Moral Disengagement of Pure Bullies and Bully/Victims: Shared and Distinct Mechanisms

Short Title: MORAL DISENGAGEMENT IN PURE BULLIES AND BULLY/VICTIMS

Kevin C. Runions[a, b], Thérèse Shaw[a] , Kay Bussey[c], Robert Thornberg[d], Christina Salmivalli[e], & Donna S.Cross[a,f]

Kevin Runions: Telethon Kids Institute, Perth Children's Hospital, PO Box 855, West Perth, Western Australia 6009, Australia, (e) kevin.runions@telethonkids.org.au; (t) 618 6319 1478

Thérèse Shaw: Telethon Kids Institute, Perth Children's Hospital, PO Box 855, West Perth, Western Australia 6009, Australia, (e) therese.shaw@telethonkids.org.au; (t) 618 6319 1478

Kay Bussey: Centre for Emotional Health, C3A 731, Macquarie University, NSW 2109, Australia. (e.) kay.bussey@mq.edu.au; (t) +61 2 9850 8085

Robert Thornberg: Department of Behavioural Sciences and Learning, Linköping University

SE-58183 Linköping Sweden. . (e) Robert.thornberg@liu.se; (t) +46 13 282118

Christina Salmivalli. Department of Psychology, University of Turku, Assistentinkatu 7, 20500, Turku Finland (e.) Christina.salmivalli@utu.fi; (t) +358 29 450 2715

Donna Cross: Telethon Kids Institute, Perth Children's Hospital, PO Box 855, West Perth, Western Australia 6009, Australia, (e) donna.cross@telethonkids.org.au; (t) 618 6319 1478

Author Affiliations and Research Interests (to be pasted when uploading manuscript)

Kevin C Runions is a Senior Research Fellow with the Faculty of Medicine, Dentistry and Health Sciences at the University of Western Australia and the Telethon Kids Institute. His major research interests include psychological enablers of bullying and aggression, the development of personality disorders, and school mental health intervention.

Thérèse Shaw is Biostatistician and Project Leader with the Telethon Kids Institute and has an adjunct position with the Faculty of Medicine, Dentistry and Health Sciences at the University of Western Australia. Her major research is on the rigorous design and evaluation of school-based health promotion programs.

Kay Bussey is an Associate Professor in the Department of Psychology at Macquarie University in Sydney, Australia. Her major research interests are guided by Bandura's social cognitive theory which considers both social and personal factors involved in human agency, including moral disengagement and self-efficacy.

Robert Thornberg is a Professor in the Department of Behavioural Sciences and Learning at Linköping University, Sweden. His research focuses on social and moral processes involved in bullying, bystander rationales, reactions and actions, and students' perspectives and explanations.

Christina Salmivalli is Professor of Psychology at University of Turku, Finland and Visiting Professor at Shandong Normal University, Jinan, China. Her research focus is on testing the effectiveness and sustainability of bullying prevention programs, and mediating mechanisms of program effects.

Donna Cross is a Professor with the Faculty of Medicine, Dentistry and Health Sciences at the University of Western Australia and the Telethon Kids Institute. Her major research interests include translational research, child and adolescent social and emotional wellbeing and bullying and cyberbullying prevention
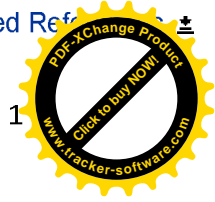
Abstract

The vast majority of adolescents recognize that bullying is morally wrong, yet bullying remains a problem in secondary schools, indicating young people may disengage from their moral values to engage in bullying. But it is unclear whether the same mechanisms enabling moral disengagement are active for bully/victims (who both bully and are bullied) as for pure bullies (who are not targets of bullying). This study tested the hypotheses that mechanisms of moral disengagement, including blaming the victim and minimizing the impact of bullying, may operate differently in bully/victims compared to pure bullies. From a sample of 1895 students from grades 7-9 (50.6% female; 83.4% from English speaking homes), 1870 provided self-reports on bullying involvement and mechanisms of moral disengagement associated with bullying. Two cut-offs were compared for bullying involvement (as perpetrator and as target of bullying) during the previous school term: a conservative cut-off (every few weeks or more often) and a liberal cut-off (once-or-twice). Using the conservative cut-off, both pure bullies and bully/victims enlisted moral disengagement mechanisms to justify bullying more than did uninvolved students and pure victims, with no significant difference in scores on any of the moral disengagement scales between pure bullies and bully/victims. For the liberal cut-off, bully/victims reported lower overall moral disengagement scores than did pure bullies, and specifically less distortion of consequences, diffusion of responsibility, and euphemistic labelling. This study advances bullying research by extending the role of moral disengagement in bullying episodes beyond pure bullies to victims, both pure victims and bully/victims. Examination of specific moral disengagement mechanisms and the extent of involvement in bullying enabled a more nuanced differentiation between the bullying groups. These results will inform future interventions aimed at reducing the use of moral disengagement mechanisms that sustain bullying and victimization. Targeted interventions are needed to challenge specific moral disengagement mechanisms from the perspectives of pure bullies and bully/victims.

Introduction

Few adolescents, if asked, would be likely to indicate that bullying is the morally right thing to do. Yet bullying behaviors remain a problem for schools. Bullying reflects intentional aggression conducted within a social relationship marked by a power imbalance leading to an inability by the target to defend him/herself (Olweus, 2013). As such, bullying takes diverse modes, including physical aggression, verbal aggression, relational aggression (the manipulation of social relationships to cause harm); and it can take place in person or from a distance via communication technologies (cyberbullying). As children grow, bullying becomes less commonly direct and physical and more commonly indirect and relational (Yeager, Fong, Lee, & Espelage, 2015). In addition, school-based bullying prevention programs can become ineffective – in some cases iatrogenic – in the adolescent years according to a developmentally sensitive meta-analysis (Yeager et al., 2015). Further inquiry into the drivers of adolescent bullying behaviors is required.

By the teenage years, most young people know that bullying is a moral transgression. In one Australian study, 99% of adolescents rated cyberbullying as "wrong" (Bussey, Fitzpatrick, & Raman, 2015). Even pre-adolescents conceptualize bullying as morally transgressive due to the harm caused (Thornberg, Thornberg, Alamaa, & Daud, 2016). This raises the question of how young people justify their bullying behavior in light of clear social norms against this behavior. Previous research indicates that adolescents who bully others are more likely to engage in moral disengagement: a set of processes to dampen the negative self-judgements that can arise from engaging in behavior one knows is morally wrong (Bandura, 1999). But to date, little is known about whether moral disengagement is equally a factor in bullying conducted by young people who are not themselves victims of bullying ("pure bullies") and those who both perpetrate and are targets of bully ("bully/victims"). Most studies do not differentiate between these two groups. However, as the motives for these two groups have been shown to diverge in past studies (e.g., Runions, Salmivalli, Shaw, Burns, & Cross, 2018), moral processes that can facilitate bullying may also diverge between these two groups. This study investigated the role of moral disengagement in bullying behavior, comparing pure bullies and bully/victims with uninvolved students and pure victims in a cross-sectional study of early adolescent students. In doing so, the study focused on differences in overall moral disengagement as well as differences in specific mechanisms of moral disengagement, with the aim of identifying specific targets for anti-bullying intervention and prevention programs amongst adolescents.

Bullying, Victimization, and their Co-occurrence

Bullying is a pervasive problem in school environments. A meta-analysis of self-reported involvement found that approximately 36% of young people have been a target of bullying victimization at some point in their lives, and 34.5% report perpetration (Modecki, Minchin, Harbaugh, Guerra, & Runions, 2014). Such estimates, however, blur the fact that victimization and perpetration of bullying are by no means mutually exclusive. Some young people are involved both as perpetrators *and* as targets of bullying, so called bully/victims. Bully/victims have been of increasing interest to peer aggression researchers since seminal work on aggressive victims in the late 20th century (Pellegrini, Bartini, & Brooks, 1999; Schwartz, 2000). As reviewed by Yang and Salmivalli (2013), prevalence estimates for bully/victim status have varied from 0.4% to 29% when examined via self-report (Solberg, Olweus, & Endresen, 2007). An international study of 40 countries used a cut-off of 2-3 times per month for bullying victimization and perpetration, and found a prevalence of 3.6% for bully/victim status (Craig et al., 2009).

