Bouman et al., 2012).

# METHOD

## Participants and procedure

Participants were drawn from the KiVa program evaluation (see Kärnä et al., 2013). The data from this project included Finnish students in grade 7–9 from 78 secondary schools that were randomly assigned to the intervention or control condition (39 schools in each condition). Active parental consent was obtained from 87.4% of the target sample. Most participants reported being born in Finland (97%). The current study used data from control schools only to avoid biases in the associations between the study variables due to the intervention and to make findings generalizable to other contexts, that is, without a formal, school-wide anti-bullying intervention. A total of 35 control schools provided data at the time points of interest.

Data collection occurred in three waves over the course of a year. Wave 1 occurred at the end of one academic year (May 2008), and Waves 2 and 3 were collected in the following December and May, respectively. Given our interest in whether defending in a specific context was associated with future victimization, we focused our analyses on the second and third waves of data (subsequently referred to as T1 and T2) as they were collected within the same school year and the students remained in the same classroom at both time points. Classrooms with fewer than 14 students and/or classrooms with participation rates lower than 60% were excluded from the current analysis, to increase the reliability of peer nomination items (Cillessen & Marks, 2011). The final sample included 4085 students (43.9% boys; T2 $M_{age}$ = 14.56, $SD$ = .75; 97% born in Finland) in grade 8 (53.1%) and grade 9 (46.9%) from 241 classrooms. The majority of students (83.5%) participated in both waves of the data collection. Youth who participated at both time points did not significantly differ from those who only responded at T1 in terms of defending, peer- or self-reported victimization, or popularity at T1 ($ps > .07$). At T1, less than 5% of data was missing on any variable. There was 6.8% missingness for T2 peer-reported victimization and 16.9% missingness for T2 self-reported victimization. Missing values were estimated for the variables of interest using the Expectation Maximization procedure, with all individual-level variables as predictors (Scheffer, 2002).

Teachers supervised the administration of the online questionnaires to students during regular school hours, and were provided with detailed instructions regarding the procedure 2 weeks prior to data collection. All students were reassured of the confidentiality of their answers and informed that participation was voluntary. Within the online questionnaires, the order of the questions, items, and scales were randomized.

## Measures

### Individual-level variables

#### *Gender*

Participants reported their self-identified gender, which was then dummy coded as 0 = *girl*, 1 = *boy*.

#### *Defending*

Defending at T1 was measured using the Participant Role Questionnaire (PRQ; Salmivalli & Voeten, 2004), which includes three items describing common behaviors that youth may engage in to defend a victimized peer (i.e., "Tries to make others stop

bullying," "Comforts the victim or encourages him/her to tell the teacher about the bullying," "Tells the others to stop bullying"). Participants could nominate an unlimited number of classmates for each item. The received nominations were summed and divided by the number of possible nominators within each class to form a proportion score for each participant. The final defending score was created by averaging the proportion scores across the three items for each student, with scores ranging from 0 to 1 (Cronbach's $\alpha$ = .89 at T1).

### *Popularity*

Participants' popularity was assessed using peer nominations at T1. Students were asked to nominate their classmates who were the "most popular." For each participant, the received (unlimited) nominations were summed and divided by the number of possible nominators to form a proportion score, with scores ranging from 0 to 1.

### *Victimization*

*Peer-reported victimization* was measured using three items from the PRQ (Salmivalli & Voeten, 2004; i.e., "s/he is called names and made fun of," "s/he is pushed and hit," "s/he is usually talked about with a bad tone"). Participants could nominate an unlimited number of classmates for each item. The received nominations were summed and divided by the number of possible nominators within each class to form a proportion score for each participant. The final victimization score was created by averaging the proportion scores across the three items for each student, with scores ranging from 0 to 1 (Cronbach's $\alpha$ = .74 and .71 at T1 and T2). *Self-reported victimization* was assessed using the revised Olweus Bully/Victim questionnaire (Olweus, 1996). Participants reported how frequently they experienced different forms of victimization (e.g., "I was hit, kicked, or pushed," "I was called nasty names or laughed in my face or hurt by insults"), using a 5-point scale (0 = *not at all*, 4 = *several times a week*). Participants' responses on the 10 items were averaged to create a total self-reported victimization score (Cronbach's $\alpha$ = .87 and .92 at T1 and T2).

## Classroom-level predictor variables

### *Bullying norms*

Bullying was assessed using three items from the PRQ (i.e., "starts bullying," "makes the others join in the bullying," "always finds new ways of harassing the victim"). The received (unlimited) nominations for each item were summed and divided by the number of possible nominators within each class. Bullying scores were created by averaging the proportion scores across the three items for each student at T1, with scores ranging from 0 to 1 (Cronbach's $\alpha$ = .93). *Bullying descriptive norms* were operationalized as the classroom-level average of peer-nominated bullying. They were computed by averaging the individual proportion scores of bullying of all students in the classroom. *Bullying popularity norms* were calculated as the within-classroom correlation between peer-nominated bullying and peer-nominated popularity.

### *Defending norms*

Defending scores were calculated as described above. *Defending descriptive norms* were operationalized as the classroom-level average of peer-nominated defending at T1. They were calculated by averaging the individual proportion scores of defending of all students in the classroom. *Defending popularity norms* were calculated as the within-classroom correlation between peer-nominated defending and peer-nominated popularity.

## Analytic plan

A series of multilevel models were performed to examine the concurrent and longitudinal effects of defending on victimization, along with potential individual-level and classroom-level moderators, using the lme4 package in R (Bates et al., 2015). Given our interest in Level 1 associations, as well as cross-level interactions, individual-level variables were centered at the classroom mean for accurate estimates of Level 1 slope, and classroom-level variables were grand-mean centered (see Enders & Tofighi, 2007 for more detailed discussion). To account for non-normality in the data, we used bootstrapping with 2500 sample replicates. We tested separate concurrent and longitudinal models with peer-reported and self-reported victimization as the dependent variables.

For each set of analyses, we first tested the unconditional means models and examined intraclass correlations (ICCs). Second, individual-level predictors were added into the unconditional model to test the effect of defending on victimization. Third, we added one within-level interaction (individual defending × individual popularity) to the model. Fourth, we added the classroom-level variables to the model. Fifth, we included the random slope between individual-level defending and victimization, to test whether there were statistically significant differences between classrooms in the effect of youth's defending on victimization. Lastly, we simultaneously added two cross-level interactions (individual defending × classroom descriptive norms and individual defending × classroom popularity norms) to the models. Separate models were run for bullying norms and defending norms. Gender was included as a control variable in all models, given known gender differences in rates of defending (i.e., girls tend to defend more: Ma et al., 2019). When testing the longitudinal associations

between defending and victimization, prior levels of victimization were also included as a control variable.

# RESULTS

## Descriptive statistics and intraclass correlations

Descriptive statistics and correlations are presented in Table 1. There was a modest, positive association between individual-level defending and popularity at T1 ($r$ = .19). There were weak, negative associations between defending and both peer- and self-reported victimization ($r$s = −.08 and − .06, respectively). Peer- and self-reported victimization were weakly to moderately correlated at both time points ($r$s = .27 and .16). Victimization was stable from T1 to T2, with higher levels of stability for peer-reported victimization ($r$ = .73) than for self-reported victimization ($r$ = .38). At the classroom level, there were weak, positive correlations between bullying descriptive norms and bullying popularity norms ($r$ = .07), as well as between defending descriptive norms and defending popularity norms ($r$ = .12). In contrast, there was a weak, negative correlation between bullying descriptive norms and defending descriptive norms ($r$ = −.07), and a stronger negative association between bullying popularity norms and defending popularity norms ($r$ = −.44).

