UvA-DARE (Digital Academic Repository)

Defending victims of bullying in the classroom: The role of moral responsibility and social costs

Pouwels, J.L.; van Noorden, T.H.J.; Caravita, S.C.S.

DOI

[10.1016/j.jesp.2019.103831](https://doi.org/10.1016/j.jesp.2019.103831)

Publication date

2019

Document Version

Final published version

Published in

Journal of Experimental Social Psychology

[Link to publication](#)

Citation for published version (APA):

Pouwels, J. L., van Noorden, T. H. J., & Caravita, S. C. S. (2019). Defending victims of bullying in the classroom: The role of moral responsibility and social costs. *Journal of Experimental Social Psychology*, 84, [103831]. <https://doi.org/10.1016/j.jesp.2019.103831>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)

Defending victims of bullying in the classroom: The role of moral responsibility and social costs

J. Loes Pouwels^{a,*,1}, Tirza H.J. van Noorden^a, Simona C.S. Caravita^b

^a Behavioural Science Institute, Radboud University, Montessorilaan 3, 6525 HR Nijmegen, the Netherlands

^b CRldee, Department of Psychology, Catholic University of the Sacred Heart (Milan & Brescia), L.go Gemelli 1, 20123 Milan, Italy

ARTICLE INFO

Keywords:

Moral courage
Defending
Bullying
Prioritizing popularity
Moral disengagement
Popularity

ABSTRACT

The current study examined the role of moral responsibility (i.e., moral disengagement) and social costs (i.e., prioritizing popularity and classroom norms) in defending victims of bullying in the classroom. Participants were 1362 students (Age 8–15) from 58 classrooms who completed self-reported measures on moral disengagement and prioritizing popularity and peer-reported measures on bullying and defending. We first examined whether the outcome of the cost-benefit analysis with regard to the social costs of defending (i.e., prioritizing popularity) only related to defending among students who did not justify bullying (i.e., low levels of moral disengagement). This hypothesis was supported among boys; multi-level analyses showed that prioritizing popularity was positively associated with defending, but only among boys with low levels of moral disengagement. We also investigated contextual effects by examining how differences in the potential negative social costs of bullying as measured by classroom norms of bullying were related to defending. Girls were more likely to defend in classrooms in which bullying was relatively common than in classrooms in which bullying was relatively uncommon. Boys and girls were more likely to defend in classrooms in which bullies were relatively popular than in classrooms in which bullies were relatively unpopular. We further tested interactions between individual and contextual characteristics. Only among boys, the effect of classroom prestige norms on defending was stronger among boys with lower levels of prioritizing popularity. These findings highlight the importance of considering interactions between moral responsibility and social costs at the individual and classroom level when trying to understand acts of moral courage, such as defending.

1. Introduction

Intervening in situations in which someone else harms others or behaves immorally in other ways, such as discrimination, aggression, mobbing, or illegal business practices requires moral courage as it potentially comes at high social costs (Osswald, Frey, & Streicher, 2012). Moral courage is a specific type of prosocial behavior, which can be defined as acting upon personal moral views and values by standing up against other people despite the anticipation of substantial negative social consequences (Greitemeyer, Osswald, Fischer, & Frey, 2007; Niesta, Greitemeyer, Fischer, & Frey, 2010). Lopez, O'Byrne, and Petersen (2003) argued that there often is a power imbalance, as morally courageous individuals tend to stand up against others who are more powerful than themselves. As most studies focused on moral courage among late adolescents and adults (see e.g., Baumert, Halmburger, & Schmitt, 2013; Galdi, Maass, & Cadinu, 2017; Osswald,

Greitemeyer, Fischer, & Frey, 2010), relatively little is known about the process of moral courage in childhood and early adolescence. However, as moral reasoning develops during this age period (Malti & Ongley, 2013), we aimed to examine to which extent an adult model of moral courage is applicable to childhood and early adolescence (Bronstein, Fox, Kamon, & Knolls, 2007). We therefore presented and provided empirical support for an adjusted seminal model of peer-reported defending as an act of perceived moral courage by peers.

1.1. Perceived moral courage in the classroom: defending victims of bullying

Bullying can be defined as a moral transgression that is a subtype of aggression, in which one or more individuals (the bullies) repeatedly attack, humiliate, or exclude other persons (the victims) who have difficulties to defend themselves (Salmivalli, 2010). Students can defend victims of bullying by trying to stop the bullying or by comforting

* Corresponding author.

E-mail addresses: j.l.pouwels@uva.nl (J.L. Pouwels), t.vannoorden@psych.ru.nl (T.H.J. van Noorden), simona.caravita@unicatt.it (S.C.S. Caravita).

¹ Present address: Amsterdam School of Communication Research, University of Amsterdam, PO Box 15,791, Amsterdam, 1001 NG, the Netherlands.

the victim (Salmivalli, Lagerspetz, Björkqvist, Österman, & Kaukiainen, 1996). Thornberg et al. (2012) concluded that most instances of defending are morally motivated as students reported that the belief that bullying is morally wrong is one of the main reasons for defending. In addition, various researchers argued that defending may be related to negative social consequences as they showed that defenders need to stand up in public against bullies who have been found to have the most powerful and popular position in the peer group (Juvonen & Galván, 2008; Pouwels, van Noorden, Lansu, & Cillessen, 2018). It has been proposed that bullies may use their peer status to punish defenders who hinder their bullying (Juvonen & Galván, 2008) as it has been found that defenders become less liked by their peers over time (Meter & Card, 2015). Because of the anticipated costs of defending, several researchers proposed that some students may be less willing to intervene because they may be afraid to become the next target of bullying or to lose peer status (Pöyhönen, Juvonen, & Salmivalli, 2010; Pronk, Goossens, Olthof, de Mey, & Willemen, 2013; Salmivalli, 2010; Spadafora, Marini, & Volk, 2018).

As standing up against bullies who are often more powerful is morally motivated and associated with potential negative social consequences, defending is a classic example of moral courage in the classroom. Likewise, researchers from different fields argued that children need courage to defend their peers who are bullied and unpopular (Bronstein et al., 2007; Peets, Poyhonen, Juvonen, & Salmivalli, 2015; Pöyhönen, Juvonen, & Salmivalli, 2010; Staub, 2003).

Students vary in the degree to which they are inclined to defend their peers and this variability can be explained by individual and contextual characteristics. The seminal model of Latané and Darley (1970) describes five stages of psychological processes that underlie helping behavior. Baumert et al. (2013) applied this model to moral courage and we further adapted and applied the model to defend as an act of perceived moral courage by peers (see Fig. 1). Our model consists of three propositions. Proposition 1 describes that, in order to display defending behavior as act of moral courage, one needs to go through the following sequential stages: a bullying situation has to be (1) witnessed and (2) interpreted as a moral violation and emergency, the witnesses must perceive themselves as (3) personally and morally responsible and (4) self-competent to intervene, (5) and in a cost-benefit analysis, the subjective costs of non-intervention need to be perceived as higher than costs of intervention. In the present study, we aimed to provide initial empirical support for the adapted seminal model. As it was beyond the scope of the present study to test all stages, we focused on stages 3 and 5 of the model. These stages were selected because they are most closely related to the moral motivation to defend and its potential costs, which are the key aspects that distinguish moral courage from helping. Previous research has acknowledged the sequential nature of the model, but did not test it empirically (Baumert et al., 2013). In line with the sequential idea of the model, we examined whether sensitivity to the outcome of the cost-benefit analysis (stage 5) would only be relevant for those students who tend to hold themselves personally and morally responsible for intervention (successfully accomplished stage 3).

Previous versions of the model mainly focused on individual characteristics and did not acknowledge the role of the social context. Proposition 2 argues that the outcomes of each stage are influenced by both individual (i.e., moral disengagement - stage 3; prioritizing popularity - stage 5) and contextual characteristics (i.e., classroom norms - stage 5). Proposition 3 even goes a step further by arguing that the individual and contextual characteristics at stage 5 interact with each other. We aimed to provide evidence for these propositions by examining contextual predictors of peer-reported defending in interaction with individual characteristics.

1.2. Associations of individual characteristics with defending

With regard to the justification of personal responsibility (stage 3),

previous research revealed that individuals who retrospectively reported about a situation in which they displayed moral courage felt themselves more personally responsible to intervene in this situation than individuals who reported about a situation in which they did not display moral courage (Greitemeyer, Fischer, Kastenmüller, & Frey, 2006). This is closely related to the concept of moral disengagement; the extent to which people disengage from their moral values, such as by blaming the victim or displacing responsibility (Bandura, Barbaranelli, & Caprara, 1996). With regard to bullying, some students justify bullying through mechanisms of moral disengagement. Through mechanisms of disengagement, students no longer experience self-sanctions for acting against their moral values, which enact them to behave in immoral ways. Ample of previous studies have shown that students who justify bullying by mechanisms of moral disengagement are more likely to display aggressive and pro bullying behavior (Caravita, Gini, & Pozzoli, 2012; Caravita, Sijtsema, Rambaran, & Gini, 2014; Mazzone, Yanagida, Caravita, & Strohmeier, 2018; van Noorden, Haselager, Cillessen, & Bukowski, 2014) and less likely to defend their peers than students who do not justify bullying (Gini, Pozzoli, & Bussey, 2015; Sijtsema, Rambaran, Caravita, & Gini, 2014; Thornberg & Jungert, 2014). We aimed to replicate the finding that students high in moral disengagement, as compared to those low in moral disengagement, are less inclined to defend their peers.