A decade of research has provided a range of characteristics that define bully/victims compared to pure bullies. Bully/victim status appears to confer accumulated risks associated with both bullying perpetration and victimization. Data from the Avon Longitudinal Study of Parents and Children provided estimates of the unique risk conferred from bullying, victimization, and their co-incidence, controlling for a wide range of covariates (Lereya, Copeland, Zammit, & Wolke, 2015). Being a victim of bullying at age 13 predicted elevated risk of psychotic experiences, depression, anxiety and other mental health problems at age 18 (Wolke, Lereya, Fisher, Lewis, & Zammit, 2014). Bully/victims are at greater risk of subsequent antisocial behavior than are pure bullies and are more likely to be socially isolated than either pure bullies or pure victims (Georgiou & Stavrinides, 2008). Mental health risks of bully/victims appear to extend into early adulthood (Sigurdson, Undheim, & Sund, 2015). These data suggest bully/victims may be at greatest risk of adverse short- and long-term social and emotional outcomes, highlighting the need for clearer intervention targets for these young people.

Along with these differential risk profiles, there is some evidence of important differences between pure bullies and bully/victims. "Aggressive victims" (who may or may not engage in bullying per se) show poor self-regulation (Toblin, Schwartz, Gorman, & Abou-Ezzeddine, 2005), although not all studies find a difference between pure bullies and bully/victims with regard to emotion regulation (e.g., Garner & Hinton, 2010) or executive function (Verlinden et al., 2014). Bully/victims have been found to be more frequently bullied than pure victims (Yang & Salmivalli, 2013). The distinct characteristics of bully/victims suggest

effective intervention strategies for bully/victims are likely to diverge from those that are successful for pure

bullies or pure victims.

For bully/victims, poor self-regulation and hostile attribution tendencies (e.g., Pouwels, Scholte, van

Noorden, & Cillessen, 2016) might suggest bullying behavior that arises more spontaneously than for pure

bullies, who might be imagined as more deliberate in their use of bullying to achieve goals. But in a recent

study, bully/victims were just as likely as pure bullies to espouse deliberate, planned reward-related motives for

aggression, and *more* likely to espouse deliberate, planned revenge as a motive (Runions et al., 2018). This

suggests that both pure bullies and bully/victims may engage in deliberate bullying knowing that bullying is

morally wrong. Thus, moral disengagement may be a requisite cognitive process to engage in bullying for both

bully/victims and pure bullies That is to say, moral disengagement provides an approach to studying how young

people might engage in bullying even when they "should know better".

Moral Disengagement and Bullying

The framework known as moral disengagement describes how individuals engage in behaviors they

know to be morally wrong without feeling remorse, guilt or other self-sanctioning emotions (Bandura,

Barbaranelli, Caprara, & Pastorelli, 1996). Moral disengagement is conceptualized as a set of socio-cognitive

mechanisms that afford the capacity to shut down self-sanctions such as shame, guilt, and negative self-

evaluation that would normally accompany violating one's own moral standards. Bandura (1999) described

mechanisms including the cognitive reframing of the morally compromised act, for example by providing a

greater moral "ends" that justifies the means (the *moral justification* mechanism). Other mechanisms serve to

reduce one's personal agentic role in the act; this can be achieved by "passing the buck" (i.e., the *displacement

of responsibility* mechanism), for example. Alternately, one may cognitively restructure the imagined

consequences of the act, for example by minimizing the harm that the act incurs, and by attributing imaginary

consequences (e.g., imagining that bullying "builds character"; *distortion of consequences*); and by blaming,

vilifying or dehumanizing those affected by the act (*victim blame*).

The role of moral disengagement in aggressive behavior has been a research focus since Bandura and

colleagues' first empirical work with adolescents on the topic (e.g., Bandura et al., 1996). The relevance of the

moral disengagement framework to bullying has been established for over a decade: an early study by Menesini

et al. (2003) found high levels of egocentric disengagement – failing to account for the impact of actions on

others – to be common amongst bullies compared to victims and outsiders. A meta-analysis found a robust

association of moral disengagement and bullying behaviors with a mean effect size of 0.25 (Gini, Pozzoli &

Hymel, 2014); the association of moral disengagement to aggression (including bullying) was found to be greatest in adolescent samples, compared to child samples.

Several studies have examined moral disengagement amongst pure bullies and bully/victims. In a Danish study of late primary school students, bullies, victims and bully/victims all showed elevated levels of moral disengagement relative to uninvolved children; no differences were found between groups (Obermann, 2011). An Australian study, including both primary and secondary schools students, found that both pure bullies and bully/victims showed increased moral disengagement compared to non-perpetrating students (including primary and secondary students; Tanrikulu & Campbell, 2015). In another study of bullying amongst adults in prison, pure bullies and bully/victims showed comparable levels of overall moral disengagement (South & Wood, 2006).

However, the earliest study of moral disengagement amongst bully/victims and pure bullies found that the role of moral disengagement in bullying perpetration depended on the extent of the young person's involvement in bullying perpetration and victimization (Hymel, Rocke-Henderson, & Bonanno, 2005). Canadian students in grades 8-10 who self-reported extensive bullying perpetration (i.e., "once a week" or more) had high levels of moral disengagement, regardless of their victimization experiences (i.e., no significant difference between pure bullies and bully/victims). But for those who reported less extensive bullying, moral disengagement scores were significantly higher amongst those who had no reported history of being victimized (i.e., pure bullies) than amongst bully/victims. Hymel et al. posited that the experience of being a victim might reduce the ability to provide rationalizations for bullying behavior; but Hymel et al. did not account for why those more extensively involved (i.e., bullying others once a week or more) did not show the same pattern. Thus, analyses that are sensitive to the degree of involvement as a bully and/or victim are required.

Moral disengagement is a multifaceted construct with conceptually distinct mechanisms. The finding that moral disengagement overall – averaged across the mechanisms – has a role in bullying behavior is arguably of limited utility in planning intervention and prevention efforts to reduce bullying. As Hymel et al. (2005) have observed, morally disengaging with regard to bullying is likely to be a gradual developmental process; consequently, moral disengagement is not a monolithic developmental acquisition. In light of this, it is remarkable that few studies have examined the specific moral disengagement mechanisms to discern patterns relevant to bullying behavior during childhood and adolescence. An Italian study of school children (aged 8 – 10 years) examined four factors of moral disengagement (regarding a range of antisocial behaviors) aligning with Bandura's formulation: cognitive restructuring, minimizing one's agentic role, disregarding/distorting the

consequences, and blaming/dehumanizing the victim (Pozzoli, Gini, & Vieno, 2012). Only cognitive restructuring significantly predicted peer-rated pro-bullying behaviors. A study by Thornberg and Jungert (2014) examined the use of moral disengagement mechanisms amongst Swedish children aged 10-14 years with a scale specific to bullying. They found bullying perpetration was best predicted by victim-blame and moral justifications for bullying. Euphemistic labeling and advantageous comparison were not associated with bullying behavior in their analysis. These findings were replicated in a recent study for direct bullying, whereas indirect bullying was only associated with victim blame (Bjärehed, Thornberg, Wänström, & Gini, 2019). This focus on specific mechanisms provides important clear targets for interventions to reduce bullying.

Extant research provides few clues to the relevance of specific moral disengagement mechanisms to the co-occurrence of bullying perpetration and victimization. Thornberg and Jungert (2014) included victimization as a control variable in their regression analysis, and their findings may therefore be more comparable to pure bullies. They also tested a regression model with victimization as the dependent variable and bullying as a control variable. It is noteworthy that the bivariate correlation matrix showed that victimization only correlated significantly and positively with distortion of consequences, $r = .14$, $p < .01$, and victim blame, $r = .18$, $p < .001$. In the final multivariate regression model predicting victimization – in which bullying perpetration was controlled—neither of these mechanisms retained significance. This suggests these mechanisms may be specifically implicated in bully/victim status.