ICCs for peer-reported victimization at T1 and T2 were .161 and .243, respectively, indicating that approximately 16% and 24% of the variance in peer-reported victimization was due to differences between classrooms at T1 and T2. The ICCs for self-reported victimization at T1 and T2 were .033 and .026, indicating that approximately 3% of the variance in self-reported victimization could be explained by differences between classrooms at T1 and T2. For both peer- and self-reported victimization, the majority of variance was due to individual differences.

## Concurrent levels of peer-reported victimization

### Effects of individual-level predictors

We first tested a model predicting T1 peer-reported victimization with only Level 1 predictors with fixed (across classrooms) slopes, which explained 5.3% of the individual-level variance in peer-reported victimization. After controlling for gender and popularity, defending was not significantly associated with T1 peer-reported victimization ($p$ = .26; Table 2, Model 1A). Next, we tested the within-level interaction of defending and popularity (Model 1B), which was significant ($p$ < .001; Figure 1). The concurrent relationship between defending and peer-reported victimization was negative at low and average levels of popularity, ($b$s = −.08 and −.04, $SE$s = .02, $p$s < .03, respectively), but not at high levels of popularity ($b$ = .01, $SE$ = .02, $p$ = .60). However, this interaction explained only 0.39% of the individual-level variance in peer-reported victimization.

### Effects of classroom-level predictors

We then included the main effects of the classroom-level variables (i.e., bullying descriptive and popularity norms, defending descriptive and popularity norms) as predictors of peer-reported victimization (not shown in Table 2), which explained 57.1% of the between-level variance in peer-reported victimization. After adding classroom-level predictors to the model, we tested the random slope for the association between defending and T1 peer-reported victimization. The model fit was significantly better with the random effect of defending ($p$ < .01). Bullying descriptive norms ($p$ < .001) were positively associated with T1 peer-reported victimization (see Table 2, Model

**TABLE 1** Descriptive statistics and correlations of individual (Level 1) and classroom-level (Level 2) variables

| | 1 | 2 | 3 | 4 | 5 | *M* (*SD*) |
|---|---|---|---|---|---|---|
| Individual level | | | | | | |
| 1. Defending T1 | — | | | | | .08 (.10) |
| 2. Popularity T1 | .19*** | — | | | | .11 (.17) |
| 3. Peer-reported victimization T1 | −.08*** | −.14*** | — | | | .06 (.08) |
| 4. Self-reported victimization T1 | −.06*** | −.03 | .27*** | — | | .19 (.43) |
| 5. Peer-reported victimization T2 | −.06*** | −.12*** | .73*** | .20*** | — | .05 (.07) |
| 6. Self-reported victimization T2 | −.05** | −.00 | .15*** | .38*** | .16*** | .21 (.55) |
| Classroom level | | | | | | |
| 1. Bullying descriptive norm | — | | | | | .05 (.03) |
| 2. Bullying popularity norm | .07*** | — | | | | .33 (.32) |
| 3. Defending descriptive norm | −.07*** | −.08*** | — | | | .08 (.05) |
| 4. Defending popularity norm | −.16*** | −.44*** | .12*** | | | .19 (.32) |

*Note*: *$p$ < .05. **$p$ < .01. ***$p$ < .001.

**TABLE 2** Concurrent associations between defending and victimization

| | Peer-reported victimization (T1) | | | | | Self-reported victimization (T1) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Model 1A** | **Model 1B** | **Model 1C** | **Model 1D** | **Model 1E** | **Model 2A** | **Model 2B** | **Model 2C** | **Model 2D** | **Model 2E** |
| | ***b* (*CI*)** | ***b* (*CI*)** | ***b* (*CI*)** | ***b* (*CI*)** | ***b* (*CI*)** | ***b* (*CI*)** | ***b* (*CI*)** | ***b* (*CI*)** | ***b* (*CI*)** | ***b* (*CI*)** |
| Intercept | .06*** (*.06, .07*) | .06*** (*.06, .07*) | .06*** (*.06, .06*) | .06*** (*.06, .06*) | .06*** (*.06, .06*) | .19*** (*.17, .21*) | .19*** (*.17, .21*) | .19*** (*.17, .21*) | .19*** (*.17, .21*) | .19*** (*.17, .21*) |
| Level 1 | | | | | | | | | | |
| Boy | .02*** (*.02, .03*) | .02*** (*.02, .02*) | .02*** (*.02, .03*) | .02*** (*.02, .03*) | .02*** (*.02, .03*) | .09*** (*.06, .12*) | .09*** (*.07, .12*) | .09*** (*.07, .12*) | .10*** (*.07, .12*) | .09*** (*.06, .12*) |
| Defending | −.02 (*−.05, .01*) | −.04* (*−.07, −.00*) | −.03 (*−.07, .00*) | −.03 (*−.06, .01*) | −.04 (*−.08, .00*) | .02 (*−.17, .20*) | .03 (*−.16, .21*) | .01 (*−.20, .22*) | .06 (*−.14, .28*) | .01 (*−.23, .24*) |
| Popularity | −.08*** (*−.10, −.07*) | −.09*** (*−.10, −.07*) | −.09*** (*−.10, −.07*) | −.09*** (*−.10, −.07*) | −.08*** (*−.10, −.07*) | −.08* (*−.16, −.01*) | −.08 (*−.16, .00*) | −.08 (*−.16, .00*) | −.08 (*−.16, .00*) | −.07 (*−.16, .01*) |
| Def × Pop | | .28*** (*.14, .41*) | .31*** (*.17, .44*) | .31*** (*.18, .45*) | .37*** (*.23, .52*) | | −.13 (*−.87, .61*) | −.08 (*−.88, .70*) | .028 (*−.77, .81*) | .10 (*−.74, .93*) |
| Level 2 | | | | | | | | | | |
| BDN | | | .75*** (*.64, .87*) | .78*** (*.67, .89*) | .75*** (*.64, .86*) | | | 1.66*** (*1.16, 2.16*) | 1.73*** (*1.28, 2.22*) | 1.65*** (*1.16, 2.12*) |
| BPN | | | .00 (*−.01, .02*) | .00 (*−.01, .02*) | .00 (*−.01, .02*) | | | −.03 (*−.08, .02*) | −.04 (*−.09, .01*) | −.03 (*−.09, .02*) |
| DDN | | | .01 (*−.05, .08*) | .01 (*−.05, .08*) | .01 (*−.05, .08*) | | | −.25 (*−.50, .02*) | −.26 (*−.53, .02*) | −.25 (*−.54, .03*) |
| DPN | | | −.00 (*−.01, .01*) | −.00 (*−.01, .01*) | −.00 (*−.01, .01*) | | | −.03 (*−.09, .03*) | −.03 (*−.08, .02*) | −.03 (*−.08, .02*) |
| Cross-level interactions | | | | | | | | | | |
| Defending × BDN | | | | −1.20* (*−2.09, −.32*) | | | | | −5.78* (*−11.27, −.46*) | |
| Defending × BPN | | | | .08 (*−.02, .18*) | | | | | .65* (*.09, 1.26*) | |
| Defending × DDN | | | | | .07 (*−.47, .63*) | | | | | .12 (*−3.39, 3.48*) |
| Defending × DPN | | | | | −.13* (*−.23, −.02*) | | | | | −.34 (*−.95, .30*) |
| Variance component | | | | | | | | | | |
| Between-class | .001 | .001 | .000 | .000 | .000 | .006 | .006 | .004 | .004 | .004 |
| Within-class | .005 | .005 | .005 | .005 | .005 | .177 | .177 | .175 | .174 | .17 |
| Defending slope | | | .005 | .003 | .004 | | | .240 | .209 | .228 |

*Note*: A total of 33 participants were excluded from the models with Level 2 predictors, as they were in classrooms that did not complete the relevant peer nominations for the popularity norms. *$p<.05$. **$p<.01$. ***$p<.001$.
Abbreviations: BDN, bullying descriptive norm; BPN, bullying popularity norm; DDN, defending descriptive norm; DPN, defending popularity norm.