With regard to the cost-benefit analysis (stage 5), students' prioritizing of popularity may be related to the extent to which they perceive defending as costly social behavior. Researchers have argued that children need courage to defend their peers because they might lose peer status by defending (Pöyhönen et al., 2010; Pronk et al., 2013; Salmivalli, 2010). Therefore, some students may withhold themselves from defending as they perceive that the costs of defending (i.e., losing peer status) are higher than the costs of non-intervention. However, not all students may be concerned with this cost to the same degree. Although striving for social status in the peer group is an important goal for many adolescents, adolescents vary in the degree to which they prioritize being popular over other goals, such as behaving in prosocial ways and maintaining friendships (LaFontana & Cillessen, 2010). Prioritizing popularity has been found to be negatively associated with prosocial behavior (Cillessen, Mayeux, Ha, & De Bruyn, 2014; Li & Wright, 2014; van den Broek, Deutz, Schoneveld, Burk, & Cillessen, 2016) and defending in particular (Duffy, Penn, Nesdale, & Zimmer-Gembeck, 2017). We therefore aimed to replicate the finding that prioritizing popularity is negatively related to defending in bullying situations.

1.3. Associations of contextual characteristics with defending

In addition to individual characteristics, we proposed that contextual characteristics may also alter the degree to which students defend their peers (Proposition 2). With regard to the cost-benefit analysis (stage 5), we argue that the costs of defending may vary between classroom, as they depend on classroom norms. Two types of classroom norms of bullying may alter the degree to which defending may be associated with negative social consequences: descriptive norms of bullying and social prestige norms of bullying.

Descriptive norms of bullying can be defined as the average level of bullying in a classroom (Cialdini, 2007; Cialdini, Reno, & Kallgren, 1990; Dijkstra & Gest, 2015). The more common bullying is in a classroom, the higher the frequency of bullying and the more pervasive the bullying incidents are to which students are exposed to. Pozzoli, Gini, and Vieno (2012) argued that adolescents conform to norms because they have the desire to be accepted by their peers and do not want to run the risk to lose status. Students who defend their peers in a context in which bullying is relatively common have been found to behave against the group norm and experience higher levels of psychosocial problems than students who defend in a context in which bullying is less common (Lambe, Hudson, Craig, & Pepler, 2017).

Proposition 1: Defending behavior as act of moral courage depends on 5 subsequential stages

Proposition 2: The outcome of each state depends on both individual and contextual characteristics

Proposition 3: Individual and contextual characteristics interact with each other

— H1 & H2 —> - - - H3 - - -> — H4 & H5 —> H6	Test Proposition 2: Replication individual effects found in previous research Test Proposition 1: The outcomes of stage 5 will only be relevant for those who accomplished stage 3 Test Proposition 2: Replication contextual effects found in previous research Test Proposition 3: Interaction individual and contextual characteristics

Fig. 1. Seminal model of defending behavior as act of moral courage.

Note: H1 to H6 refer to hypothesis 1 to hypothesis 6.

Adapted from Latané and Darley (1970)

Pozzoli et al. (2012) argued that the high costs of defending in classrooms in which bullying is relatively common may withhold students from defending. In line with this idea, it has been found that the more common bullying is in a classroom or peer group (i.e., high descriptive norms of bullying) the less likely students are to defend or intervene (Espelage, Green, & Polanin, 2012; Peets et al., 2015).

In addition to descriptive norms, classroom social prestige norms of bullying also impact the degree to which defending may be related to negative social outcomes. The level of bullying of popular students in a classroom can be defined as social prestige norms of bullying (Dijkstra, Lindenberg, & Veenstra, 2008). Dijkstra et al. (2008) argued that students vary in the extent to which they have impact on the establishment of classroom social norms of bullying. As popular adolescents have been found to be socially dominant and the most powerful members of a peer group (LaFontana & Cillessen, 1998), they appear to have the strongest impact on the acceptance of social behavior in a group (Dijkstra & Gest, 2015). Dijkstra et al. (2008) therefore argued that adolescents especially tend to imitate their popular peers as this may help them to increase in status themselves. Likewise, the level of bullying among popular students had a stronger impact on social acceptance of bullying than the general level of bullying in a classroom (Dijkstra & Gest, 2015). With regard to defending, it can be argued that behaving against the bullying norm that is set by popular adolescents may be a risk in terms of peer status that students do not want to take. Accordingly, Peets et al. (2015) found that the more popular bullies were in a classroom, the less likely students were to defend. The present study aimed to replicate the associations of descriptive and prestige norms with peer-reported defending that have been found by Peets et al. (2015).

1.4. Interactions between contextual and individual characteristics associated with defending

The third proposition of the seminal model of defending as act of moral courage is that contextual characteristics are likely to interact with individual characteristics. With regard to classroom norms of defending, it has been argued that not all adolescents conform themselves to classroom norms of social behaviors. Previous studies have found that popular adolescents are more likely to conform themselves to social norms related to bullying than unpopular students (Lucas-Molina, Giménez-Dasí, Fonseca-Pedrero, & Pérez-Albéniz, 2018; Peets et al., 2015; Yun & Graham, 2018). Whereas these findings are often explained by the idea that popular adolescents are concerned with a potential loss of social status if they do not conform to social norms (Dijkstra et al., 2008; Peets et al., 2015; Yun & Graham, 2018), not all popular students have been found to strive for popularity (Cillessen et al., 2014). Juvonen and Galván (2008) argued that the effects of group norms on behavior may be strongest among those adolescents who strive for peer status. However, as previous research on classroom norms examined students' actual level of popularity rather than how much they strive for popularity, we do not know yet whether the associations of classroom norms of bullying with defending are stronger among students with relatively high levels of prioritizing popularity than among students with low levels of prioritizing popular. This study therefore extended the findings of previous research by testing the interaction between prioritizing popularity and classroom norms.

1.5. The role of gender

The individual and contextual characteristics that impact stages 3 (moral responsibility) and stage 5 (social costs) of the seminal model of defending may be different for boys than for girls. We therefore tested separate models for boys and girls, in order to explore gender differences. Especially the cost-benefit analysis may be different for boys than girls. Defending has been found to be more common and normative among girls than boys (Pouwels et al., 2018; Salmivalli,

Lappalainen, & Lagerspetz, 1998). This may be due to fact that people have gender-specific norms for prosocial behavior. As compared to boys, girls are more often socialized in ways that promote prosocial and empathic behavior, such as defending. Accordingly, prosocial behavior, such as defending, may be better accepted among girls than boys, whereas the costs of defending may be higher for boys (Eagly & Wood, 1991; Hoffman, 1977; Ickes, Gesn, & Graham, 2000). This suggests that boys may be more concerned with factors influencing the potential costs of defending than girls (stage 3), as defending may be more harmful for them. Individual characteristics impacting the cost benefit analysis, such as prioritizing popularity and classroom norms of bullying, may therefore be more strongly related to defending among boys than girls. Accordingly, the association between prioritizing popularity and prosocial behavior (Cillessen et al., 2014; Li & Wright, 2014; van den Broek et al., 2016) and defending in particular (Duffy et al., 2017) has been found to be stronger among boys than girls. Less is known about gender differences in the association of classroom norms with defending, but like prioritizing popularity, these norms may be more strongly associated with defending among boys than girls. We therefore aimed to examine whether prioritizing popularity and classroom norms of bullying are significantly associated with defending among boys, but not significant or relatively weakly associated with defending among girls.

With regard to the association between moral disengagement and defending, the role of gender is less clear. Although most studies did not examine whether the association of moral disengagement with defending was moderated by gender, one study showed that in late-childhood moral disengagement was negatively associated with defending only among girls (Caravita et al., 2012). However, this gender difference was not found in early adolescence. We therefore examined the association between moral disengagement and defending separately for boys and girls.

1.6. The present study

The present study aimed to empirically test stages 3 and 5 of the adjusted seminal model by Latané and Darley (1970) of defending as an act of perceived moral courage by peers (see Fig. 1). We did this by investigating how peer-reported levels of defending were associated with characteristics at the individual and classroom level that were assumed to affect stage 3 (holding one selves morally responsible) and 5 (cost-benefit analysis) of the seminal model of defending.

At the individual level, according with previous studies, we expected that the more students justify bullying through mechanisms of moral disengagement, the less likely it would be that they will defend their peers (hypothesis 1, individual level, stage 3) (Gini et al., 2015; Sijtsma et al., 2014; Thornberg & Jungert, 2014). We also expected that the more students prioritize popularity over other social priorities, the less likely it would be that they will defend their peers (hypothesis 2, individual level, stage 5) (Cillessen et al., 2014; Duffy et al., 2017; Li & Wright, 2014; van den Broek et al., 2016).