To date, only one study has directly examined specific mechanisms of moral disengagement and compared pure bullies with bully/victims. Perren, Gutzwiller-Helfenfinger, Malti, and Hymel (2012) examined the moral justifications of 516 Swiss students (aged 12-18) as a function of self-reported bullying and victimization. Using a vignette involving relational aggression, students were asked to explain the perpetrator's thinking. These accounts were coded into morally responsible and morally disengaged justifications. In this study, pure bullies ($n = 74$) and bully/victims ($n = 20$) showed no significant differences in use of morally responsible or disengaged justifications. Examination of mean values, however, suggests that pure bullies had notably higher scores for these mechanisms than did bully/victims. However, the limited power due to the small sample size may have resulted in Type 2 (false negative) error, and responses to hypothetical vignettes may diverge from self-report responses. Further research using moral disengagement items specific to bullying that accounts for the co-occurrence of bullying perpetration and victimization is needed.
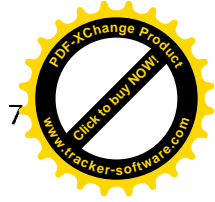
The Current Study

It is currently unclear how bullying-related moral disengagement may differentially be associated with bullying behaviors amongst pure bullies and bully/victims. Based on Perren et al. (2012), pure bullies were hypothesized to report higher levels of moral disengagement overall than bully/victims. The findings of Thornberg and Jungert (2014) suggest some moral disengagement mechanisms may be more pronounced amongst pure bullies and other mechanisms may be common to anyone who perpetrates bullying. Additionally, there may be mechanisms that are more pronounced amongst bully/victims. To investigate this possibility, this study examined if pure bullies and bully/victims endorse different mechanisms of bullying-relevant moral disengagement. As noted, Thornberg and Jungert (2014) found distortion of consequences and victim-blame to be positively correlated with victimization experiences. For bully/victims, the cognitively dissonant belief that being bullied is "no big deal" may arise to enable coping with one's own victimization. It may be that distorting consequences of bullying and identifying reasons why others "deserve" to be bullied are important to push past one's own victimization experiences to perpetrate bullying against others. Thus, distortion of consequences and victim blame were hypothesized to be elevated for bully/victims.

Furthermore, compared to bully/victims, who are more likely to be socially isolated and rejected by peers (Georgiou & Stavrinides, 2008), pure bullies may be more likely to be a part of groups who jointly engage in bullying (Pouwels, Lansu, & Cillessen, 2018). Some bullying researchers (e.g., Salmivalli, Lagerspetz, Bjorkqvist, Osterman, & Kaukiainen., 1996) have divided bullying perpetration roles into those who initiate bullying (ringleaders) and those who take part once initiated (variously called "assistants", "henchmen" or "followers"). Whether bullying perpetration occurs as initiator or follower, this collective phenomenon may afford the opportunity for diffusion of responsibility that is not as likely for bully/victims who are more likely to be socially isolated. Consequently, pure bullies were hypothesized to be especially likely to engage in minimization of their personal agentic role in bullying, as reflected in the diffusion of responsibility mechanism.

In light of the pioneering work of Hymel et al (2005), which found that the role of moral disengagement varied depending on the extent of bullying involvement, the current study aimed to examine how robust the role of moral disengagement is in bullying perpetration. Involvement in bullying may be assessed by a range of methods. Some researchers have used ad hoc cut-offs for continuous scales (e.g., one standard deviation above the mean; Sampasa-Kanyinga, Chaput, Hamilton, & Colman, 2018). Others have followed Solberg and Olweus (2003) in adopting a cut-off for "meaningful" bullying involvement of 2-3 times per month (e.g., Jackson & Vaughn, 2018). However, even the seminal Solberg and Olweus study found significantly

greater problems with depressive tendencies, negative self-evaluation and social disintegration amongst students who reported being bullied "once or twice", compared to students not involved in bullying at all (albeit less severe problems than those who experience greater bullying). Similarly, those who reported bullying perpetration "once or twice" had significantly higher general aggression and antisocial behavior scores. Rather than prematurely setting a cut-off that may not capture fully the phenomenon, cut-off points can be manipulated to evaluate the robustness of findings. In the current study, two cut-offs were compared: a liberal cut-off of any self-reported involvement (i.e., *once or twice* a school term or greater) and a more conservative cut-off (i.e., *every few weeks* or greater).

Methods

Recruitment and Participants

Following institutional human ethics approval and securing permissions from the Department of Education, the Catholic Education office, and from individual schools, students from grades 7–9 (ages 12–15 years) were recruited from 12 Australian government and non-government schools. Government schools required active parental (opt-in) consent; non-government schools permitted parental opt-out consent for survey completion. For opt-in recruitment, school newsletter items were provided, and schools sent two emails to parents and one letter was sent home with a internet link to a web-based consent portal. No inducements to participation (e.g., entry in a draw for returning forms) were permitted to encourage parental consent. For the non-government schools, families were contacted on three occasions via two rounds of email/SMS and a final round of hard copy letters providing a link to the consent portal. For the opt-in schools (n = 7), out of an estimated 4785 eligible students, consent was received for 339 participants (7.08%). By contrast, for the opt-out schools (n = 5), of an estimated potential 1996 students, 1597 (80.01%) parents consented. This resulted in an overall 28.6% consent rate.

The two samples differed in demographic characteristics. Compared to the non-government school students, relatively more of the government-school students were in grade 7 (40.4% vs 33.5%) and fewer in grade 9 (23.7% vs 32.4%; $\chi^2 = 10.8$, $p = .005$); relatively more came from homes where English was the only language spoken (91.1% versus 82.6%; $\chi^2 = 14.9$, $p < .001$). Overall, half of the students were from affluent families based on the Family Affluence Scale (Boyce, Torsheim, Currie, & Zambon, 2006); fewer government-school students were from high socio-economic status families (41.3% versus 52.1%). No differences were found between the groups on gender ($\chi^2 = 0.8$, $p = .370$) or Indigenous status ($\chi^2 = 0.4$, $p = .513$). With regard to key variables in the analysis, the two groups did not differ significantly on their mean scores for the total score
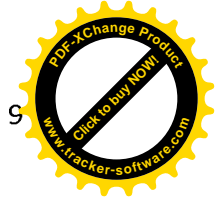
for moral disengagement ($t(1820) = 1.2$, $p = .229$) or on any of the subscales, and did not differ in the distribution of cases between the bully/victim groups for the conservative cut-off ($\chi^2 = 3.6$, $p = .307$), but did differ based on the liberal cut-off ($\chi^2 = 10.3$, $p = .017$) – the definition of the conservative and liberal involvement cut-offs is provided below. Relatively fewer non-involved students were from the government schools compared to the non-government schools (46.5% vs. 50.2%), more students were classified into the pure victim group (42.9% versus 35.6%) and fewer in the pure bully group (1.8% versus 4.6%), while the percentages in the bully/victim group were similar (8.7% versus 9.6%). In summary, the opt-in consent procedure used in government schools resulted in an underrepresentation of students from more culturally diverse and higher socio-economic status families and students in grade 9. Additionally, students who reported being a target (even once or twice) but not a perpetrator of bullying, were over-represented (by 7.3%) in government schools. Hence, differences between the consent groups were controlled for these in the analyses.

The full sample included 1936 students from 12 schools in Western Australia. However, at the end of the survey, a single item asked if the participant had been honest in responding. We excluded cases who indicated they were "not honest at all" or "honest once in a while" (n = 30). Students who completed the entire survey in five minutes or less (n = 11) were also excluded: The median time to complete the survey was 20 minutes and the middle 50% of students took 16.4 - 25.2 minutes); completing the survey in under five minutes indicates the responses were not based on reading the questions or reflecting on their experiences. These 41 excluded cases did not differ significantly from those retained on grade level, family socio-economic status, home language or Indigenous status, but did differ on gender ($\chi^2 = 5.7$, $p = .017$), with 2.9% of boys excluded versus 1.3% of girls. The dropped cases scored significantly higher (based on a Mann-Whitney test) on the victim-blame moral disengagement subscale (Z = -3.6, $p < .001$; dropped cases $M = 2.1$, $SD = 1.42$; retained cases $M = 1.3$, $SD = 0.63$). No differences were found on the total score or other moral disengagement subscales. There was also no significant association with bully/victim group ($\chi^2 = 4.3$, $p = .230$) using the liberal cut-off. (There were too few cases within the involved groups amongst the dropped cases to test this association based on the conservative cut-off.) Given the questionability of the validity of the responses to the survey questions, interpretation of any differences between the responses for the cases dropped and those retained in the analyses is fraught as it is unclear whether differences are real or a consequence of the dishonest or careless responses. This is particularly pertinent for the scale measuring moral disengagement as this scale was placed towards the end of the survey.