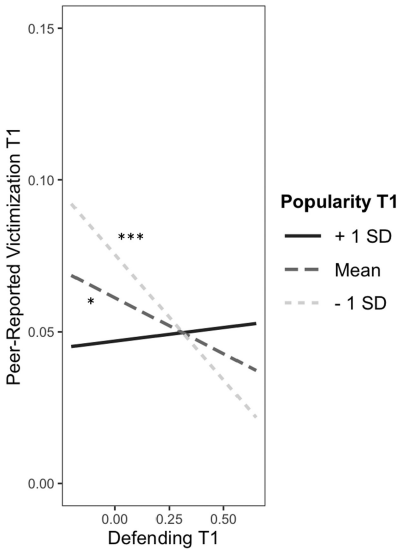


**FIGURE 1** Interaction between defending and individual popularity T1 predicting peer-reported victimization T1

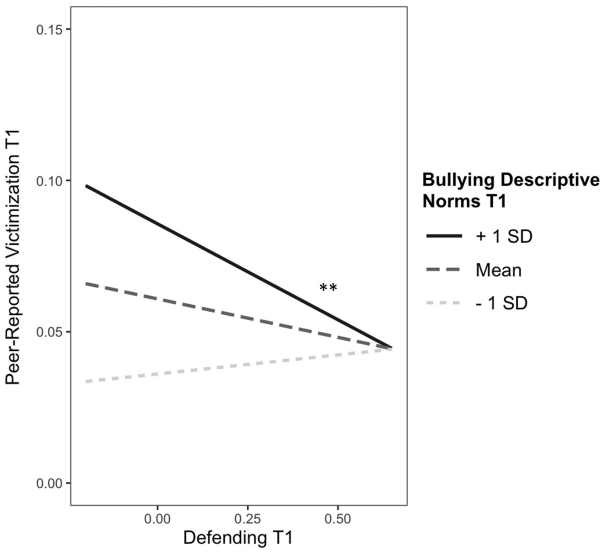


**FIGURE 2** Interaction between defending and bullying descriptive norms T1 predicting peer-reported victimization T1

1C). Bullying popularity norms, defending descriptive norms, and defending popularity norms were not significantly associated with peer-reported victimization (*p*s > .49).

Next, we tested two cross-level interactions (between defending and descriptive norms, and defending and popularity norms) in separate models for bullying norms and defending norms (Table 2, Models 1D and 1E, respectively). The association between defending and T1 peer-reported victimization was significantly moderated by bullying descriptive norms ($p$ = .006; Figure 2). When bullying was very common (high bullying descriptive norms), there was a negative association between defending and T1 peer-reported victimization ($b$ = −.06, $SE$ = .02, $p$ = .004). In classrooms with low and average bullying descriptive norms, there was no significant association between defending and T1 peer-reported victimization ($b$s = −.01 and −.03, $SE$s = .02, *p*s > .15). This cross-level interaction explained 35.8% of the variance in the random slope for the association between defending and peer-reported victimization. Bullying popularity norms did not moderate the association between defending and T1 peer-reported victimization.

Defending descriptive norms also were not a significant moderator ($p$ = .81; Table 2, Model 1E). There was, however, a significant interaction between defending and defending popularity norms ($p$ = .02), which explained 24.6% of the variance in the random slope. At high levels of defending popularity norms, there was a negative association between defending and T1 peer-reported victimization ($b$ = −.07, $SE$ = .03, $p$ = .005; Figure 3). At low and average levels of defending popularity norms, there was no significant association between defending and T1 peer-reported victimization ($b$s = .00 and −.04, $SE$s = .03 and .02, *p*s > .07).

## Concurrent levels of self-reported victimization

### Effects of individual-level predictors

The same model-building procedure was followed for self-reported victimization. We first tested a model predicting T1 self-reported victimization, with only Level 1 predictors (Table 2, Model 2A), which explained 1.2% of the individual-level variance in T1 self-reported victimization. After controlling for gender and popularity, defending was not significantly associated with T1 self-reported victimization ($p$ = .85). The within-level interaction between defending and popularity (Table 2, Model 2B) was also not significant ($p$ = .75).

### Effects of classroom-level predictors

Next, we included the main effects of the classroom-level variables (i.e., bullying norms, defending norms) as predictors of T1 self-reported victimization (not shown in Table 2), which explained 54.3% of the between-level variance in self-reported victimization. After adding classroom-level predictors to the model, we tested the random slope for the association between defending and T1 self-reported victimization, which significantly improved the model ($p$ = .01). Bullying descriptive norms were positively associated with T1 self-reported victimization ($p$ < .001; Table 2, Model 2C). Bullying popularity norms, defending descriptive norms, and defending popularity norms were not significantly associated with self-reported victimization (*p*s > .07).

Two cross-level interactions (between defending and descriptive norms, and defending and popularity norms) were tested in separate models for bullying