Second, in line with the sequential idea of the model (Proposition 1), we extended previous research at the individual level, by examining whether the outcome of the cost-benefit analysis (stage 5) would only be relevant among those adolescents with low levels of moral disengagement (students who are more likely to have accomplished stage 3, hypothesis 3). Specifically, we expected that students who tend to justify bullying by mechanisms of moral disengagement would be relatively unlikely to defend their peers, regardless of their priority of popularity. With regard to students who do not tend to employ mechanisms of moral disengagement, we expected that their level of defending would be related to the outcomes of the cost benefit analysis (prioritizing popularity).

Third, we tested whether contextual characteristics also altered the degree to which students defend their peers (Proposition 2). We aimed to replicate the finding that students were relatively unlikely to defend

in classrooms in which bullying was relatively common (i.e., high descriptive norms, hypothesis 4, classroom level, stage 5) and in classrooms in which bullies were relatively popular (i.e., high social prestige norms, hypothesis 5, classroom level, stage 5) (Peets et al., 2015).

Fourth, we extended the findings of previous research by examining the interaction between personal and contextual characteristics (Proposition 3, hypothesis 6, stage 5). We expected that students who do prioritize popularity would only be likely to defend when the descriptive and social prestige norms of bullying are low (i.e., low potential costs of defending). In contrast, we expected that students who do not prioritize popularity would always defend their peers, regardless of the classroom norms, as they may not be concerned with the risk associated with the bullying norms within their classroom.

Finally, separate models were tested for boys and girls, as the associations of prioritizing popularity (Duffy et al., 2017) and classroom norms of bullying may be stronger among boys than girls. Specifically, we expected that the associations of prioritizing popularity and classroom norms with defending would be significant in the model for boys, but not significant or relatively weak in the model for girls, as the costs of defending may be higher for boys than girls (Pouwels et al., 2018; Salmivalli et al., 1998).

2. Method

2.1. Participants and procedure

The potential sample consisted of 896 students in 36 classrooms in fifteen primary schools (Grades 3 to 6) and 545 students in 23 classrooms in two secondary schools (Grades 7 and 8). This sample was recruited by sending a letter to schools in the Eastern part of the Netherlands explaining the study and asking teachers to participate. All schools that were willing to participate were included in the study. After the school principals gave permission for participation, parents received a letter explaining the study. They were asked to return the letter if they did not want their child to participate in the study. One parent did not want to let their child participate in the study. Assent was obtained from all students. One primary school classroom was excluded as the peer nominations procedure was not reliable in this classroom because 46% of the students were absent during the data collection (Marks, Babcock, Cillessen, & Crick, 2013). We also excluded one participant who did not speak Dutch and one participant who very recently switched classrooms. In the remaining sample, 52 students were absent during data collection and five students were excluded due to missing data. Therefore, the final sample consisted of 1362 students ranging in age from eight to fifteen years (52% girls, M age = 11.85 years, SD = 1.49) from 58 classrooms (M classroom size = 23.48 students, SD = 3.87). The sample was predominantly Caucasian: 96% of the children were born in the Netherlands. Secondary school students were in different educational tracks: Prevocational track (VMBO, 33%), prevocational and intermediate general secondary education (VMBO-T/HAVO, 20%) and intermediate general secondary education and college preparatory education (HAVO/VWO, 46%).

The study was approved by the institutional review board. Data were collected during a one-hour classroom session that took place between 2013 and 2016. One part of the primary school students completed the measures on paper. The other part of the primary school students and all secondary school students participated in a computerized peer nomination assessment. Each student was provided with a mini-laptop computer on which they filled in the questionnaire in their own classroom. For each peer-nomination question, the classroom was the reference group and students could nominate both same-gender and other-gender peers. They were allowed to nominate an unlimited number of peers. Students were ensured that their answers to questions were confidential and processed in an anonymous way. More details on the peer nomination procedure can be found in van den Berg and

Cillessen (2012). Part of the data on bullying, defending and popularity has been published before in Pouwels et al. (2018). However, Pouwels et al. (2018) covered a different topic as they did not examine any classroom level effects and did not predict defending from students' individual characteristics (i.e., moral disengagement and prioritizing popularity).

2.2. Measures

In addition to the measures that were used in the current study and that are listed below, additional measures were administered in certain classrooms. We collected peer-nominations on friendship, social preference, and socio-emotional functioning, and self-reported measures on students' socio-emotional well-being. A full list of measures that were used in the study can be found in Online Supplement S2.

2.2.1. Self-reported variables

2.2.1.1. Prioritizing popularity

Prioritizing popularity was measured with a Dutch translation of the prioritizing popularity scale (LaFontana & Cillessen, 2010). Students completed twenty items belonging to ten vignettes that presented them a dilemma. Each dilemma had two possible actions: one action consisted of behavior that maintained or increased popularity, while the other action benefited one out of five other social domains: maintaining a same-sex friendship, conforming to norms for behavior, achieving personal athletic or academic success, showing compassion for a rejected peer, or pursuing romantic relationships. There were two vignettes per domain. An example vignette is: "Imagine that you are invited to a party. All the popular boys/girls will be there. You ask if you can bring your best friend, but you are told that your friend is not welcome to come. You really want to go, but you know your friend wants to go too." The two possible actions belonging to this vignette were "How likely are you to tell them that you can't go because your friend can't come?" and "How likely are you to go to the party anyway without your friend?". Students rated how likely it was that they would choose each action on a 6-point scale, ranging from (1) "definitely not" to (6) "definitely". After recoding and averaging, higher scores indicated that students were more likely to prioritize popularity in general over the other priorities. Cronbach's alpha was 0.83.

2.2.1.2. Moral disengagement

Moral disengagement regarding bullying was assessed with the Dutch translation (van Noorden et al., 2014) of the 15 item version of the moral disengagement scale (Almeida, Correia, & Marinho, 2010; Hymel, Rocke-Henderson, & Bonanno, 2005). Students rated the items (e.g., "Some kids get bullied because they deserve it") on a 4-point scale (1 = completely disagree, 4 = completely agree). A mean score was calculated so that higher scores reflected higher levels of moral disengagement. Cronbach's alpha was 0.78.

2.2.2. Peer-reported variables

2.2.2.1. Popularity

Popularity was measured by asking students to nominate classmates who were popular ("Who are most popular?") and unpopular ("Who are least popular?"). For each question, students had to nominate at least one classmate. For each student, proportion scores were calculated per item by dividing the total number of nominations received per item by the number of nominators in a classroom. Subsequently, a score for popularity was computed by subtracting the proportion of nominations received for "least popular" from the proportion of nominations received for "most popular" (Cillessen, 2009; Parkhurst & Hopmeyer, 1998).

Fig. 2. Distribution of classroom social prestige norms across classrooms.

2.2.2.2. Bullying and defending

Bullying and defending were measured with the bullying and defending subscale of the Dutch version (Pouwels, Lansu, & Cillessen, 2016) of the shortened participant role questionnaire (Kärnä et al., 2013; Salmivalli & Voeten, 2004). Each scale consisted of three questions (e.g., “Who starts bullying?”; “Who tells to others to stop bullying?”). For all items, students could also nominate no one. For each student, proportion scores were calculated per item by dividing the total number of nominations received per item by the number of nominators in a classroom. Subsequently, we computed a scale score for bullying and defending by averaging the proportion scores of the bullying and defending items, respectively. Cronbach's alpha was 0.84 for the bullying scale and 0.94 for the defending scale.

2.2.2.3. Classroom norms of bullying

Classroom descriptive norms of bullying were computed by averaging the individual proportion scores of bullying of all students in the classroom (Dijkstra & Gest, 2015; Peets et al., 2015). Classroom social prestige norms of bullying were determined by calculating the within classroom correlation between proportion scores of bullying and proportion scores of popularity (Dijkstra & Gest, 2015; Peets et al., 2015).

In classroom with high levels of social prestige norms of bullying, there is a strong positive correlation between bullying and popularity, suggesting that bullies are relatively popular in these classrooms (as compared to their peers). In classrooms with low social prestige norms, bullying is unrelated or even negatively related to popularity, suggesting that bullies are not particularly popular or even relatively unpopular (Dijkstra & Gest, 2015). Correlations ranged from -0.51 to 0.84 (see Fig. 2).

2.3. Analytic approach

All data-analyses were conducted after all data were collected. More details on the power of the study can be found in the Online Supplement S1. We conducted multilevel analyses in order to predict students' defending from prioritizing popularity, moral disengagement and classroom norms of bullying. A multi-level approach was used to control for the fact that students (Level 1) were nested within classrooms (Level 2). We estimated separate models for boys and girls. The

models were estimated using the lmer function of the lme4 package (Bates, Maechler, Bolker, & Walker, 2017) in R (R Core Team, 2016). Models were estimated by maximum likelihood estimation and p -values were obtained using the Satterthwaite's degrees of freedom method of the package lmerTest (Kuznetsova, Brockhoff, & Bojesen, 2018). For both boys and girls, we used a model building approach, as recommended by Hox, Van De Schoot, and Moerbeek (2018). We first compared different models against each other and tested whether the complex model had a significantly better fit than the simplified model, using the likelihood ratio test. Only if the model had a significant better fit, we interpreted the fixed effects.