The final analysis sample included 1895 students. Of these 50.6% (n = 959) were girls. Students were from grade 7 (age 12-13; n = 658; 34.7%), grade 8 (age 13-14; n = 652; 34.4%), and grade 9 (age 14-15; n = 585; 30.9%). Five students (0.3%) did not indicate their grade level. Most respondents were from English-only speaking homes (n = 1580; 83.4%). The remaining children's families spoke a wide range of languages other than English, with no one language dominant. A small number of students identified as Australian Aboriginal (n = 35; 1.8%) or Torres Strait Islander (n = 14; 0.87); or both (n = 4; 0.2%). One half of the students (n = 874, 50.2%) were from high socio-economic status families (based on a 6-item Family Affluence Scale; Hartley, Levin, & Currie, 2016).

Measures

Global bullying involvement. Participants were provided with a pictorial overview of the definition of bullying emphasizing the intention, repetitive and power-imbalanced nature of bullying; these representations also provided illustrations depicting different forms of bullying (e.g., verbal, relational, physical, property damage) by both offline and online modes, and distinguishing bullying from playful teasing and fights between people equal in power (Shaw, Dooley, Cross, & Zubrick, 2013). Students were then asked, "In the [previous term] at school, how often did you, on your own or in a group, bully (including cyberbully) another young person(s)". Response options were *I did NOT bully another* young person; *Once or twice*; *Every few weeks; About once a week;* and *Several times a week*. Similarly, participants were asked how often in the previous term they were bullied using the same response options.   From this, pure bully and bully/victim status was computed in four categories: uninvolved; pure victim, pure bully; bully/victim.

These bullying roles were examined via liberal and conservative cut-offs to assess the robustness of the role of moral disengagement mechanisms in bullying roles. The liberal cut-offs were computed for those with any reported experience of bullying (i.e., *once or twice* or greater) in the previous school term. However, this risks capturing trivial levels of normative adolescent aggression. The more conservative criterion included only self-reports of victimization or perpetration "every few weeks" or more often. This approach risks misclassifying students who are underreporting their experiences (e.g., their perpetration) as not involved in bullying, and risks reducing the power due to smaller subgroup (i.e., cell) sizes in statistical analyses. Results are reported based on both liberal and conservative cut-offs to examine consistency of findings (see SELF_IDENTIFYING REFERENCE OMITTED).

Moral Disengagement. Participants completed the Moral Disengagement in Bullying Scale (MDBS; Thornberg & Jungert, 2014), an 18-item self-report scale which assesses the extent to which young people

endorse distinct mechanisms of moral disengagement with regard to peer victimization. Each item is rated on a seven-point scale from 1 ("strongly disagree") to 5 ("strongly agree").

The MDBS is comprised of several subscales. To reduce respondent burden, based on the findings of Thornberg and Jungert (2014), only five subscales were included, resulting in a 13-item reduced version of the MDBS. The *Moral Justification* subscale included two items: "it's okay to harm another person a couple of times a week if you do that to protect your friends," and "It's okay to hurt a person a couple of times a week if you do that in order to help your friends"; split-half reliability coefficient for the current sample = 0.81. The *Distorting Consequences* subscale involved four items including "Teasing a person a couple times a week is no big deal because you don't really hurt the person" and "People who get teased don't really get too sad about it"; Cronbach's alpha = 0.82. *Diffusion of Responsibility* was measured using two items: "If my friends begin to bully a classmate, I can't be blamed for being with them and bullying that person too" and "A student can't help that he/she is bullying another student if all his/her friends are doing it"; split-half reliability coefficient = 0.58. *Victim Blame* was assessed via 3 items including "If people are weird, it is their own fault if they get bullied", and "If you can't be like everybody else, you have to blame yourself if you get bullied"; Cronbach's alpha = 0.82. Finally, *Euphemistic labelling* was measured with 2 items: "There's nothing wrong with name-calling a person a bit every day as long as you just do it as a joke", and "Saying mean things to a certain person a couple of times a week doesn't matter. It's just about joking a little with the person"; split-half reliability coefficient = 0.77. The MDBS was validated on school children and has been shown to differentiate between bullies and uninvolved students and between defending students and uninvolved students (Thornberg & Jungert, 2014).

The expected five-factor structure was tested via confirmatory factor analysis (CFA) in MPlus. Data were treated as multinomial categorical using the WLSMV estimator. CFA results indicated the five-factor model provided good fit for the data (CFI = .99; TLI = .99; RMSEA = .018, 90% CI = 0.014 − 0.023).

Analytic Plan

Using the global bullying items, membership in uninvolved, pure victim, pure bully and bully/victim groups was calculated using both the liberal and conservative cut-offs. Structural equation models were conducted in Mplus, testing for differences in levels of moral disengagement (as latent variables) based on membership in the pure bully, bully/victim, pure victim and uninvolved categories. SEM models were applied to appropriately model the moral disengagement scale and subscales as correlated latent variables. Firstly, a second-order SEM model was fitted with moral disengagement as the second-order factor and the five forms of moral disengagement as first-order factors, with the second-order factor predicted by bully/victim group (Gini,

Pozzoli & Bussey, 2015). Secondly, to test hypotheses regarding specific mechanisms of moral disengagement on which pure bully and bully/victim groups differed, a five-factor CFA model for the subscales was fitted, allowing the factors to co-vary and with each predicted by bully/victim group and the demographic variables. The items were treated as categorical and WLSMV estimation was used, with nestedness within schools accounted for via robust standard error estimation with the COMPLEX option.

To fully interrogate the data, different reference groups were systematically varied. First, pure victim, bully/victim and pure bully involvement was examined, with *uninvolved* students as the reference group. Second, pure victim was set as the reference group to compare against bully/victim and pure bully status. Finally, pure bully was set as the reference group to compare against bully/victim status. Several demographic covariates were included. First, as gender has been shown to predict overall moral disengagement, with boys scoring higher than girls on average (Gini et al., 2015), gender was controlled for. Second, socioeconomic status has shown a small but significant association with moral disengagement (Bao, Zhang, Lai, Sun, & Wang, 2015), thus it was also entered as a covariate. Parental consent process (opt-in vs opt-out parental consent) was also controlled for. There is no presumption of prediction or causality in this modeling given cross-sectional data cannot speak to causality.

Results

Bullying Roles

In total 1,870 students completed both the bullying victimization and perpetration questions and could be classified in bully/victim groups. Using the liberal cut-off, a half reported no involvement in bullying as either a victim or perpetrator (927; 49.6%); 690 (36.9%) reported they had been bullied, but had not bullied others; 77 (4.1%) indicated they had bullied others but had not been bullied themselves; and 176 (9.4%) reported both victimization and perpetration at least once in the last school term. Using the more conservative cut-off (i.e., at least "every few weeks"), the great majority (n = 1538; 82.2%) of students were classified as uninvolved; 295 (15.8%) students were classified as pure victims; 20 (1.1%) as pure bullies; and 17 (0.9%) as bully/victims.

In the interests of understanding the bullying role distributions, demographic covariates were examined. Examination of the conservative cut-offs resulted in cell sizes that were too small to enable testing, so only results for the liberal cut-offs are presented. There were significant differences between bullying behaviors of boys and girls (contingency coefficient = .15, $p < .001$). A majority of boys were uninvolved (54.5%); whereas only 45.3% of girls were uninvolved ($p < .05$). Girls were significantly more likely to be pure victims

(boys: 30.4%; girls: 42.6%, $p < .05$). Boys were significantly more likely to be pure bullies (6.0% vs 2.3% for girls; $p < .05$). Boys and girls were equally likely to be bully/victims (9.0% vs 9.8%, respectively, *n.s.*). Bully role was not dependent on grade level (contingency coefficient = .06, *n.s.*). Children with languages other than English were more likely to report being uninvolved in bullying (contingency coefficient = .08, $p < .05$).