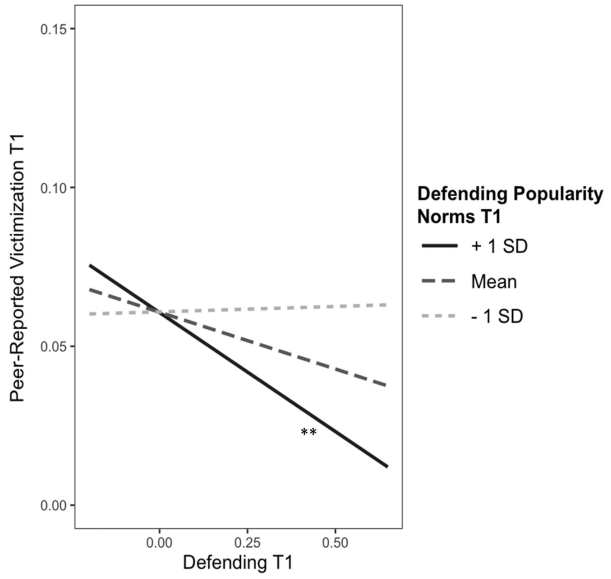


**FIGURE 3** Interaction between defending and defending popularity norms T1 predicting peer-reported victimization T1

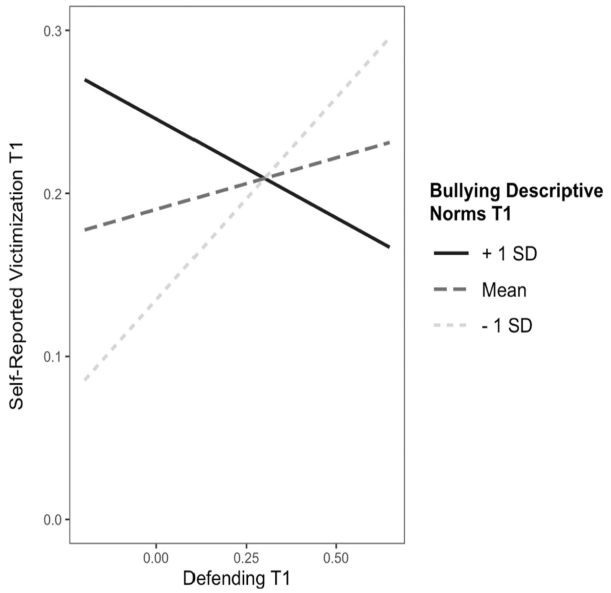


**FIGURE 4** Interaction between defending and bullying descriptive norms T1 predicting self-reported victimization T1

norms and defending norms (Table 2, Models 2D and 2E, respectively). Bullying descriptive norms were a significant moderator ($p$ = .03; Figure 4). However, this interaction only explained 1.4% of the variance in the random slope, and none of the simple slopes reached statistical significance ($b$s = −.12, .06, .25, $SE$ = .13, .11, and .15, $p$s > .09 for high, average, and low levels of defending, respectively). For those high in defending, levels of T1 self-reported victimization were higher in classrooms with lower levels of bullying descriptive norms than in classrooms with higher levels of bullying descriptive norms.

The association between defending and T1 self-reported victimization was also significantly moderated by bullying popularity norms ($p$ = .03). In classrooms with high levels of bullying popularity norms ($b$ = .28, $SE$ = .16, $p$ = .07), defending was associated with higher levels of T1 self-reported victimization, compared to classrooms with low or average levels of bullying popularity norms ($b$s = −.15 and .06, $SE$ = .13 and .11, $p$s > .25; Figure 5). This cross-level interaction explained 12.9% of the variance in the random slope for the association between defending and self-reported victimization. Neither defending descriptive norms nor defending popularity norms moderated the association between defending and T1 self-reported victimization ($p$s > .22).

## Prospective levels of peer-reported victimization

### Effects of individual-level predictors

As shown in Table 3 (Model 3A), defending was not significantly associated with T2 peer-reported victimization ($p$ = .77), when controlling for gender, T1 peer-reported victimization, and popularity. In addition, there was no significant interaction between defending and popularity ($p$ = .37; Table 3, Model 3B).

### Effects of classroom-level predictors

Next, we added the classroom-level predictors and tested the random slope for the association between defending and T2 peer-reported victimization, which

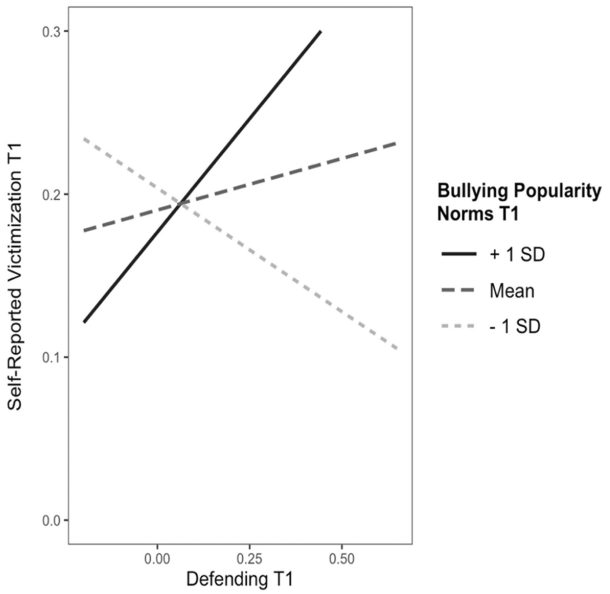


**FIGURE 5** Interaction between defending and bullying popularity norms T1 predicting self-reported victimization T1

**TABLE 3** Longitudinal associations between defending and victimization

| | Peer-reported victimization (T2) | | | | | Self-reported victimization (T2) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Model 3A** | **Model 3B** | **Model 3C** | **Model 3D** | **Model 3E** | **Model 4A** | **Model 4B** | **Model 4C** | **Model 4D** | **Model 4E** |
| | ***b* (*CI*)** | ***b* (*CI*)** | ***b* (*CI*)** | ***b* (*CI*)** | ***b* (*CI*)** | ***b* (*CI*)** | ***b* (*CI*)** | ***b* (*CI*)** | ***b* (*CI*)** | ***b* (*CI*)** |
| Intercept | .05*** (*.05, .06*) | .05*** (*.05, .06*) | .05*** (*.05, .06*) | .05*** (*.05, .06*) | .05*** (*.05, .06*) | .21*** (*.19, .23*) | .21*** (*.19, .23*) | .21*** (*.19, .23*) | .21*** (*.19, .23*) | .21*** (*.19, .23*) |
| Level 1 | | | | | | | | | | |
| Boy | .00 (*−.00, .00*) | .00 (*−.00, .00*) | .00 (*−.00, .00*) | .00 (*−.00, .00*) | .00 (*−.00, .00*) | .09*** (*.06, .13*) | .09*** (*.06, .13*) | .09*** (*.06, .13*) | .09*** (*.06, .13*) | .09*** (*.06, .13*) |
| Victim T1 | .63*** (*.61, .65*) | .63*** (*.61, .65*) | .63*** (*.61, .65*) | .63*** (*.61, .65*) | .63*** (*.61, .64*) | .46*** (*.42, .50*) | .46*** (*.42, .50*) | .46*** (*.42, .49*) | .46*** (*.42, .50*) | .46*** (*.42, .49*) |
| Defending | .00 (*−.01, .02*) | −.00 (*−.02, .02*) | .00 (*−.02, .02*) | .00 (*−.02, .02*) | −.01 (*−.03, .02*) | −.16 (*−.37, .06*) | −.15 (*−.37, .07*) | −.18 (*−.43, .07*) | −.18 (*−.42, .06*) | −.20 (*−.49, .08*) |
| Popularity | −.01** (*−.02, −.00*) | −.01*** (*−.02, −.01*) | −.01** (*−.02, −.01*) | −.01** (*−.02, −.00*) | −.01** (*−.02, −.00*) | .04 (*−.05, .14*) | .05 (*−.05, .14*) | .05 (*−.04, .15*) | .05 (*−.04, .15*) | .04 (*−.06, .14*) |
| Def × Pop | | .04 (*−.04, .11*) | .05 (*−.03, .13*) | .06 (*−.02, .14*) | .06 (*−.02, .15*) | | −.13 (*−1.01, .78*) | −.27 (*−1.18, .66*) | −.24 (*−1.17, .71*) | −.42 (*−1.47, .59*) |
| Level 2 | | | | | | | | | | |
| BDN | | | .49*** (*.35, .63*) | .48*** (*.34, .62*) | .49*** (*.36, .63*) | | | 1.87*** (*1.27, 2.49*) | 1.86*** (*1.24, 2.50*) | 1.87*** (*1.28, 2.46*) |
| BPN | | | .02* (*.00, .03*) | .01 (*−.00, .03*) | .02* (*.00, .03*) | | | −.03 (*−.10, .03*) | −.03 (*−.10, .04*) | −.03 (*−.10, .03*) |
| DDN | | | .01 (*−.07, .09*) | .01 (*−.07, .09*) | −.00 (*−.08, .08*) | | | .04 (*−.33, .37*) | .03 (*−.34, .37*) | .02 (*−.33, .40*) |
| DPN | | | .01 (*−.00, .03*) | .01 (*−.00, .03*) | .01 (*−.00, .03*) | | | .03 (*−.04, .10*) | .02 (*−.04, .10*) | .02 (*−.05, .09*) |
| Cross-level interactions | | | | | | | | | | |
| Defending x BDN | | | | .22 (*−.32, .73*) | | | | | .58 (*−5.85, 7.33*) | |
| Defending × BPN | | | | .03 (*−.03, .09*) | | | | | −.04 (*−.79, .66*) | |
| Defending × DDN | | | | | .20 (*−.11, .52*) | | | | | .57 (*−3.44, 4.61*) |
| Defending × DPN | | | | | −.02 (*−.08, .04*) | | | | | −.38 (*−.38, 1.13*) |
| Variance component | | | | | | | | | | |
| Between-class | .001 | .001 | .001 | .001 | .001 | .010 | .011 | .008 | .008 | .008 |
| Within-class | .002 | .002 | .002 | .002 | .002 | .252 | .252 | .249 | .249 | .249 |
| Defending slope | | | .002 | .002 | .002 | | | .287 | .288 | .285 |