We first examined the relative amount of within and between person variance in defending (Model 1). Subsequently, we examined main effects at the individual level (Model 2). We controlled for students' individual level of bullying and popularity, because previous research found that the lower students' levels of bullying and the higher their levels popularity, the more likely they were to defend (Peets et al., 2015). We also added moral disengagement (hypothesis 2) and prioritizing popularity (hypothesis 3) which were the substantive predictors (proposition 2, individual level). Next, we added the interaction between prioritizing popularity and moral disengagement in line with the sequential idea of the seminal model (proposition 1, Model 3, hypothesis 3). All individual level variables were centered around their group mean according to the guidelines by Enders and Tofghi (2007). Subsequently, classroom level effects were examined (Model 4). We controlled for grade and classroom size, as the level of defending depends on grade (Pouwels et al., 2018) and defending is less common in larger classrooms (Peets et al., 2015). Grade was dummy coded with primary school as the comparison group. The main variables of interest at the second level were descriptive norms (proposition 2, classroom level, hypothesis 4) and social prestige norms of bullying in the classroom (proposition 2, classroom level, hypothesis 5). All classroom level variables were centered around their grand mean. Finally, we entered a random slope for prioritizing popularity and two-way cross-level interactions of descriptive norms and social prestige norms, respectively, with prioritizing popularity (Model 5, proposition 3, hypothesis 6).

Table 1Means, standard deviations, and correlations between student-level variables for boys ($N = 654$) and girls ($N = 708$).

	1	2	3	4	5	<i>M</i> (girls)	<i>SD</i> (girls)
1. Prioritizing popularity	–	0.27***	0.22***	0.17***	–0.06	2.69	0.72
2. Moral disengagement	0.33***	–	0.05	0.18***	–0.02	1.46	0.34
3. Popularity	0.19***	0.02	–	0.30***	0.14***	–0.02	0.34
4. Bullying	0.25***	0.15***	0.32***	–	–0.16***	0.03	0.06
5. Defending	–0.22***	–0.10*	0.01	–0.25***	–	0.22	0.14
<i>M</i> (boys)	2.86	1.57	0.05	0.07	0.13		
<i>SD</i> (boys)	0.75	0.40	0.34	0.12	0.12		

Note. Means, standard deviations and correlations for girls are presented above the diagonal. Means, standard deviations and correlations for boys are presented below the diagonal.

* $p < .05$.

*** $p < .001$.

3. Results

3.1. Descriptive statistics and correlations

Table 1 presents the descriptive statistics and correlations between the student-level variables above the diagonal for girls, and below the diagonal for boys. Gender differences between the means were determined using independent t -tests. Boys scored significantly higher than girls on prioritizing popularity, $t(1360) = 4.40$, $p < .001$, moral disengagement, $t(1284) = 5.15$, $p < .001$, popularity $t(1360) = 3.52$, $p < .001$, and bullying, $t(926.71) = 8.02$, $p < .001$. In contrast, girls scored significantly higher than boys on defending, $t(1345.38) = 13.13$, $p < .001$.

The pattern of correlations also seemed to differ between boys and girls (see Table 1). These gender differences were examined using the r -to- z transformations for independent correlations. Among boys, defending was significantly negatively related to prioritizing popularity and moral disengagement, whereas these correlations were not significant among girls. However, only the negative correlation between prioritizing popularity and defending was significantly stronger for boys than girls, $z = -3.01$, $p = .002$. In contrast, defending was significantly positively related to popularity, but only among girls and not among boys, $z = 2.41$, $p = .016$. Defending was negatively related to bullying and the strength of this correlation did not differ between boys and girls. Correlations between prioritizing popularity, moral disengagement, popularity and bullying did not differ between boys and girls. To summarize, among boys, defending was correlated with lower levels of prioritizing popularity, moral disengagement and bullying. Among girls, defending was correlated with higher levels of popularity and lower levels of bullying.

Descriptive statistics and correlations between the classroom-level variables, including the classroom level of defending, are presented in Table 2. Grade was negatively correlated with classroom descriptive norms of bullying, indicating that the average level of bullying in classrooms was higher in primary schools than in secondary schools. Defending was not significantly related to grade and classroom descriptive norms of bullying, whereas significant negative associations of defending with classrooms size and social prestige norms of bullying

were found. These associations indicate that the average level of defending in a classroom was higher in smaller classrooms and in classrooms in which bullies were less popular.

3.2. Predicting defending from individual and contextual characteristics

As correlations confound individual differences and classroom dependent differences, all final conclusions are based on the multi-level analyses and not on the bivariate correlations. Tables 3 and 4 present the results of the multi-level models for boys and girls, respectively. We started the analyses by testing an intercept-only model to estimate the relative amount of variance in defending at the student and classroom levels (Model 1). The intraclass correlation was 0.71 among boys and 0.63 among girls. This indicates that 71% and 63% of the variance was due to differences between classrooms among boys and girls, respectively. In contrast 29% and 27% of the variance was due to individual differences between students.

Next, variables at the individual level were added to the model (Model 2). The likelihood ratio test showed that the fit of Model 2 was significantly better than the fit of Model 1. Among both boys and girls defending was positively associated with popularity and negatively associated with bullying and prioritizing popularity (hypothesis 2). Only among boys, defending was also negatively associated with moral disengagement (hypothesis 1).

Subsequently, in line with the sequential idea of the model (proposition 1) we tested the interaction between moral disengagement and prioritizing popularity (Model 3, hypothesis 3). The likelihood ratio test showed that Model 3 had a significantly better fit than Model 2 among boys but not among girls. Likewise, a significant interaction between prioritizing popularity and moral disengagement was found among boys. In order to interpret this interaction, we determined the significance of the effect of prioritizing popularity on defending among boys with low (mean – 1 SD) and high values of moral disengagement (mean + 1 SD) using the web utility of Preacher, Curran, and Bauer (2006). The interaction is plotted in Fig. 3a. Prioritizing popularity was significantly positively related to levels of defending among adolescents with low levels of moral disengagement, $t(648) = -5.43$, $p < .001$, but not significantly related to defending among adolescents with high

Table 2Means, standard deviations, and correlations between classroom-level variables ($N = 58$).

	1	2	3	4	5	<i>M</i>	<i>SD</i>
1. Grade	–					0.40	0.49
2. Classroom size	–0.16	–				24.52	4.16
3. Classroom descriptive norms of bullying	–0.51***	–0.21	–			0.05	0.03
4. Classroom social prestige norms of bullying	0.15	0.18	0.09	–		0.37	0.29
5. Defending (classroom level)	0.14	–0.38**	0.06	–0.34**	–	0.19	0.11

Note. Grade was dummy-coded (0 = primary school; 1 = secondary school).

** $p < .01$.

*** $p < .001$.

Table 3

Multilevel model predicting defending among boys from individual social-cognitive characteristics and classroom bullying norms.

	Defending among boys									
	Model 1 Empty model		Model 2 P2: individual Level		Model 3 P1: sequential nature		Model 4 P2: contextual level		Model 5 P3: individual * contextual level	
	<i>b</i>	<i>SE</i>	<i>b</i>	<i>SE</i>	<i>b</i>	<i>SE</i>	<i>b</i>	<i>SE</i>	<i>b</i>	<i>SE</i>
Individual level (1)										
Intercept	0.134***	0.014	0.134***	0.014	0.132***	0.014	0.126***	0.019	0.127***	0.018
Popularity			0.045***	0.007	0.045***	0.007	0.045***	0.007	0.043***	0.007
Bullying			−0.249***	0.021	−0.249***	0.021	−0.249***	0.021	−0.243***	0.021
Prioritizing popularity (PP)			−0.016***	0.003	−0.016**	0.003	−0.016***	0.003	−0.016***	0.004
Moral disengagement (MD)			−0.014*	0.006	−0.021***	0.007	−0.021**	0.007	−0.020**	0.007
PP * MD					0.028**	0.008	0.028**	0.008	0.027**	0.008
Classroom level (2)										
Grade							0.005	0.034	0.003	0.034
Classroom size							−0.007*	0.003	−0.007*	0.003
Descriptive norm (DN)							−0.199	0.660	−0.231	0.660
DN * PP									−0.336*	0.158
Social prestige norm (PN)							−0.107*	0.048	−0.106*	0.048
PN * PP									0.029*	0.014
Level 1 variance										
Intercept		0.005	0.003		0.003		0.003		0.003	
Level 2 variance										
Intercept		0.011	0.011		0.011		0.009		0.009	
Slope prioritizing popularity									< 0.001	
Fit statistics										
AIC	−1473.480		−1649.839		−1659.731		−1663.997		−1666.164	
BIC	−1460.031		−1618.457		−1623.866		−1610.200		−1594.434	
Deviance	−1479.480		−1663.839		−1675.731		−1687.997		−1698.164	
χ^2			184.36*** df = 4		11.891** df = 1		12.267* df = 4		10.167* df = 4	

Note. P1 = Proposition 1. P2 = Proposition 2. P3 = Proposition 3. We controlled for popularity, bullying, grade, and classrooms size. Grade was dummy coded (0 = primary school; 1 = secondary school).

*** $p < .001$.

** $p < .01$.

* $p < .05$.

levels of moral disengagement $t(648) = -1.29$, $p = .202$.