Next, mechanisms of moral disengagement scores for the whole sample were examined. On average, the young people expressed disagreement with moral disengagement statements, reflected in overall low mean scores (on the scale of one to five). The mean values were similar across the following five subscales: euphemistic labelling ($M = 2.00$; $SD = 1.03$), with a mean score corresponding with the "disagree" anchor on the five-point response scale; moral justification ($M = 1.78$; $SD = 0.83$); diffusion of responsibility ($M = 1.71$; $SD = 0.87$); distortion of consequences ($M = 1.57$; $SD = 0.74$), and victim blame ($M = 1.35$; $SD = 0.67$). No mean differences in subscales were observed.

Moral Disengagement of Bullies and Bully/Victims

< Please insert Table 1 about here>

Moral disengagement, averaged across mechanisms, was higher on average amongst both pure bullies and bully/victims when compared with students who reported no bullying involvement; while pure victims had similar mean values to uninvolved students (see Table 1). These patterns were observed using both the liberal and conservative involvement criteria and statistically significant when tested (see Table 2 and 3, respectively). With pure victim status as the reference group, both pure bully and bully/victim groups reported significantly greater moral disengagement; this was also replicated across the liberal and conservative involvement cut-offs. Use of liberal and conservative cut-offs did, however, reveal disjunctions in the use of moral disengagement overall between pure bullies and bully/victims. The use of a conservative cut-off resulted in no difference between these bullying categories (see Table 2); the use of a liberal cut-off, on the other hand, revealed that pure bullies espoused more moral disengagement overall than did bully/victims (see Table 3).

<Please insert Tables 2&3 about here>

Mean values for the specific moral disengagement mechanisms were elevated for all students who reported involvement in bullying (see Table 1). The analyses using the conservative thresholds indicated that both pure bullies and bully/victims espoused significantly more moral disengagement on all mechanisms — moral justification, distortion of consequences, diffusion of responsibility, euphemistic labelling and victim blaming — compared to uninvolved students (see Table 2). Compared to the uninvolved students, pure victims

also reported significantly more moral justification for bullying, and more victim blame. Similarly, both pure bullies and bully/victims espoused more moral disengagement on all mechanisms than did pure victims. No significant differences were found on any moral disengagement subscales between pure bullies and bully/victims when deploying the conservative threshold for bullying involvement.

The same analyses were conducted using the liberal thresholds (i.e., including students involved "once or twice" in the previous school term) for bullying perpetration and victimization. The findings were comparable with those using the conservative threshold with regard to the uninvolved students as a reference group: both bully/victims and pure bullies reported more moral disengagement for all five subscales. Again, the pure victims reported greater moral justification and victim blame than the uninvolved students. Also, when contrasted with the pure victim group, both bully/victim and pure bully groups had greater overall moral disengagement and higher scores on all five subscales. Use of the liberal threshold revealed significant differences between bully/victims and pure bullies on specific moral disengagement mechanisms, with significantly less distortion of consequences, less diffusion of responsibility, and less euphemistic labelling amongst the bully/victim group compared to the pure bullies (see also Table 1 for mean differences by mechanism).

Discussion

Prior research has provided evidence that adolescents who engage in bullying have a greater overall tendency toward moral disengagement than do uninvolved adolescents (e.g., Gini, Pozzoli, & Hymel, 2014), and that some mechanisms (e.g., victim blame; moral justification) appear more pertinent to bullying than others (Thornberg & Jungert, 2014). While the overall moral disengagement of bully/victims has been found to be comparable to that of pure bullies (e.g., Obermann, 2011; Tanrikulu & Campbell, 2015), no previous study has examined specific mechanisms of moral disengagement, and whether differences in bully/victims and pure bullies were present in specific mechanisms. Furthermore, some prior research has indicated that moral disengagement may be differentially implicated in bullying for young people who are less extensively involved in bullying (Hymel et al., 2005). In the current study, bullying and victimization roles (pure bully, bully/victim, pure victim, uninvolved) were examined in relation to specific moral disengagement mechanisms, with a comparison of analyses using a more liberal cut-off for involvement compared to a conservative cut-off,. These analyses enabled a robust assessment of the role of moral disengagement in adolescent bullying.

Moral Disengagement, Pure Bullies and Bully/Victims

Overall, moral disengagement was common to pure bullies and to bully/victims alike when contrasted with young people who self-reported no involvement in bullying or who reported only victimization. When including only those students who self-reported extensive involvement in bullying, moral disengagement was comparable between bully/victims and pure bullies. But when students who reported moderate involvement in bullying were included, pure bullies were more likely to score higher in moral disengagement than were bully/victims. These findings replicate those of Hymel et al. (2005): when pure bullying and bully/victim statsu were based on extensive bullying involvement, no differences in moral disengagement were observed. But when young people who were less extensively involved in bullying were examined, pure bullies reported more moral disengagement than did bully/victims. Although the current findings, based on cross-sectional data, cannot speak to the causal role of moral disengagement in enabling bullying perpetration , but it is possible that moderate bulling involvement may arise from processes of moral disengagement. Those mechanisms may be relatively less important in moderate bullying perpetration for young people who are themselves also victims of bullying. (Of course, this cross-sectional study cannot rule out the hypothesis that bullying is causal in the development of moral disengagement.) Further prospective study is required to unravel the likely causality of these relationships.

Support for the specific hypotheses was mixed. First, no differences between pure bullies and bully/victims was observed using the conservative cut-off for any of the mechanisms. The finding that pure bullies either had equivalent (if using the conservative cut-off) or greater (if using the liberal cut-off) tendencies toward victim-blame and distortion of consequences, suggests there are no substantive differences in reliance on moral disengagement amongst students most involved in bullying perpetration. Further research is needed to directly address how bully/victims can justify their bullying perpetration in light of their own histories of victimization, and whether mechanisms of cognitive dissonance are implicated.

The hypotheses about divergence on specific mechanisms did receive support when analyses included adolescents less extensively involved in bullying, with differences observed for three of the five mechanisms. However, the hypotheses that bully/victims would rely more on distortion of consequences and victim blame were not supported. Instead, pure bullies reported significantly greater espousal of distortion of consequences as well as euphemistic labelling. Lacking first-hand experiences of being victimized, it might be easier for pure bullies to perceptually ignore, minimize or distort the harm bullying inflicts on the victim. Also these pure bullies may treat the bullying less seriously by talking about it as a trivial "joking" situation rather than as a bullying situation. Interpreting the actions as "just kidding" might be less convincing if one has some first-hand

experiences as a victim of such actions. Further, the higher level of diffusion of responsibility among pure bullies than bully/victims in the current findings could be related to research showing that, unlike bully/victims who tend to be more socially isolated and less accepted by peers than non-victimized students (Georgiou & Stavrinides, 2008), pure bullying perpetration may have captured group bullying processes, with initiators and assistants, which afforded the opportunity for diffusion of responsibility (Salmivalli et al., 1996). Diffusion of responsibility may also be more applicable to those less active in initiating bullying: the liberal cut-off may have captured more 'henchmen' or 'assistant' participants in bullying: students who do not initiate the bullying but do join in once it has begun. The judgement that they are not the ones responsible may be afforded from this role in particular. Unfortunately, 'ringleader' versus 'henchman' roles were not differentially assessed, and the extent to which the liberal and conservative cut-offs relate to that role distinction is unclear.

Overall, it bears asking why bully/victims who are more extensively involved in bullying have comparable scores on moral disengagement as pure bullies, although those bully/victims who are less involved show differences with pure bullies. It may be that occasional experiences of victimization are inadequate to fuel a process of cognitive dissonance that might apply when only those more extensively involved students are counted (as per our conservative cut-off). Alternately, or in addition, less extensively involved bully/victims may be driven by other psychological processes. Bully/victims appear to be more driven by revenge, rage, and recreational motives for aggression than are pure bullies (Runions et al., 2018). These motives – which capture aggressive tendencies more specifically than does the reactive and proactive dichotomy (Runions, 2013) – may be sufficient to drive moderate levels of bullying perpetration, without a need for evoking moral disengagement mechanisms. Further research is warranted to understand how moral disengagement can arise to enable bullying perpetration amongst those who have experienced the impact of bullying themselves. This likely entails a closer look at how experiences of victimization might afford moral disengagement, and what role such a response to victimization might play in turning victims into bully/victims; this would require prospective assessment of bullying over a substantial period of development.