*Note*: A total of 33 participants were excluded from the models with Level 2 predictors, as they were in classrooms that did not complete the relevant peer nominations for the popularity norms. *$p < .05$. **$p < .01$. ***$p < .001$.
Abbreviations: BDN, bullying descriptive norm; BPN, bullying popularity norm; DDN, defending descriptive norm; DPN, defending popularity norm.

significantly improved the model ($p = .002$; see Table 3, Model 3C). Both bullying descriptive norms ($p < .001$) and bullying popularity norms ($p = .04$) were positively associated with T2 peer-reported victimization. Neither defending popularity norms nor defending descriptive norms ($ps > .17$) were significantly associated with T2 peer-reported victimization. None of the cross-level interactions between defending and bullying norms (Table 3, Model 3D) and between defending and defending norms (Table 3, Model 3E) were significant ($ps > .22$).

### Prospective levels of self-reported victimization

#### Effects of individual-level predictors

As shown in Table 3 (Model 4A), defending was not a significant predictor of T2 self-reported victimization ($p = .16$). Popularity also did not significantly moderate the association between defending and T2 self-reported victimization ($p = .76$; Table 3, Model 4B).

#### Effects of classroom-level predictors

We then added the classroom-level predictors and tested the random slope for the association between defending and T2 self-reported victimization, which significantly improved the model ($p < .001$). As shown in Table 3 (Model 4C), bullying descriptive norms were positively associated with T2 self-reported victimization ($p < .001$). Bullying popularity norms, defending descriptive norms and defending popularity norms were not significantly associated with T2 self-reported defending ($ps > .32$). Finally, none of the cross-level interactions between defending and bullying norms, and between defending and defending norms, were significant (Table 3, Models 4D and 4E, respectively; $ps > .30$).

### Sensitivity analyses

Defending can be enacted in different ways (e.g., Reijntjes et al., 2016; Trach et al., 2010), and direct defending (e.g., directly confronting a bully) in particular may be more likely to be positively associated with concurrent and prospective victimization. Therefore, we conducted sensitivity analyses to address this possibility. In these analyses, we focused on the two items of direct defending ("Tries to make others stop bullying," "Tells the others to stop bullying") when measuring individual-level defending and classroom defending norms. All analyses were repeated with new variables representing individual-level direct defending and classroom norms of direct defending, and all patterns of findings remained the same as with the main set of analyses.

## DISCUSSION

Many anti-bullying interventions encourage youth to defend peers who are bullied; however, defending is often discussed as a potentially risky behavior in the research literature. Although some studies have found a positive correlation between defending and victimization (Ma et al., 2019), no clear pattern of findings has been found for whether defending actually predicts victimization over time. Furthermore, there are key individual- and classroom-level moderators that might affect whether defending is related to concurrent and future victimization. Thus, the current study examined whether defending was related to concurrent and prospective peer-reported and self-reported victimization, while considering individual-level popularity and classroom-level bullying and defending norms as potential moderators. We focused on status and classroom norms as moderators, as these are likely to be particularly relevant in early adolescence, given that peer relationships become more salient during this developmental period (Laursen & Veenstra, 2021).

Overall, our findings did not support the assumption that defending would be a risk factor for concurrent or prospective victimization. In fact, the main effect of defending on concurrent levels of victimization was negative for peer-reported victimization and non-significant for self-reported victimization. Therefore, defending itself was not associated with higher levels of victimization. Instead, youth with a reputation for defending were less likely to also have a reputation among their peers as a victim. This supports the idea that defending can signal to classmates that youth are assertive and willing to stand up to bullies, which could protect them from being victimized.

An important individual-level characteristic that could influence whether defending is related to victimization is youth's social status. Past research has found positive associations between defending and popularity (e.g., Garandeau et al., 2022), and we reasoned that defending may only be related to victimization for youth who do not have high enough status to defend their peers without experiencing social consequences. In the current study, popularity was a significant moderator of the association between defending and concurrent levels of peer-reported victimization, but not self-reported victimization. For popular youth, who typically are less likely to be seen as highly victimized by their peers (e.g., Malamut, Trach, et al., 2021), defending was indeed not related to their concurrent levels of peer-reported victimization. However, for less popular youth, defending actually appears to provide some level of protection against concurrent victimization, as those with a stronger reputation for defending were lower in peer-reported victimization compared to lower status youth who defended less. Contrary to our expectations, defending did not appear to put low-status youth at risk of being victimized, but actually mitigated their likelihood of being bullied (according to peers). Due to the concurrent

nature of this finding, it is also possible, however, that when youth are low or average in popularity, they may be more likely to defend when they are not victimized (or low in victimization) than when they are highly victimized. Alternatively, youth low or average in popularity may only be seen by peers as defenders under the condition that peers do not also see them as victims. It is important to acknowledge that this interaction did not explain a substantial amount of variance (less than 1%), so it should be interpreted with caution.

We also considered classroom-level characteristics (i.e., bullying descriptive and popularity norms, defending descriptive and popularity norms) that could affect whether defending is associated with victimization. Only bullying descriptive norms and defending popularity norms significantly moderated the association between defending and concurrent peer-reported victimization. Contrary to our expectations, defending was *negatively* associated with peer-reported victimization in classrooms with a higher frequency of bullying, whereas defending was not significantly related to victimization in classrooms with lower levels of bullying. However, as seen in Figure 2, this interaction was driven by differences in youth with low levels of defending. Youth who did not defend were higher on peer-reported victimization in classrooms with a higher level of bullying. In contrast, at high levels of defending, youth experienced similar levels of peer-reported victimization regardless of the level of bullying in their classroom. In classrooms with high levels of bullying, defending others likely signals to peers that an individual is assertive and capable of standing up for themselves, and therefore defenders are unlikely to also have a reputation as a victim.