Next, we examined classroom level effects in addition to individual level effects (Model 3, Proposition 2). For both boys and girls, the fit of Model 4 was significantly better than the fit of Model 3. Fixed effects indicated that the more popular bullies were in a classroom (i.e., higher social prestige norms, hypothesis 5), the less likely boys and girls were to defend. The associations of the other classroom levels variables with defending varied between boys and girls. Boys were relatively unlikely to defend in classrooms that were relatively large, whereas girls were more likely to defend in secondary school than in primary school and in classrooms in which bullying was relatively common (i.e., high descriptive norms, hypothesis 4) rather than uncommon.

Finally, we tested proposition 3 (hypothesis 6) by examining the interaction between individual characteristics (i.e., prioritizing popularity) and contextual predictors (i.e., classroom prestige and descriptive norms) (Model 5). These interactions significantly improved the fit of Model 5 as compared to Model 4 among boys, but not among girls. Although a significant two-way interaction between descriptive norms of bullying and prioritizing popularity was found, simple effects showed that classroom descriptive norms of bullying were neither related to defending among boys with low levels of prioritizing popularity, $t(642) = 0.01$, $p = .994$, nor among boys with high levels of prioritizing popularity (see Fig. 3b), $t(642) = 0.71$, $p = .484$ (see Fig. 3b). We also found a significant two-way interaction between social prestige norms of bullying and prioritizing popularity. Fig. 3c shows that level of defending among boys with low levels of prioritizing popularity was lower in classrooms in which bullying was strongly positively related to popularity (i.e., high prestige norms) than in classrooms in which bullying was weakly or negatively related to popularity, $t(642) = 2.57$, $p = .013$. In contrast, among boys with high levels of

prioritizing popularity, classrooms prestige norms of bullying were not significantly associated with their defending status, $t(642) = 1.77$, $p = .082$.

4. Discussion

The current study examined the role of moral responsibility (i.e., moral disengagement) and social costs (i.e., prioritizing popularity and classroom norms) in defending victims of bullying in the classroom. Three propositions of the seminal model of defending were tested: (1) defending as act of moral courage depends on sequential stages, (2) the outcomes of stages 3 (moral responsibility) and 5 (social costs) of the seminal model of defending as act of moral courage (see Fig. 1) are influenced by both individual (i.e., moral disengagement – stage 3; prioritizing popularity – stage 5) and contextual factors (i.e., classroom norms – stage 5), (3) individual and contextual factors related to stage 5 interact with each other.

4.1. Individual characteristics related to defending

We first examined individual characteristics associated with defending. As stage 3 of the seminal model of defending states that one needs to feel morally responsible in order to display moral courage, our first hypothesis was that moral disengagement would be negatively associated with defending. Although previous studies have found negative associations between moral disengagement and defending, most studies did not examine this association separately for boys and girls (Gini et al., 2015; Sijtsema et al., 2014; Thornberg & Jungert, 2014; Thornberg, Wänström, Hong, & Espelage, 2017). We extended the findings of previous studies by showing that moral disengagement was

Table 4

Multilevel model predicting defending among girls from individual social-cognitive characteristics and classroom bullying norms.

	Defending among girls									
	Model 1 Empty model		Model 2 P2: individual level		Model 3 P1: sequential nature		Model 4 P2: contextual level		Model 5 P3: individual * contextual level	
	<i>b</i>	<i>SE</i>	<i>b</i>	<i>SE</i>	<i>b</i>	<i>SE</i>	<i>b</i>	<i>SE</i>	<i>b</i>	<i>SE</i>
Individual Level 1										
Intercept	0.233***	0.015	0.233***	0.015	0.234***	0.015	0.196***	0.018	0.192***	0.018
Popularity			0.099***	0.009	0.100***	0.009	0.100***	0.009	0.102***	0.009
Bullying			−0.602***	0.056	−0.600***	0.056	−0.600***	0.056	−0.596***	0.055
Prioritizing popularity (PP)			−0.014**	0.005	−0.015**	0.005	−0.015**	0.005	−0.016**	0.005
Moral disengagement (MD)			−0.015	0.01	−0.011	0.010	−0.011	0.010	−0.010	0.010
PP * MD					−0.021	0.013	−0.021	0.013	−0.018	0.013
Classroom level (2)										
Grade							0.080*	0.034	0.088*	0.034
Classroom size							−0.007	0.004	−0.007	0.004
Descriptive norm (DN)							1.402*	0.669	1.490*	0.668
DN * PP									0.117	0.190
Social prestige norm (PN)							−0.129*	0.048	−0.132**	0.048
PN * PP									0.009	0.017
Level 1 variance										
Intercept	0.007		0.006		0.006		0.006		0.006	
Level 2 variance										
Intercept	0.013		0.013		0.013		0.009		0.009	
Slope prioritizing popularity									0.003	
Fit statistics										
AIC	−1262.191		−1424.378		−1424.513		−1436.981		−1432.684	
BIC	−1248.504		−1392.441		−1388.014		−1382.232		−1359.685	
Deviance	−1268.191		−1438.378		−1440.513		−1460.981		−1464.684	
χ^2			170.187***	df = 4	2.135	df = 1	20.468***	df = 4	3.703	df = 4

Note. P1 = Proposition 1. P2 = Proposition 2. P3 = Proposition 3. We controlled for popularity, bullying, grade, and classrooms size. Grade was dummy coded (0 = primary school; 1 = secondary school).

*** $p < .001$.

** $p < .01$.

* $p < .05$.

negatively related to defending among boys but not among girls. In addition to perceiving oneself as morally responsible in a situation, stage 5 of the seminal model states that one also needs to have the courage to behave according to this moral principle by standing up against bullying despite potential negative social costs, such as a potential loss in popularity (Hypothesis 2). In line with our hypothesis, we replicated the study of Duffy et al. (2017); Prioritizing popularity was negatively associated with defending.

Next, our study extended the field of moral courage and defending by providing initial empirical support for the sequential idea of the seminal model (proposition 1). That is, we examined whether the outcome of the cost-benefit analysis (stage 5) would only be relevant among students who successfully accomplished stage 3 of the model by perceiving bullying as morally wrong (i.e., having low levels of moral disengagement, hypothesis 3). For boys, but not for girls, the results were in line with the sequential idea of the model. Specifically, we found that boys who justified bullying by mechanisms of moral disengagement were relatively unlikely to defend their peers, regardless of their priority of popularity. With regard to boys who did not justify bullying by mechanisms of moral disengagement, a positive association between prioritizing popularity and defending was found.

A potential explanation for the absence of a main effect and interaction including moral disengagement among girls may be due to fact that on average girls perceive bullying as more morally wrong than boys. Like previous studies (Caravita et al., 2012; Gini et al., 2015; Sijtsema et al., 2014; Thornberg & Jungert, 2014), we have shown that in general levels of moral disengagement with regard to bullying were lower among adolescent girls than boys and the variance in moral disengagement was smaller among girls than boys ($F(653) = 1.38$, $p < .001$). This is often explained by the fact that girls have higher

levels of moral competence and higher resistance against direct aggression than boys (Paciello, Fida, Tramontano, Lupinetti, and Caprara, 2008; Archer, 2004; Eisenberg, 2000). Associations between moral disengagement and defending may therefore be weaker among girls than boys due to a floor effect; a large part of the girls already successfully accomplished stage 3 of the model and hold themselves morally responsible for intervention. This floor effect may be particularly pronounced in the adolescent age period, as there is a general decrease in moral disengagement from childhood to late adolescence, perhaps due to improvements in moral reasoning (Eisenberg, 2000). We therefore stress the theoretical importance of testing gender differences with regard to stage 3 (moral responsibility) of the seminal model across different age groups to get further insight in the extent to which the seminal model is applicable to different age periods.

4.2. Classroom characteristics related to defending

We further extended the seminal model in proposition 2 by acknowledging that in addition to individual factors, contextual factors may also impact students' likelihood of defending (e.g., classroom norms). At the classroom level, we aimed to replicate the finding that students were less likely to defend in classrooms in which bullying was more common rather than less common (i.e., high descriptive norms, hypothesis 4) (Espelage et al., 2012; Peets et al., 2015), because defending has been found to be associated with more negative social consequences in classrooms in which bullying is relatively common (Lambe et al., 2017; Yun & Graham, 2018). This hypothesis was not supported; descriptive norms of bullying were not related to defending among boys. For girls, we even found an effect in the opposite direction; girls were not less, but more likely to defend in classrooms in which

a. Defending Boys: Prioritizing Popularity X Moral Disengagement

b. Defending Boys: Prioritizing Popularity X Classroom Descriptive Norms

c. Defending Boys: Prioritizing Popularity X Classroom Prestige Norms

Fig. 3. Cross-level interaction between classroom prestige norms of bullying, prioritizing popularity, and moral disengagement predicting defending. Note. *** $p < .001$. Simple slopes were computed at values 1 SD below and above the means according to the method of Preacher et al. (2006).

bullying was relatively common. This finding may be explained by the fact that there may be more opportunities to defend in classrooms in which bullying is more common than in classroom in which it is less common. This explanation applies to girls, because girls are more socialized to behave more prosocially than boys, are more prone to defend their victimized peers (Pouwels et al., 2018; Salmivalli et al., 1998), and may perhaps more strongly prefer low risky forms of defending (e.g., comforting the victim). Thus, our study contributes to the literature that high levels of bullying may not always reduce students' likelihood of defending, and may even increase the likelihood of defending. By examining the interaction between descriptive norms (i.e., frequency) and self-reported perceived costs and opportunities of defending, future research may reveal in which cases high descriptive norms may lead to increasing risks versus opportunities.