Moral Disengagement Amongst Pure Victims

Although not the primary focus of the study, the findings for pure victims bear some consideration. Regardless of whether the more conservative or the more liberal threshold is used, pure victims were found to be more inclined than non-involved peers to express moral justification and victim blame for bullying. At first, these findings might seem surprising but if bullying is considered as a social influence and learning situation for the victim, this may be reasonable. Ethnographic studies of school bullying cases and qualitative interview

studies with children and adolescents indicate that in bullying episodes, participants often use dehumanizing and oddness-related labels in an effort to explain and justify their bullying towards the victim (e.g., Thornberg, 2018). This suggests that bullying events may function as socialization or learning situations in which the victims internalize the collectively shared and verbally expressed moral disengagement mechanisms of moral justification, dehumanization and victim blame attributed to them by bullies. These justifications by the bullies for their bullying behavior provide the basis for victims' interpretation of why bullying happens to them and forms part of their growing self-image (Thornberg, 2018). In particular, in interviews conducted by Thornberg Halldin, Bolmsjö, and Petersson (2013), former victims of school bullying reported how they started to develop a sense of not fitting in and a self-blame (i.e., internalizing dehumanizing and oddness-related labeling into self-labeling, and victim blaming into self-blaming) as a result of being bullied. The positive link between being a victim of bullying and self-blaming has been found in other studies (Schacter, White, Chang, & Juvonen, 2015) as well as a negative link between victimization and self-efficacy for avoiding self-blame (Singh & Bussey, 2011). Together, these findings might help to explain why pure victims scored higher than uninvolved peers in moral justification and victim blame in the present study; further longitudinal research will disentangle the causative link between these variables.

Limitations

As with all cross-sectional research, the findings herein are causally ambiguous, and little valid inference can be made about the causal role either developmentally or in vivo of the moral disengagement mechanisms and the purported aggression motives. Longitudinal research will help to provide causal understandings, and intervention research—to test potential causal mechanisms—is recommended to test hypotheses about causal processes. In this regard, it is important to consider that bully/victims may not 'begin' as victims of bullying, but may start as perpetrators who end up bullied themselves.

Students who are neither self-reported victims nor bullies have been referred to in this study as "uninvolved". However, the social dynamics of bullying are such that these students may well be involved in bullying as enablers of the bully, defenders of the victim, or as witnesses (Salmivalli et al., 1996). The analyses aimed to examine the factors contributing to explicit perpetation of bullying. Further research across a wider range of roles is warranted.

The use of a global measure of bullying also ignores potentially important distinctions in the use of moral disengagement. Bullies and bully/victims who engage in physical bullying may demonstrate different moral disengagement patterns compared to those who bully via relational aggression or verbal aggression,

especially by adolescence when physical bullying is less common (Yeager et al., 2015). Similarly, cyberbullying may differentially evoke – or even facilitate – moral disengagement in ways that face-to-face encounters do not (Runions & Bak, 2015). Further inquiry examining diverse forms of bullying – possibly using person-centered analyses (e.g., latent class analysis)—may shed light on such factors.

Self-report was used for all measures in the analyses, which is therefore vulnerable to the risk of inflated associations between variables as a function of shared method variance. Although moral disengagement is too subjective for valid and reliable assessment by other raters, bullying is an objectifiable outcome, and peer or teacher nominations may provide a valuable replication extension. However, in most countries, secondary school teacher nominations are challenging due to the distributed nature of classes, with no single teacher acting as the key contact point who might be expected to be aware of bullying issues. In some jurisdictions, bullying behavior is being classified as illegal, and education departments prohibited inquiry into illegal behaviors, further challenging progress into addressing the problem. Peer reports would provide an index of 'convergent' (i.e., based on both peer- and self-reports) victims, bullies and bully/victims (Scholte, Burk, & Overbeek, 2013). Research that has collected both peer and self-reports have begun to map out important differences in students who are self-reported, peer-reported, or both on these bullying roles. Peer-nominated victims have comparable adjustment outcomes to non-victims, compared to self-identified (or convergently identified) victims; self-nominated victims are more accepted by peers than peer-identified victims (Graham & Juvonen, 1998; Scholte et al., 2013). Self-identified victims do not expect their peers to defend them as much as peer-identified victims (Dawes, Chen, Farmer, & Hamm, 2017). Our findings are limited to self-identified bullying; future research that captures both peer- and self-nominations would provide important insights into moral disengagement for these divergent groups.

Finally, the composition of the sample was problematic. The low consent rate from government-school parents under the opt-in procedure poses a risk to the generalizability of the findings to students within the government sector. This would be of particular concern were school sector a focus of this study, however this study did not seek to draw comparisons between the sectors. Since the under-represented students within the opt-in group are present within the opt-out consent group from non-government schools, they are represented within the broader sample. By controlling for consent procedure in the analyses, this study mitigated any impact from the low consent rate in government schools on the conclusions of the study with regard to the broader student population.

Implications

Bullying is not a homogeneous phenomenon, but arises from distinct psychological enablers. This study provides some insight into how moral disengagement might differentially influence bullying for adolescents who are and are not themselves bullied. Reducing and preventing bullying is likely to require specific strategies that are sensitive to these possibly subtle differences in collective moral disengagement at the classroom level (Gini et al., 2015) and individual moral disengagement risk. By understanding the nuanced moral disengagement mechanisms involved in bullying, direct targeted interventions may be possible to help students find alternative social means to achieve their social goals. Given that a meta-analysis has shown that many anti-bullying interventions for adolescents are ineffective or counterproductive (Yeager et al., 2015), the finding that moral disengagement may be a common mechanism for both pure bullies and bully/victims provides a direction for intervention development. Specifically, these results indicate the need for targeted intervention and prevention strategies to address specific moral disengagement mechanisms overall, but with possible targeted components that address distinct modes of bullying aggression. Addressing pure bullying may require intervention into the potential harm of what they perceive to be harmless jokes. Addressing bully/victim bullying may be more productive if it focused on their own victimization experiences, breaking down moral justification and victim blaming through anger management, conflict resolution, social skills, and motivational interviewing to support students to meet their social goals via other means.

Motivational interviewing may be particularly potent in working with bullies and bully/victims whose bullying arises via moral disengagement. Motivational interviewing is a process of working with people who experience problems that are resistant to change, and that may provide an effective means of working with bullies (Cross, Runions, Resnicow, Britt, & Gray, 2018). The first stages of MI involve working with the client to dismantle "resistance" talk: reasons why the behavior is not "really" a problem. Moral disengagement mechanisms provide fertile cognitive grounds for resistance talk, as they provide internal justifications for why the behavior – in this case bullying - was warranted (e.g., moral justification) or even deserved (e.g., victim blame), or alternately, why those behaviors are not acknowledged as being the responsibility of the young person (diffusion of responsibility). However, implementing motivational interviewing can be challenging in schools (Pennell, Campbell, Tangen, Runions, Brooks, & Cross, 2018) given limits to the capacity to deliver selective interventions.

The moral agency framework developed by Bandura is based in large part on historical examples (e.g., Bandura, 1999), and the mechanisms identified are, in principle, common to world leaders and class clowns alike. For teachers and parents, historical examples provide case studies that are distant enough in time and

space to be discussed safely with young people, with gradual integration of those lessons to the daily lives of the students themselves, and how moral disengagement may arise in their own lives. The *Beyond Bali* curriculum package is an example of integration of the moral disengagement framework into the curriculum (Taylor, Taylor, Karnovsky, Aly, & Taylor, 2017). The integration of moral disengagement as a common framework for the teaching of history, civics and social-emotional competencies, and whether such an integrated framework could address a range of moral and ethical goals, remains unclear.

Finally, bullying researchers are encouraged to measure moral disengagement in their program evaluations even for interventions that do not explicitly aim to change moral disengagement, to determine whether individual and collective moral disengagement (Gini et al., 2015) mediates changes in bullying behavior in schools. Discussions of bullying and victimization are likely to challenge core moral disengagement mechanisms either directly via curriculum or indirectly via peer and other whole-school processes that arise, for example to improve bystander responses. The search for the essential underlying mechanisms in behavioral change away from bullying remains a core goal for bullying prevention and intervention.