In addition, in classrooms where the association between defending and popularity was high, defending was negatively associated with peer-reported victimization. However, defending was unrelated to peer-reported victimization in classrooms with low or average defending popularity norms. These results demonstrate that youth who defend are unlikely to have a reputation among peers as a victim in certain contexts. This effect was not found in classrooms with a low prevalence of bullying and classrooms where defenders do not tend to be popular. Still, defending was not positively associated with peer-reported victimization in these contexts either.

Nevertheless, youth who defend may report different experiences. Although we were not able to statistically compare the effects for peer- and self-reported victimization, a different pattern of results was found for the effect of defending on concurrent, self-reported victimization than was found for peer-reported victimization. This is perhaps not surprising, given previous research showing that outcomes of victimization varied depending on the assessment method used (Bouman et al., 2012). Only bullying norms (not defending norms) moderated the association between defending and self-reported victimization. Defending was associated with higher levels of concurrent self-reported victimization in classrooms where bullying was rewarded with popularity (high bullying popularity norms) compared to classrooms where bullying was not rewarded with popularity. Youth who defend were more likely to report being victimized in contexts where bullying behavior was valued. In such contexts, defenders may realize that they are going against the values of the classroom, and thus report *feeling* more targeted or isolated. At the same time, it is also conceivable that their tendency to get involved in social conflicts could contribute to being at higher risk for experiencing victimization themselves. Indeed, self-reports of victimization might indicate either biased perceptions or actual experiences of victimization (Graham & Juvonen, 1998). There was also a significant interaction between defending and bullying descriptive norms. As with peer-reported victimization, defending was positively associated with self-reported victimization in classrooms with low bullying descriptive norms, and negatively associated with self-reported victimization in classrooms with high bullying descriptive norms. However, none of the individual simple slopes were significant and this interaction only explained 1.4% of the variance. Thus, this interaction may have little practical significance.

Taken together, our attempts to clarify the concurrent associations between defending and peer victimization found that children who are known as defenders are also less likely to be seen as victimized by their classmates. Whereas defending may protect youth from having a reputation for being bullied, it also appears to put youth at risk for feeling targeted in classrooms where bullying is supported by the peer group (particularly in classes where youth who bully are also popular). The different findings for peer- versus self-reported victimization emphasize the importance of considering both informants of victimization if we are to fully understand the complexity of youth's bullying experiences. Not only are peer- and self-reports of victimization only modestly correlated, but they also identify different profiles of victimized youth with divergent adjustment (e.g., Dawes et al., 2017; Malamut, Dawes, et al., 2021).

While these findings help clarify the concurrent association between defending and victimization, our next goal was to examine whether defending was a risk factor for future victimization. Defending was not a significant predictor of prospective peer- or self- reported victimization. Even when considering several possible moderators, we did not find any evidence of defending being positively associated with future victimization. Although fear of retaliation is often given as a reason why youth are hesitant to defend (e.g., Strindberg et al., 2020; Thornberg et al., 2012), our findings do not indicate that *actual* experiences of victimization are more likely after defending, regardless of whether victimization is measured via peer- or self- report. However, the current study focused on overall levels of peer- and self- reported victimization; thus, it is still possible that youth who defend could be at

risk for retaliation from the specific perpetrator that was confronted (Huitsing et al., 2014).

## Strengths, limitations, and future directions

The strengths of this study include a multi-informant, longitudinal design, and the examination of personal and contextual moderators of the association between defending and victimization. There are also a number of limitations. First, even though the current study provides no evidence that defending is risky in terms of future victimization, it is possible that the timing of our data assessments did not allow us to detect such effects. There were only a few months between the waves, but any risks associated with defending may occur much sooner (i.e., over the span of days or weeks, rather than months). In addition, some students may have already experienced victimization as a result of their defending and subsequently stopped defending. In other words, it could be that only those who have never experienced negative consequences for their defending behavior keep defending. Future research could consider person-centered analyses of defending trajectories (e.g., stable defending, increasing defending, decreasing defending) to see whether these groups differ on victimization experiences.

Second, peer-reports are not ideal to assess change, as they are generally stable indicators of reputation (Olweus, 2013). Peer ratings or items that assess frequency/severity of victimization are more suited to assess change in peers' perceptions of youth's victimization. By including self-reported victimization, however, we still had some measure of frequency. Although we included both informants, it is important to note that our design did not allow us to make direct statistical comparisons between the effects for peer- versus self- reported victimization. Also related to our peer-reports, our study only had information regarding who is "most popular," not who is "least popular." Numerous studies have used a single item of "most popular" (e.g., Caravita & Cillessen, 2012; Garandeau et al., 2022; Parkhurst & Hopmeyer, 1998; Peets et al., 2015), including for calculating popularity norms (e.g., Garandeau et al., 2022; Peets et al., 2015). Nevertheless, having information regarding both the most and least popular students would be preferable to more clearly differentiate those who are not especially popular from those who are strongly unpopular.

Third, there are aspects related to the classroom context that could be important to consider. Although this study examined important classroom norms, we did not consider differences in how teachers handle bullying in the classroom. Relatedly, we examined defending norms in general, but another important factor could be whether, during a specific bullying incident, a student is the only person in the classroom defending the victim. That is, regardless of the norms in the classroom, students may be more likely to be victimized after defending if they were the only person to stand up for the victim in that specific instance. Thus, it is important for future studies to also consider the number of defenders in specific bullying instances, in addition to focusing on the classroom as the contextual unit of analysis.

Fourth, we did not examine who defends whom against whom. We accounted for characteristics of youth who defend and of the classroom, but the consequences of defending may also depend on characteristics of the victim or the perpetrator (e.g., their peer status). Furthermore, dyadic nominations could also better account for youth who have a certain role (e.g., defender) in one bullying incident and a different role (e.g., victim) in another.

Fifth, we did not assess whether youth's defending was actually successful, which could have important implications for the consequences of defending. Defenders who are unable to successfully deter the bully may be at greater risk for victimization. Future research should account for whether some youth are more effective at defending than others, and how this relates to potential consequences of defending. Relatedly, our study used peer-reports of defending, and it could be argued that peer-reported defending may be more likely to encapsulate successful defending as these represent the youth who are recognized by peers as someone who defends others. In other words, youth who are less successful at defending may also be less likely to have a reputation in the peer group for defending.

Lastly, this study included a sample of ethnically and economically relatively homogenous Finnish adolescents, which was representative of the Finnish population at the time of data collection. It is important for future research to replicate these findings to ensure that they generalize to other cultures, including less homogenous samples. For example, in contexts with higher levels of social inequality, other factors, such as ethnicity and socioeconomic status, may contribute more strongly to the dynamics of defending and victimization. Moreover, we examined the association between defending and victimization in a relatively narrow age group in early adolescence. We focused on early adolescents given our interest in popularity and classroom norms as moderators, as prioritization of status and conformity to peer norms peak in early adolescence (e.g., Laursen & Veenstra, 2021). Future research should examine whether our findings generalize to younger children or older adolescents who may be less concerned about status. However, we would have expected defending to be the riskiest in adolescence, when "adult-approved" behavior such as defending may be more likely to be punished by peers (which was generally not the case in this study). Thus, we would expect similar results in other age groups.