In contrast to descriptive norms of bullying, but in line with our expectation (hypothesis 5 and proposition 2), social prestige norms of

bullying were negatively related to defending. In line with Peets et al. (2015), students were less likely to defend in classrooms in which bullies were relatively popular (i.e., high prestige norms) than in classrooms in which bullies were not particularly popular or even relatively unpopular (i.e., low prestige norms). Although this finding suggests that students need more courage to defend victims in classrooms in which bullies are more popular than in classrooms in which they are less popular, we cannot directly draw this conclusion from the present study as we did not directly examine the negative consequences of defending. Future longitudinal or experimental studies are needed to directly examine the negative consequences of defending in classrooms in which bullying is strongly related to popularity, and how these consequences are related to students' subsequent motivation to defend. Thus, based on this study we can conclude that the perceived costs in situations that require moral courage (i.e., defending) are not only related to individual characteristics but also by contextual characteristics.

4.3. Interactions between individual characteristics and classroom characteristics

We finally tested the third proposition of the seminal model of defending by examining the interaction between individual characteristics (i.e., prioritizing popularity) and classroom characteristics (i.e., classroom descriptive and prestige norms of bullying, hypothesis 6). Although the associations between classroom norms and defending depended on individual characteristics among boys but not among girls, the direction of these interactions were not in line with our hypothesis. We expected that students who did prioritize popularity would only defend when the descriptive and social prestige norms of bullying are low (i.e., low potential costs of defending). However, no significant associations between classroom norms of bullying and defending were found among boys who prioritized popularity to a greater extent. Students who prioritize popularity may be relatively unlikely to defend their victimized peers, even when the potential costs of defending are low. Perhaps this may be due to the fact that prosocial behavior may be less effective than aggressive or coercive behavior in obtaining their social goal (i.e., increasing or maintaining status) (Reijntjes et al., 2018) even in contexts in which popularity is not so closely related to bullying. Moreover, the finding that popular adolescents are more likely to conform themselves to social norms related to bullying than unpopular students (Lucas-Molina et al., 2018; Peets et al., 2015; Yun & Graham, 2018), may be more likely to be impacted by actually being popular rather than prioritizing popularity.

We expected that the likelihood of defending among boys who did not prioritize popularity would not be affected by the potential negative costs in terms of status, as they are not concerned with a popular position in the peer group. In contrast to our hypothesis, we found that social prestige norms of bullying were positively related to defending among boys with low levels of prioritizing popularity. This suggests that even though these boys may be relatively unconcerned with a potential loss in social status by defending their peers, as they do not prioritize popularity, the level of bullying among their most popular classmates still seem to have an impact on their level of defending. Perhaps they may be afraid of other social consequences of defending rather than a potential loss in status, such as becoming the next target of bullying if they stand up for other victims.

In line with our hypotheses, we did not find significant interactions for girls. These findings were in line with the idea that associations of prioritizing popularity and classroom norms with defending would be stronger for boys than girls, as the costs of defending may be higher for boys (Pouwels et al., 2018; Salmivalli et al., 1998). Factors influencing the social costs of defending may therefore be more important for boys than girls. Thus, our study contributes to the literature by highlighting the relevance of examining gender differences related to the social costs of bullying.

4.4. Strengths, limitations, and implications for future research and practice

The present study extended previous research by examining moral courage in a relatively young sample of early adolescents. A main strength is that our design provided more insight into moral courage perceived by classmates. Thus, self-reports of prioritizing popularity and moral disengagement meaningfully related to how their defending was described by peers. As defending was measured by means of peer nominations, we tapped into students' reputation as defender in their classroom rather than observable defending behavior. It has been argued that peer-reports may be biased as they may be based on a students' general reputation in a classroom rather than on observable behaviors (Hymel, Wagner, & Butler, 1990). It is therefore still unclear to which extent student's reputation corresponds to observable defending behavior in a classroom setting. Behavioural paradigms could be used in future research to directly measure defending behavior. For example, defending could be measured in an experimental cyberball paradigm, in

which some classmates are socially excluded. This paradigm has been used before in a classroom session to measure the psychological outcomes of socially excluded children (Sandstrom et al., 2017). By excluding the classmates rather than the participant him or herself, this paradigm could be used to measure to what extent peers try to include their excluded classmates in the game as indicator of defending.

Another limitation of our assessment method is that we have used classroom-level determinants of costs (stage 5), which is a standard methodology in peer relations research. We found that the more popular bullies were in a classroom, the less likely students were to defend. However, by measuring classroom norms, we were not able to directly assess whether bullies' popularity is related to students' perceived potential social costs of defending. A suggestion for future research is therefore to manipulate the costs of defending and bullying norms by using a modified version of the survivor game. The idea behind the survivor game is that students need to stand up against bullying, and accordingly, they will be evaluated by peer judges (Reijntjes, Kamphuis et al., 2011; Reijntjes, Thomaes et al., 2011). By making some adjustments to the design of this paradigm, participants could be first presented with a narrative consisting of examples of how peers responded to other players who defended victims of bullying. The costs of defending could be manipulated by providing them with either negative examples (i.e., peers who aggress upon the defenders and retaliate against them) or neutral examples. Likewise, the norms of defending could be manipulated by providing them either examples of players who defend victims of bullying (players who try to comfort the victims and retaliate against bullying) or players who do not display defending behavior. After the manipulation, actual acts of defending could be measured, for example by using a cyberball paradigm.

The first theoretical contribution of our study is that we shed new light on the seminal model by providing initial evidence for the sequential nature of social-cognitive processes related to moral courage (proposition 1). It should be noted that although our study found that the effect of the cost-benefit analysis (stage 5) depends on students' level of moral disengagement (stage 3), we could not directly test the sequence of these stages due to the cross-sectional design of our study (i.e., whether students justified bullying at Time 1, before they made a cost-benefit analysis at Time 2). We recommend researchers to provide further empirical evidence for the sequential nature of all stages of the seminal model of defending (i.e., stages 1 to 5), as well as the sequence of stages related to other examples of moral courage (i.e., standing up against discrimination). This could be tested in longitudinal studies in which defending and the different stages are measured multiple times within one school year.

The second theoretical contribution of our study is that the original seminal model of defending only focused on individual characteristics related to the different stages of helping (Latané & Darley, 1970) and moral courage (Baumert et al., 2013). Although not all the interactions were in line with the hypotheses, the findings of the present study do provide support for the importance of contextual characteristics and the interaction between individual and contextual characteristics (proposition 2 and 3). The present study has taught us that stages 3 and 5 of the model of defending are impacted by both individual and contextual factors. Future research could extend upon this finding by testing individual and contextual characteristics to other phases of the model. For example, the adjusted seminal model of defending also underlines the importance of self-competence or self-efficacy (see Fig. 1, stage 4). Self-efficacy has been positively related to defending in previous research (Pronk et al., 2013). Moreover, retrospective reports revealed that intervention skills play a more important role in moral courage situations than helping situations (Greitemeyer et al., 2006). Especially if bullies are highly popular (i.e., contextual characteristics - social costs - stage 5), students may perceive themselves as incapable to stop bullying given the power that popular bullies have (i.e., individual characteristic - self-competence, stage 4). Accordingly, they may be less likely to defend.

The third theoretical contribution is that we shed more light on defending from a moral courage perspective. We argued that defending is an example of moral courage (i.e., outcome of the model), due to its high social costs and moral motivation. However, in real life, the distinction between defending as act of moral courage versus defending as act of helping may be subtler. With regard to the social costs of defending (stage 5), it should be noted that the costs may vary between different types of defending (i.e., the outcome of the model). Types of defending that are oriented at the bullies, such as confronting the bullies in public, may be related to more negative social consequences than types of defending that are oriented at victims, such as consoling the victim. As a result, students may need higher levels of courage to display bully-oriented defending than victim-oriented defending. With regard to the moral motivation and personal responsibility of intervention (stage 3), it should be noted that instead of defending out of the moral principle that bullying is wrong, it has been argued that some youth may also defend out of self-interest (Huitsing, Snijders, van Duijn, & Veenstra, 2014). As this implies that not all defending may be a form of moral courage, a suggestion for future research is to directly examine students' underlying motivation of defending in stage 3, so that defending out of moral principles could be distinguished from defending out of self-interest.

A suggestion for future research is to examine how the effects of individual and contextual characteristics related to each of the stages of the model, vary between different outcome variables. For example, future studies could examine to which extent moral disengagement, prioritizing popularity, and classroom norms are related to different types of defending, such as victim-oriented versus bully-oriented defending and direct versus indirect defending (Pronk et al., 2013; Reijntjes et al., 2016). Likewise, other behaviors than defending could be investigated, such as standing up to discrimination and illegal business practices.