Conclusion

Developmentally, adolescents have internalized the moral frameworks that condemn bullying, but nevertheless some continue to engage in bullying others, even if they themselves are targets of bullying. The current study sought to address several gaps in the literature with regard to the role of moral disengagement mechanisms in adolescent bullying. Specifically, the study examined pure bullies and bully/victims to determine whether the latter are less reliant on moral disengagement, due to their first-hand knowledge of the impact of victimization. These results suggest that moral disengagement plays an equivalent role in bullying perpetration— regardless of the concurrent victimization status—for those most extensively involved. For those bully/victims less extensively involved in bullying, pure bullies espouse more moral disengagement. Given that interventions to reduce bullying seem to lose efficacy in adolescence, these findings point to moral disengagement as a possible target for improving the efficacy of bullying prevention and intervention programs for adolescents.

References

Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review, 3*(3), 193-209. doi:10.1207/s15327957pspr0303_3

Bandura, A., Barbaranelli, C., Caprara, G. V. , & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology, 71*(2), 364-374. doi:10.1037/0022-3514.71.2.364

Bjärehed, M., Thornberg, R., Wänström, L., & Gini, G. (2019). Mechanisms of moral disengagement and their associations with indirect bullying, direct bullying, and pro-aggressive bystander behavior. *Journal of Early Adolescence.* Advance online publication. doi: 10.1177/0272431618824745

Boyce, W., Torsheim, T., Currie, C., & Zambon, A. (2006). The family affluence scale as a measure of national wealth: Validation of an adolescent self-report measure. *Social Indicators Research, 78*(3), 473-487. doi:10.1007/s11205-005-1607-6

Bussey, K., Fitzpatrick, S., & Raman, A. (2015). The role of moral disengagement and self-efficacy in cyberbullying. *Journal of School Violence, 14*(1), 30-46. doi:10.1080/15388220.2014.954045

Craig, W., Harel-Fisch, Y., Fogel-Grinvald, H., Dostaler, S., Hetland, J., Simons-Morton, B., . . . Due, P. (2009). A cross-national profile of bullying and victimization among adolescents in 40 countries. *International Journal of Public Health, 54*, 216-224. doi:10.1007/s00038-009-5413-9

Cross, D. S., Runions, K. C., Resnicow, K. A., Britt, E. F., & Gray, C. (2018). Motivational interviewing as a positive response to high-school bullying. *Psychology in the Schools, 55*(5), 464-475. doi:10.1002/pits.22120

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*(11), 2273-2288. doi:10.1007/s10964-017-0688-2

Garner, P. W., & Hinton, T. S. (2010). Emotional display rules and emotion self-regulation: Associations with bullying and victimization in community-based after school programs. *Journal of Community & Applied Social Psychology, 20*(6), 480-496. doi:10.1002/casp.1057

Georgiou, S. N., & Stavrinides, P. (2008). Bullies, victims and bully/victims psychosocial profiles and attribution styles. *School Psychology International, 29*(5), 574-589. doi:10.1177/0143034308099202

Gini, G., Pozzoli1, T & Hymel, S. (2014). Moral disengagement Among children and youth: A meta-analytic review of links to aggressive behavior. *Aggressive Behavior, 40*, 56–68. doi:10.1002/ab.21502

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*(3), 441-452. doi:10.1007/s10802-014-9920-7

Graham, S., & Juvonen, J. (1998). Self-blame and peer victimization in middle school: An attributional analysis. *Developmental Psychology, 34*(3), 587-599. doi:10.1037/0012-1649.34.3.587

Hartley, J. E. K., Levin, K., & Currie, C. (2016). A new version of the HBSC Family Affluence Scale - FAS III: Scottish qualitative findings from the International FAS Development Study. *Child Indicators Research, 9*(1), 233-245. doi:10.1007/s12187-015-9325-3

Hymel, S., Rocke-Henderson, R. A., & Bonanno, R. A. (2005). Moral disengagement: A framework for understanding bullying among adolescents. *Journal of Social Sciences, 8*, 33-45.

Jackson, D. B., & Vaughn, M. G. (2018). The bully/victim overlap and nutrition among school-aged youth in North America and Europe. *Children and Youth Services Review, 90*, 158-165. doi:10.1016/j.childyouth.2018.05.027

Lahelma, E. (2004). Tolerance and understanding? Students and teachers reflect on differences at school. *Educational Research and Evaluation, 10*, 3-19.

Lereya, S. T., Copeland, W. E., Zammit, S., & Wolke, D. (2015). Bully/victims: A longitudinal, population-based cohort study of their mental health. *European Child & Adolescent Psychiatry, 24*(12), 1461-1471. doi:10.1007/s00787-015-0705-5

Menesini, E., Sanchez, V., Fonzi, A., Ortega, R., Costabile, A., & Lo Feudo, G. (2003). Moral emotions and bullying: A cross-national comparison of differences between bullies, victims and outsiders. *Aggressive Behavior, 29*(6), 515-530. doi:10.1002/ab.10060

Modecki, K. L., Minchin, J., Harbaugh, A. G., Guerra, N. G., & Runions, K. C. (2014). Bullying prevalence across contexts: A meta-analysis measuring cyber and traditional bullying. *Journal of Adolescent Health, 55*(5), 602-611. doi:10.1016/j.jadohealth.2014.06.007

Obermann, M.-L. (2011). Moral disengagement in self-reported and peer-nominated school bullying. *Aggressive Behavior, 37*(2), 133-144. doi:10.1002/ab.20378

Olweus, D. (2013). School bullying: Development and some important challenges. In S. Nolen-Hoeksema (Ed.), *Annual Review of Clinical Psychology, Vol 9* (Vol. 9, pp. 751-780).

Pellegrini, A. D., Bartini, M., & Brooks, F. (1999). School bullies, victims, and aggressive victims: Factors relating to group affiliation and victimization in early adolescence. *Journal of Educational Psychology, 91*(2), 216-224. doi:10.1037//0022-0663.91.2.216

Pennell, D., Campbell, M.A., Tangen, D., Runions, K., Brooks, J., & Cross D.S. (2018). Facilitators and barriers to the implementation of motivational interviewing (MI) for bullying perpetration in school settings. *Scandinavian Journal of Psychology*. DOI: 10.1111/sjop.12502

Perren, S., Gutzwiller-Helfenfinger, E., Malti, T., & Hymel, S. (2012). Moral reasoning and emotion attributions of adolescent bullies, victims, and bully/victims. *British Journal of Developmental Psychology, 30*(4), 511-530. doi:10.1111/j.2044-835X.2011.02059.x

Pouwels, J. L., Lansu, T.A.M., & Cillessen, A.H.N. (2018). A developmental perspective on populatiry and the group process of bullying. *Aggression and Violent Behavior, 43*, 64-70. https://doi.org/10.1016/j.avb.2018.10.003

Pouwels, J. L., Scholte, R. H. J., van Noorden, T. H. J., & Cillessen, A. H. N. (2016). Interpretations of bullying by bullies, victims, and bully/victims in interactions at different levels of abstraction. *Aggressive Behavior, 42*(1), 54-65. https://doi:10.1002/ab.21605

Pozzoli, T., Gini, G., & Vieno, A. (2012). Individual and class moral disengagement in bullying among elementary school children. *Aggressive Behavior, 38*(5), 378-388. doi:10.1002/ab.21442

Runions, K. C. (2013). Toward a conceptual model of motive and self-control in cyber-aggression: Rage, revenge, reward, and recreation. *Journal of Youth and Adolescence, 42*(5), 751-771. doi:10.1007/s10964-013-9936-2

Runions, K. C., & Bak, M. (2015). Online moral disengagement, cyberbullying, and cyber-aggression. *Cyberpsychology Behavior and Social Networking, 18*(7), 400-405. doi:10.1089/cyber.2014.0670

Runions, K. C., Salmivalli, C., Shaw, T., Burns, S., & Cross, D. (2018). Beyond the reactive-proactive dichotomy: Rage, revenge, reward, and recreational aggression predict early high school bully and bully/victim status. *Aggressive Behavior, 44*(5), 501-511. doi:10.1002/ab.21770