# CONCLUSIONS

The current findings provide preliminary good news for anti-bullying programs that encourage youth to defend:

in most contexts, defending was not positively associated with victimization concurrently or over time. Our results add to other recent findings that demonstrate no clear longitudinal evidence that defending itself is costly (Malamut, Trach, et al., 2021). Still, this is an important question that should continue to be examined. Even though defending did not predict being victimized over time, it is still important to keep in mind that youth who are both highly victimized and who frequently defend others are more at risk for developing internalizing problems (Malamut, Trach, et al., 2021). Furthermore, it is important for school professionals to pay attention to the popularity norms of their classrooms, because youth who defend in classrooms where popular students bully others were more likely to self-report being (concurrently) victimized. This is crucial, as many intervention efforts aim to increase youth's perceived self-efficacy on defending (i.e., belief that they can successfully defend others; Pöyhönen et al., 2010). If youth who defend also feel more victimized in some situations, then they may be less likely to defend again in the future (perhaps due to decreased defending self-efficacy), regardless of whether or not their peers recognize their victimization. Thus, youth who defend need ongoing peer and adult support if we want them to continue standing up for their bullied peers.

## ORCID

*Sarah T. Malamut* https://orcid.org/0000-0001-5907-2752
*Jessica Trach* https://orcid.org/0000-0002-1604-8813
*Claire F. Garandeau* https://orcid.org/0000-0002-0395-020X
*Christina Salmivalli* https://orcid.org/0000-0002-6253-146X

## REFERENCES

Andreou, E. (2006). Social preference, perceived popularity and social intelligence: Relations to overt and relational aggression. *School Psychology International*, *27*, 339–351. https://doi.org/10.1177/0143034306067286

Andrews, N. C., Hanish, L. D., Updegraff, K. A., Martin, C. L., & Santos, C. E. (2016). Targeted victimization: Exploring linear and curvilinear associations between social network prestige and victimization. *Journal of Youth and Adolescence*, *45*, 1772–1785. https://doi.org/10.1007/s10964-016-0450-1

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models Usinglme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Bouman, T., van der Meulen, M., Goossens, F. A., Olthof, T., Vermande, M. M., & Aleva, E. A. (2012). Peer and self-reports of victimization and bullying: Their differential association with internalizing problems and social adjustment. *Journal of School Psychology*, *50*, 759–774. https://doi.org/10.1016/j.jsp.2012.08.004

Brechwald, W. A., & Prinstein, M. J. (2011). Beyond homophily: A decade of advances in understanding peer influence processes. *Journal of Research on Adolescence*, *21*, 166–179. https://doi.org/10.1111/j.1532-7795.2010.00721.x

Caravita, S. C., & Cillessen, A. H. (2012). Agentic or communal? Associations between interpersonal goals, popularity, and bullying in middle childhood and early adolescence. *Social Development*, *21*, 376–395. https://doi.org/10.1111/j.1467-9507.2011.00632.x

Christina, S., Magson, N. R., Kakar, V., & Rapee, R. M. (2021). The bidirectional relationships between peer victimization and internalizing problems in school-aged children: An updated systematic review and meta-analysis. *Clinical Psychology Review*, *85*, 101979. https://doi.org/10.1016/j.cpr.2021.101979

Cillessen, A. H., & Marks, P. E. (2011). Conceptualizing and measuring popularity. *Popularity in the Peer System*, 25–56.

Cillessen, A. H., & Rose, A. J. (2005). Understanding popularity in the peer system. *Current Directions in Psychological Science*, *14*, 102–105. https://doi.org/10.1111/j.0963-7214.2005.00343.x

Dawes, M., Chen, C. C., Farmer, T. W., & Hamm, J. V. (2017). Self-and peer-identified victims in late childhood: Differences in perceptions of the school ecology. *Journal of Youth and Adolescence*, *46*, 2273–2288. https://doi.org/10.1007/s10964-017-0688-2

Dawes, M., & Malamut, S. (2018). No one is safe: Victimization experiences of high-status youth. *Adolescent Research Review*, *5*, 27–47. https://doi.org/10.1007/s40894-018-0103-6

Dijkstra, J. K., Lindenberg, S., & Veenstra, R. (2008). Beyond the class norm: Bullying behavior of popular adolescents and its relation to peer acceptance and rejection. *Journal of Abnormal Child Psychology*, *36*, 1289–1299. https://doi.org/10.1007/s10802-008-9251-7

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*, 121–138. https://doi.org/10.1037/1082-989X.12.2.121

Garandeau, C. F., Vermande, M. M., Reijntjes, A. H., & Aarts, E. (2022). Classroom bullying norms and peer status: Effects on victim-oriented and bully-oriented defending. *International Journal of Behavioral Development*, *46*, 401–410. https://doi.org/10.1177/0165025419894722

Graham, S., & Juvonen, J. (1998). Self-blame and peer victimization in middle school: An attributional analysis. *Developmental Psychology*, *34*, 587–599. https://doi.org/10.1037/0012-1649.34.3.587

Gromann, P. M., Goossens, F. A., Olthof, T., Pronk, J., & Krabbendam, L. (2013). Self-perception but not peer reputation of bullying victimization is associated with non-clinical psychotic experiences in adolescents. *Psychological Medicine*, *43*, 781–787. https://doi.org/10.1017/S003329171200178X

Huitsing, G., Snijders, T. A., Van Duijn, M. A., & Veenstra, R. (2014). Victims, bullies, and their defenders: A longitudinal study of the coevolution of positive and negative networks. *Development and Psychopathology*, *26*, 645–659. https://doi.org/10.1017/S0954579414000297

Kärnä, A., Voeten, M., Little, T. D., Alanen, E., Poskiparta, E., & Salmivalli, C. (2013). Effectiveness of the KiVa antibullying program: Grades 1–3 and 7–9. *Journal of Educational Psychology*, *105*, 535–551. https://doi.org/10.1037/a0030417

LaFontana, K. M., & Cillessen, A. H. N. (2010). Developmental changes in the priority of perceived status in childhood and adolescence. *Social Development*, *19*, 130–147. https://doi.org/10.1111/j.1467-9507.2008.00522.x

Lambe, L. J., Della Cioppa, V., Hong, I. K., & Craig, W. M. (2019). Standing up to bullying: A social ecological review of peer defending in offline and online contexts. *Aggression and Violent Behavior*, *45*, 51–74. https://doi.org/10.1016/j.avb.2018.05.007

Laursen, B., & Veenstra, R. (2021). Toward understanding the functions of peer influence: A summary and synthesis of recent empirical research. *Journal of Research on Adolescence*, *31*, 889–907. https://doi.org/10.1111/jora.12606

Logis, H., Rodkin, P. C., Gest, S., & Ahn, H.-J. (2013). Popularity as an organizing factor of preadolescent friendship networks: Beyond prosocial and aggressive behavior. *Journal of Research on Adolescence*, *23*, 413–423. https://doi.org/10.1111/jora.12033

Ma, T. L., & Chen, W. T. (2019). The benefits of being defended: Perceived bystander participant roles and victims' emotional and psychosocial adjustment. *Journal of School Violence*, *18*, 77–91. https://doi.org/10.1080/15388220.2017.1387132

Ma, T. L., Meter, D. J., Chen, W. T., & Lee, Y. (2019). Defending behavior of peer victimization in school and cyber context during childhood and adolescence: A meta-analytic review of individual

and peer-relational characteristics. *Psychological Bulletin*, *145*, 891–928. https://doi.org/10.1037/bul0000205

Malamut, S. T., Dawes, M., van den Berg, Y. H. M., Lansu, T. A. M., Schwartz, D., & Cillessen, A. H. N. (2021). Adolescent victim types across the popularity status hierarchy: Differences in internalizing symptoms. *Journal of Youth and Adolescence*, *50*, 2444–2455. https://doi.org/10.1007/s10964-021-01498-w