The findings of the present study also have practical implications for the prevention of bullying and promotion of defending. In line with the sequential idea of the seminal model, our results suggest that the negative associations of prioritizing popularity with defending are stronger among students with low rather than high levels of moral disengagement. Thus, defending may be promoted by targeting moral cognitions related to bullying among students with relatively low levels of prioritizing popularity. In addition, our findings suggest that it may be important to promote a prosocial classroom context in which bullies are not popular as students may be more likely to defend in such a context. Previous experimental studies have shown that intergroup attitudes and positive behaviors such as confronting prejudice (Paluck, 2011) and water consumption (Smit, de Leeuw, Bevelander, Burk, & Buijzen, 2016) could be promoted by identifying and training influential peers. Likewise, in order to promote prosocial norms and obtain additional evidence for the effects of the social context on defending, it could be examined whether defending as morally courageous behavior could be diffused across peer networks through influencing popular peer trainers.

To summarize, we have shown that when considering moral courage in childhood and adolescence, such as defending the victims in bullying, individual differences related to moral disengagement for intervention, as well as an analysis of the costs and benefits of intervention are likely to promote moral courage. In line with the sequential idea of the model, we found that the outcome of the costs benefit analysis only alters the likelihood of defending among those students who held themselves personally responsible for intervention. In addition, people's social context may also determine to which extent they are likely to defend. We can conclude that it is important to consider interactions between moral responsibility and social costs at the individual and classroom level when trying to understand acts of moral courage, such as defending.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jesp.2019.103831>.

Acknowledgements

The authors are grateful to the school administrators, teachers, parents, and students who made this research possible.

References

- Almeida, A., Correia, I., & Marinho, S. (2010). Moral disengagement, normative beliefs of peer group, and attitudes regarding roles in bullying. *Journal of School Violence*, 9, 23–36. <https://doi.org/10.1080/15388220903185639>.
- Archer, J. (2004). Sex differences in aggression in realworld settings: A meta-analytic review. *Review of General Psychology*, 8, 291–322.
- Bandura, A., Barbaranelli, C., & Caprara, G. V. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology*, 71, 364–374. <https://doi.org/10.1037/0022-3514.71.2.364>.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2017). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-17. <https://cran.r-project.org/web/packages/lme4/lme4.pdf>.
- Baumert, A., Halmberger, A., & Schmitt, M. (2013). Interventions against norm violations: Dispositional determinants of self-reported and real moral courage. *Personality and Social Psychology Bulletin*, 39, 1053–1068. <https://doi.org/10.1177/0146167213490032>.
- van den Berg, Y. H. M., & Cillessen, A. H. N. (2012). Computerized sociometric and peer assessment: An empirical and practical evaluation. *International Journal of Behavioral Development*, 37, 68–76. <https://doi.org/10.1177/0165025412463508>.
- van den Broek, N., Deutz, M. H. F., Schoneveld, E. A., Burk, W. J., & Cillessen, A. H. N. (2016). Behavioral correlates of prioritizing popularity in adolescence. *Journal of Youth and Adolescence*, 45, 2444–2454. <https://doi.org/10.1007/s10964-015-0352-7>.
- Bronstein, P., Fox, B. J., Kamon, J. L., & Knolls, M. L. (2007). Parenting and gender as predictors of moral courage in late adolescence: A longitudinal study. *Sex Roles*, 56, 661–674. <https://doi.org/10.1007/s11199-007-9182-8>.
- Caravita, S. C. S., Gini, G., & Pozzoli, T. (2012). Main and moderated effects of moral cognition and status on bullying and defending. *Aggressive Behavior*, 38, 456–468. <https://doi.org/10.1002/ab.21447>.
- Caravita, S. C. S., Sijtsma, J. J., Rambaran, J. A., & Gini, G. (2014). Peer influences on moral disengagement in late childhood and early adolescence. *Journal of Youth and Adolescence*, 43, 193–207. <https://doi.org/10.1007/s10964-013-9953-1>.
- Cialdini, R. B. (2007). Descriptive social norms as underappreciated sources of social control. *Psychometrika*, 72, 263–268. <https://doi.org/10.1007/s11336-006-1560-6>.
- Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, 58, 1015–1026. <https://doi.org/10.1037/0022-3514.58.6.1015>.
- Cillessen, A. H. N. (2009). Sociometric methods. In K. H. Rubin, W. M. Bukowski, & B. Laursen (Eds.), *Handbook of peer interactions, relationships, and groups*. New York: The Guilford Press.
- Cillessen, A. H. N., Mayeux, L., Ha, T., & De Bruyn, E. H. (2014). Aggressive effects of prioritizing popularity in early adolescence. *Aggressive Behavior*, 40, 204–213. <https://doi.org/10.1002/ab.21518>.
- Dijkstra, J. K., & Gest, S. D. (2015). Peer norm salience for academic achievement, prosocial behavior, and bullying: Implications for adolescent school experiences. *The Journal of Early Adolescence*, 35, 79–96. <https://doi.org/10.1177/0272431614524303>.
- 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>.
- Duffy, A. L., Penn, S., Nesdale, D., & Zimmer-Gembeck, M. J. (2017). Popularity: Does it magnify associations between popularity prioritization and the bullying and defending behavior of early adolescent boys and girls? *Social Development*, 26, 263–277. <https://doi.org/10.1111/sode.12206>.
- Eagly, A. H., & Wood, W. (1991). Explaining sex differences in social behavior: A meta-analytic perspective. *Personality and Social Psychology Bulletin*, 17, 306–315. <https://doi.org/10.1177/0146167291173011>.
- Eisenberg, N. (2000). Emotion, regulation, and moral development. *Annual Review of Psychology*, 51, 665–697.
- 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>.
- Espelage, D., Green, H., & Polanin, J. (2012). Willingness to intervene in bullying episodes among middle school students: Individual and peer-group influences. *The Journal of Early Adolescence*, 32, 776–801. <https://doi.org/10.1177/0272431611423017>.
- Galdi, S., Maass, A., & Cadinu, M. (2017). Defending the victim of sexual harassment: The influence of civil courage and media exposure. *Psychology of Women Quarterly*, 41, 338–351. <https://doi.org/10.1177/0361684317709770>.
- Gini, G., Pozzoli, T., & Bussey, K. (2015). The role of individual and collective moral disengagement in peer aggression and bystanding: A multilevel analysis. *Journal of Abnormal Child Psychology*, 43, 441–452. <https://doi.org/10.1007/s10802-014-9920-7>.
- Greitemeyer, T., Fischer, P., Kastenmüller, A., & Frey, D. (2006). Civil courage and helping behavior. *European Psychologist*, 11, 90–98. <https://doi.org/10.1027/1016-9040.11.2.90>.
- Greitemeyer, T., Osswald, S., Fischer, P., & Frey, D. (2007). Civil courage: Implicit

- theories, determinants, and measurement. *Journal of Positive Psychology*, 2, 115–119. <https://doi.org/10.1080/17439760701228789>.
- Hoffman, M. L. (1977). Sex differences in empathy and related behaviors. *Psychological Bulletin*, 84, 712–722. <https://doi.org/10.1037/0033-2909.84.4.712>.
- Hox, J. J., Van De Schoot, R., & Moerbeek, M. (2018). *Multilevel analysis: Techniques and applications*. New York: Routledge.
- Huitsing, G., Snijders, T. A. B., van Duijn, M. A. J., & Veenstra, R. (2014). Victims, bullies, and their defenders: A longitudinal study of the co-evolution of positive and negative networks. *Development and Psychopathology*, 26, 645–659. <https://doi.org/10.1017/S0954579414000297>.
- Hymel, S., Rocke-Henderson, N., & Bonanno, R. A. (2005). Moral disengagement: A framework for understanding bullying among adolescents. *Journal of Social Sciences*, 8, 1–11.
- Hymel, S., Wagner, E., & Butler, L. J. (1990). Reputational bias: View from the peer group. In S. R. Asher, & J. D. Coie (Eds.). *Peer rejection in childhood* (pp. 156–186). New York: Cambridge University Press.
- Ickes, W., Gesn, P. R., & Graham, T. (2000). Gender differences in empathic accuracy: Differential ability or differential motivation? *Personal Relationships*, 7, 95–109. <https://doi.org/10.1111/j.1475-6811.2000.tb00006.x>.
- Juvonen, J., & Galván, A. (2008). Peer influence in involuntary social groups: Lessons from research on bullying. In M. J. Prinstein, & K. A. Dodge (Eds.). *Understanding peer influence in children and adolescents* (pp. 225–244). New York: Guilford Press.
- Kärnä, A., Voeten, M., Little, T. D., Alanan, 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>.
- Kuznetsova, A., Brockhoff, P. B., & Bojesen, R. H. (2018). lmerTest: Tests in linear mixed effects models. R package version 3.0-1. <https://cran.r-project.org/web/packages/lmerTest/lmerTest.pdf>.
- LaFontana, K. M., & Cillessen, A. H. N. (1998). The nature of children's stereotypes of popularity. *Social Development*, 7, 301–320. <https://doi.org/10.1111/1467-9507.00069>.
- 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., Hudson, C. C., Craig, W. M., & Pepler, D. J. (2017). Does defending come with a cost? Examining the psychosocial correlates of defending behaviour among bystanders of bullying in a Canadian sample. *Child Abuse & Neglect*, 65, 112–123. <https://doi.org/10.1016/j.chiabu.2017.01.012>.
- Latané, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?* New York, NY: Appleton.
- Li, Y., & Wright, M. F. (2014). Adolescents' social status goals: Relationships to social status insecurity, aggression, and prosocial behavior. *Journal of Youth and Adolescence*, 43, 146–160. <https://doi.org/10.1007/s10964-013-9939-z>.
- Lopez, S., O'Byrne, K. K., & Petersen, S. (2003). Profiling courage. In S. Lopez, & C. R. Snyder (Eds.). *Positive psychology assessment: A handbook of models and measures* (pp. 185–197). Washington: APA.
- Lucas-Molina, B., Giménez-Dasí, M., Fonseca-Pedrero, E., & Pérez-Albéniz, A. (2018). What makes a defender? A multilevel study of individual correlates and classroom norms in explaining defending behaviors. *School Psychology Review*, 47, 34–44. <https://doi.org/10.17105/SPR-2017-0011.V47-1>.
- Malti, T., & Ongley, S. F. (2013). The development of moral emotions and moral reasoning. In M. Killen, & J. Smetana (Eds.). *Handbook of moral development*. New York: Taylor and Francis.
- Marks, P. E. L., Babcock, B., Cillessen, A. H. N., & Crick, N. R. (2013). The effects of participation rate on the internal reliability of peer nomination measures. *Social Development*, 22, 609–622. <https://doi.org/10.1111/j.1467-9507.2012.00661.x>.
- Mazzone, A., Yanagida, T., Caravita, S. C. S., & Strohmeier, D. (2018). Moral emotions and moral disengagement: Concurrent and longitudinal associations with aggressive behavior among early adolescents. *The Journal of Early Adolescence*, 0272431618791276. <https://doi.org/10.1177/0272431618791276>.
- 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>.
- Niesta, D., Greitemeyer, T., Fischer, P., & Frey, D. (2010). Why mood affects help giving, but not moral courage: Comparing two types of prosocial behaviour. *European Journal of Social Psychology*, 40, 1136–1157. <https://doi.org/10.1002/ejsp.717>.
- van Noorden, T. H. J., Haselager, G. J. T., Cillessen, A. H. N., & Bukowski, W. M. (2014). Dehumanization in children: The link with moral disengagement in bullying and victimization. *Aggressive Behavior*, 40, 320–328. <https://doi.org/10.1002/ab.21522>.
- Osswald, S., Frey, D., & Streicher, B. (2012). Moral courage. In E. Kals, & J. Maes (Eds.). *Justice and conflicts* (pp. 391–405). Berlin Heidelberg: Springer.
- Osswald, S., Greitemeyer, T., Fischer, P., & Frey, D. (2010). What is moral courage? Definition, explication, and classification of a complex construct. In C. Pury, & S. Lopez (Eds.). *The psychology of courage* (pp. 149–164). Washington, DC: American Psychological Association.
- Paciello, M., Fida, R., Tramontano, C., Lupinetti, C., & Caprara, G. V. (2008). Stability and change of moral disengagement and its impacts on aggression and violence in late adolescence. *Child Development*, 79, 1288–1309.
- Paluck, E. L. (2011). Peer pressure against prejudice: A high school field experiment examining social network change. *Journal of Experimental Social Psychology*, 47, 350–358. <https://doi.org/10.1016/j.jesp.2010.11.017>.
- Parkhurst, J. T., & Hopmeyer, A. (1998). Sociometric popularity and peer-perceived popularity: Two distinct dimensions of peer status. *The Journal of Early Adolescence*, 18, 125–144. <https://doi.org/10.1177/0272431698018002001>.
- Peets, K., Poyhonen, 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., Lansu, T. A. M., & Cillessen, A. H. N. (2016). Participant roles of bullying in adolescence: Status characteristics, social behavior, and assignment criteria. *Aggressive Behavior*, 42, 239–253. <https://doi.org/10.1002/ab.21614>.
- 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*. <https://doi.org/10.1111/sode.12294>.
- Pöyhönen, V., Juvonen, J., & Salmivalli, C. (2010). What does it take to stand up for the victim of bullying? The interplay between personal and social factors. *Merrill-Palmer Quarterly*, 56, 143–163. <https://doi.org/10.1353/mpq.0.0046>.
- 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>.
- Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interaction effects in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, 31, 437–448. <https://doi.org/10.3102/10769986031004437>.
- Pronk, J., Goossens, F. A., Olthof, T., de Mey, L., & Willems, A. M. (2013). Children's intervention strategies in situations of victimization by bullying: Social cognitions of outsiders versus defenders. *Journal of School Psychology*, 51, 669–682. <https://doi.org/10.1016/j.jsp.2013.09.002>.
- R Core Team (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <http://www.R-project.org/>.
- Reijntjes, A., Kamphuis, J. H., Prinzie, P., Boelen, P. A., van der Schoot, M., & Telch, M. J. (2011). Prospective linkages between peer victimization and externalizing problems in children: A meta-analysis. *Aggressive Behavior*, 37, 215–222. <https://doi.org/10.1002/ab.20374>.
- Reijntjes, A., Thomaes, S., Kamphuis, J. H., Bushman, B. J., De Castro, B. O., & Telch, M. J. (2011). Explaining the paradoxical rejection-aggression link: The mediating effects of hostile intent attributions, anger, and decreases in state self-esteem on peer rejection-induced aggression in youth. *Personality and Social Psychology Bulletin*, 37, 955–963. <https://doi.org/10.1177/0146167211410247>.
- Reijntjes, A., Vermande, M., Olthof, T., Goossens, F. A., & Aleva, L. (2016). Defending victimized peers: Opposing the bully, supporting the victim, or both? *Aggress. Behav.* 42, 585–597. <https://doi.org/10.1002/ab.21653>.
- Reijntjes, A., Vermande, M. M., Olthof, T., Goossens, F. A., Vink, G., Aleva, L., & van der Meulen, M. (2018). Differences between resource control types revisited: A short term longitudinal study. *Social Development*, 27, 187–200. <https://doi.org/10.1111/sode.12257>.
- 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>.
- Salmivalli, C., Lagerspetz, K., Björkqvist, K., Österman, K., & Kaukiainen, A. (1996). Bullying as a group process: Participant roles and their relations to social status within the group. *Aggressive Behavior*, 22, 1–15. <https://doi.org/10.1002/%28SICI%291098-2337%281996%2922:1%3C1::AID-AB1%3E3.0.CO;2-T>.
- Salmivalli, C., Lappalainen, M., & Lagerspetz, K. M. J. (1998). Stability and change of behavior in connection with bullying in schools: A two-year follow-up. *Aggressive Behavior*, 24, 205–218. <https://doi.org/10.1002/%28SICI%291098-2337%281998%2924:3%3C205::AID-AB5%3E3.0.CO;2-J>.
- 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>.
- Sandstrom, M. J., Deutz, M. H. F., Lansu, T. A. M., van Noorden, T. H. J., Karremans, J. C., & Cillessen, A. H. N. (2017). Unanimous versus partial rejection: How the number of excluders influences the impact of ostracism in children. *Aggressive Behavior*, 43, 190–203. <https://doi.org/10.1002/ab.21674>.
- Sijtsema, J. J., Rambaran, J. A., Caravita, S. C. S., & Gini, G. (2014). Friendship selection and influence in bullying and defending: Effects of moral disengagement. *Developmental Psychology*, 50, 2093–2104. <https://doi.org/10.1037/a0037145>.
- Smit, C. R., de Leeuw, R. N. H., Bevelander, K. E., Burk, W. J., & Bijl, M. (2016). A social network-based intervention stimulating peer influence on children's self-reported water consumption: A randomized control trial. *Appetite*, 103, 294–301. <https://doi.org/10.1016/j.appet.2016.04.011>.
- Spadafora, N., Marini, Z., & Volk, A. (2018). 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* 082957351879375. <https://doi.org/10.1177/0829573518793752>.
- Staub, E. (2003). *The psychology of good and evil. Why children, adults, and groups help and harm others*. Cambridge: Cambridge University Press.
- Thornberg, R., & Jungert, T. (2014). School bullying and the mechanisms of moral disengagement. *Aggressive Behavior*, 40, 99–108. <https://doi.org/10.1002/ab.21509>.
- 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>.
- Thornberg, R., Wänström, L., Hong, J. S., & Espelage, D. L. (2017). Classroom relationship qualities and social-cognitive correlates of defending and passive bystanding in school bullying in Sweden: A multilevel analysis. *Journal of School Psychology*, 63, 49–62. <https://doi.org/10.1016/j.jsp.2017.03.002>.
- Yun, H., & Graham, S. (2018). Defending victims of bullying in early adolescence: A multilevel analysis. *Journal of Youth and Adolescence*, 47, 1926–1937. <https://doi.org/10.1007/s10964-018-0869-7>.