Malamut, S. T., Luo, T., & Schwartz, D. (2020). Prospective associations between popularity, victimization, and aggression in early adolescence. *Journal of Youth and Adolescence*, *49*, 2347–2357. https://doi.org/10.1007/s10964-020-01248-4

Malamut, S. T., Trach, J., Garandeau, C. F., & Salmivalli, C. (2021). Examining the potential mental health costs of defending victims of bullying: A longitudinal analysis. *Research on Child and Adolescent Psychopathology*, *49*, 1197–1210. https://doi.org/10.1007/s10802-021-00822-z

Meijs, N., Cillessen, A. H., Scholte, R. H., Segers, E., & Spijkerman, R. (2010). Social intelligence and academic achievement as predictors of adolescent popularity. *Journal of Youth and Adolescence*, *39*, 62–72. https://doi.org/10.1007/s10964-008-9373-9

Meter, D. J., & Card, N. A. (2015). Effects of defending: The longitudinal relations among peer-perceived defending of victimized peers, victimization, and liking. *Social Development*, *24*, 734–747. https://doi.org/10.1111/sode.12129

Olweus, D. (1996). *The revised Olweus bully/victim questionnaire*. 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

Parkhurst, J. T., & Hopmeyer, A. (1998). Sociometric popularity and peer-perceived popularity: Two distinct dimensions of peer status. *Journal of Early Adolescence*, *18*, 125–144. https://doi.org/10.1177/0272431698018002001

Peets, K., Pöyhönen, V., Juvonen, J., & Salmivalli, C. (2015). Classroom norms of bullying alter the degree to which children defend in response to their affective empathy and power. *Developmental Psychology*, *51*, 913–920. https://doi.org/10.1037/a0039287

Pouwels, J. L., van Noorden, T. H. J., Lansu, T. A. M., & Cillessen, A. H. N. (2018). The participant roles of bullying in different grades: Prevalence and social status profiles. *Social Development*, *27*, 732–747. https://doi.org/10.1111/sode.12294

Pöyhönen, V., Juvonen, J., & Salmivalli, C. (2010). What does it take to defend the victimized peer? The interplay between personal and social factors. *Merrill-Palmer Quarterly*, *56*, 143–163. https://digitalcommons.wayne.edu/mpq/vol56/iss2/4

Pozzoli, T., Gini, G., & Vieno, A. (2012). The role of individual correlates and class norms in defending and passive bystanding behavior in bullying: A multilevel analysis. *Child Development*, *83*, 1917–1931. https://doi.org/10.1111/j.1467-8624.2012.01831.x

Reijntjes, A., Vermande, M., Olthof, T., Goossens, F. A., Aleva, L., & van der Meulen, M. (2016). Defending victimized peers: Opposing the bully, supporting the victim, or both? *Aggressive Behavior*, *42*, 585–597. https://doi.org/10.1002/ab.21653

Salmivalli, C., Laninga-Wijnen, L., Malamut, S. T., & Garandeau, C. F. (2021). Bullying prevention in adolescence: Solutions and new challenges from the past decade. *Journal of Research on Adolescence*, *31*, 1023–1046. https://doi.org/10.1111/jora.12688

Salmivalli, C., & Voeten, M. (2004). Connections between attitudes, group norms, and behaviour in bullying situations. *International Journal of Behavioral Development*, *28*, 246–258. https://doi.org/10.1080/01650250344000488

Salmivalli, C., Voeten, M., & Poskiparta, E. (2011). Bystanders matter: Associations between defending, reinforcing, and the frequency of bullying in classrooms. *Journal of Clinical Child and Adolescent Psychology*, *40*, 668–676. https://doi.org/10.1080/15374416.2011.597090

Scheffer, J. (2002). Dealing with missing data. *Research Letters in the Information and Mathematical Sciences*, *3*, 153–160.

Schoeler, T., Duncan, L., Cecil, C. M., Ploubidis, G. B., & Pingault, J.-B. (2018). Quasi-experimental evidence on short- and long-term consequences of bullying victimization: A meta-analysis. *Psychological Bulletin*, *144*, 1229–1246. https://doi.org/10.1037/bul0000171

Scholte, R. H., Burk, W. J., & Overbeek, G. (2013). Divergence in self-and peer-reported victimization and its association to concurrent and prospective adjustment. *Journal of Youth and Adolescence*, *42*, 1789–1800. https://doi.org/10.1007/s10964-012-9896-y

Sentse, M., Scholte, R., Salmivalli, C., & Voeten, M. (2007). Person–Group dissimilarity in involvement in bullying and its relation with social status. *Journal of Abnormal Child Psychology*, *35*, 1009–1019. https://doi.org/10.1007/s10802-007-9150-3

Spadafora, N., Marini, Z. A., & Volk, A. A. (2020). Should I defend or should I go? An adaptive, qualitative examination of the personal costs and benefits associated with bullying intervention. *Canadian Journal of School Psychology*, *35*, 23–40. https://doi.org/10.1177/0829573518793752

Strindberg, J., Horton, P., & Thornberg, R. (2020). The fear of being singled out: Pupils' perspectives on victimisation and bystanding in bullying situations. *British Journal of Sociology of Education*, *41*, 942–957. https://doi.org/10.1080/01425692.2020.1789846

Thornberg, R., Tenenbaum, L., Varjas, K., Meyers, J., Jungert, T., & Vanegas, G. (2012). Bystander motivation in bullying incidents: To intervene or not to intervene? *The Western Journal of Emergency Medicine*, *13*, 247–252. https://doi.org/10.5811/westjem.2012.3.11792

Trach, J., Hymel, S., Waterhouse, T., & Neale, K. (2010). Bystander Responses to School Bullying: A Cross-Sectional Investigation of Grade and Sex Differences. *Canadian Journal of School Psychology*, *25*, 114–130. https://doi.org/10.1177/0829573509357553.

van der Ploeg, R., Kretschmer, T., Salmivalli, C., & Veenstra, R. (2017). Defending victims: What does it take to intervene in bullying and how is it rewarded by peers? *Journal of School Psychology*, *65*, 1–10. https://doi.org/10.1016/j.jsp.2017.06.002

Wei, H. S., & Chen, J. K. (2008). Social withdrawal, peer rejection, and peer victimization in Taiwanese middle school students. *Journal of School Violence*, *8*, 18–28. https://doi.org/10.1080/15388220802067755

Williford, A., Boulton, A., Noland, B., Little, T. D., Kärnä, A., & Salmivalli, C. (2012). Effects of the KiVa anti-bullying program on adolescents' depression, anxiety, and perception of peers. *Journal of Abnormal Child Psychology*, *40*, 289–300. https://doi.org/10.1007/s10802-011-9562-y

Wright, J. C., Giammarino, M., & Parad, H. W. (1986). Social status in small groups: Individual–group similarity and the social "misfit". *Journal of Personality and Social Psychology*, *50*, 523–536. https://doi.org/10.1037/0022-3514.50.3.523

**How to cite this article:** Malamut, S. T., Trach, J., Garandeau, C. F., & Salmivalli, C. (2022). Does defending victimized peers put youth at risk of being victimized? *Child Development*, *00*, 1–15. https://doi.org/10.1111/cdev.13866