Salmivalli, C., Lagerspetz, K., Bjorkqvist, K., Osterman, K., & Kaukiainen, A. (1996). Bullying as a group process: Participant roles and their relations to social status within the group. *Aggressive Behavior, 22*(1), 1-15. doi:10.1002/(sici)1098-2337(1996)22:1<1::aid-ab1>3.0.co;2-t

Sampasa-Kanyinga, H., Chaput, J. P., Hamilton, H. A., & Colman, I. (2018). Bullying involvement, psychological distress, and short sleep duration among adolescents. *Social Psychiatry and Psychiatric Epidemiology, 53*(12), 1371-1380. doi:10.1007/s00127-018-1590-2

Schacter, H. L., White, S. J., Chang, V. Y., & Juvonen, J. (2015). "Why Me?": Characterological self-blame and continued victimization in the first year of middle school. *Journal of Clinical Child and Adolescent Psychology, 44*(3), 446-455. doi:10.1080/15374416.2013.865194

Scholte, R. H. J., 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*(12), 1789-1800. doi:10.1007/s10964-012-9896-y

Schwartz, D. (2000). Subtypes of victims and aggressors in children's peer groups. *Journal of Abnormal Child Psychology, 28*(2), 181-192. doi:10.1023/a:1005174831561

Shaw, T., Dooley, J. J., Cross, D., & Zubrick, S. R. (2013). The Forms of Bullying Scale (FBS): Validity and reliability estimates for a measure of bullying victimization and perpetration in adolescence. *Psychological Assessment, 25*(4), 1045-1057. doi:10.1037/a0032955

Sigurdson, J. F., Undheim, A. M., & Sund, A. M. (2015). Bulling involvement in early adolescence years: Impact on mental health problems, psychosocial adjustment and use of services in adulthood. *European Child & Adolescent Psychiatry, 24*, S28-S28.

Singh, P., & Bussey, K. (2011). Peer victimization and psychological maladjustment: The mediating role of coping self-efficacy. *Journal of Research on Adolescence, 21*(2), 420-433. doi:10.1111/j.1532-7795.2010.00680.x

Solberg, M. E., Olweus, D., & Endresen, I. M. (2007). Bullies and victims at school: Are they the same pupils? *British Journal of Educational Psychology, 77*, 441-464. doi:10.1348/000709906x105689

Solberg, M. E., & Olweus, D. (2003). Prevalence estimation of school bullying with the Olweus Bully Victim Questionnaire. *Aggressive Behavior, 29*(3), 239-268. doi:10.1002/ab.10047

South, C. R., & Wood, J. (2006). Bullying in prisons: the importance of perceived social status, prisonization, and moral disengagement. *Aggressive Behavior, 32*(5), 490-501. doi:10.1002/ab.20149

Tanrikulu, I., & Campbell, M. (2015). Correlates of traditional bullying and cyberbullying perpetration among Australian students. *Children and Youth Services Review, 55*, 138-146. doi:10.1016/j.childyouth.2015.06.001

Taylor, E., Taylor, P. C., Karnovsky, S., Aly, A., & Taylor, N. (2017). "Beyond Bali": A transformative education approach for developing community resilience to violent extremism. *Asia Pacific Journal of Education, 37*(2), 193-204. doi:10.1080/02188791.2016.1240661

Thornberg, R. (2018). School bullying and fitting into the peer landscape: A grounded theory field study. *British Journal of Sociology of Education, 39*(1), 144-158. doi:10.1080/01425692.2017.1330680

Thornberg, R., Halldin, K., Bolmsjö, N., & Petersson, A. (2013). Victimising of school bullying: A grounded theory. *Research Papers in Education, 28*(3), 309-329. doi:10.1080/02671522.2011.641999

Thornberg, R., & Jungert, T. (2014). School bullying and the mechanisms of moral disengagement. *Aggressive Behavior, 40*(2), 99-108. doi:10.1002/ab.21509

Thornberg, R., Thornberg, U. B., Alamaa, R., & Daud, N. (2016). Children's conceptions of bullying and repeated conventional transgressions: moral, conventional, structuring and personal-choice reasoning. *Educational Psychology, 36*(1), 95-111. doi:10.1080/01443410.2014.915929

Toblin, R. L., Schwartz, D., Gorman, A. H., & Abou-Ezzeddine, T. (2005). Social-cognitive and behavioral attributes of aggressive victims of bullying. *Journal of Applied Developmental Psychology, 26*(3), 329-346. doi:10.1016/j.appdev.2005.02.004

Verlinden, M., Veenstra, R., Ghassabian, A., Jansen, P. W., Hofman, A., Jaddoe, V. W. V., . . . Tiemeier, H. (2014). Executive functioning and non-verbal intelligence as predictors of bullying in early elementary school. *Journal of Abnormal Child Psychology, 42*(6), 953-966. doi:10.1007/s10802-013-9832-y

Wolke, D., Lereya, S. T., Fisher, H. L., Lewis, G., & Zammit, S. (2014). Bullying in elementary school and psychotic experiences at 18 years: A longitudinal, population-based cohort study. *Psychological Medicine, 44*(10), 2199-2211. doi:10.1017/s0033291713002912

Yang, A., & Salmivalli, C. (2013). Different forms of bullying and victimization: Bully/victims versus bullies and victims. *European Journal of Developmental Psychology, 10*(6), 723-738. doi:10.1080/17405629.2013.793596

Yeager, D. S., Fong, C. J., Lee, H. Y., & Espelage, D. L. (2015). Declines in efficacy of anti-bullying programs among older adolescents: Theory and a three-level meta-analysis. *Journal of Applied Developmental Psychology, 37*, 36-51. doi:10.1016/j.appdev.2014.11.005
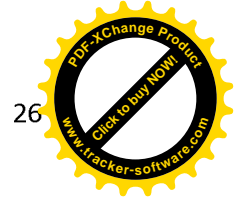
Table 1. *Mean (SD) Scores on Moral Disengagement (MD) Mechanisms by Bully/Victim Status, by Liberal and Conservative Cut-offs for Bullying Involvement*

	Uninvolved	Pure Victim	Bully/Victim	Pure Bully
MD Mechanism	Liberal cut-off ("once or twice" in the previous school term)			
Justification	1.67 (0.87)	1.74 (0.86)	2.14 (1.09)	2.40 (1.03)
Euphemism	1.94 (1.02)	1.91 (0.94)	2.29 (1.07)	2.78 (1.12)
Diffusion	1.65 (0.84)	1.66 (0.80)	1.92 (0.98)	2.29 (0.95)
Distortion	1.52 (0.69)	1.48 (0.62)	1.80 (0.87)	2.06 (0.86)
Blame	1.27 (0.59)	1.29 (0.56)	1.55 (0.84)	1.69 (0.82)
Overall MD	1.55 (0.63)	1.55 (0.57)	1.87 (0.81)	2.16 (0.76)
	Conservative cut-off ("every few weeks" or more in the previous school term)			
Justification	1.73 (0.87)	1.84 (0.96)	3.00 (1.45)	2.88 (1.30)
Euphemism	1.98 (1.01)	1.89 (0.93)	3.00 (1.29)	2.98 (1.29)
Diffusion	1.68 (0.84)	1.70 (0.81)	2.72 (1.18)	2.47 (1.09)
Distortion	1.54 (0.69)	1.48 (0.64)	2.55 (1.34)	2.49 (1.00)
Blame	1.29 (0.59)	1.35 (0.62)	2.12 (1.38)	2.20 (1.18)
Overall MD	1.59 (0.62)	1.58 (0.60)	2.63 (1.20)	2.56 (1.01)

MORAL DISENGAGEMENT IN PURE BULLIES AND BULLY/VICTIMS

Table 2. *Structural Equation Model of Bullying Perpetration and Victimization as Predictors of Moral Disengagement Mechanisms: Conservative Bullying Participation*

Cut-Offs (n = 1802)

	Moral Justification			Distortion of Consequences			Diffusion of Responsibility			Euphemistic Labelling			Victim Blame			Overall Moral Disengagement		
	B	SE	p	B	SE	p	B	SE	p	B	SE	p	B	SE	p	B	SE	p
Reference: Uninvolved																		
Victim	0.14	0.07	.043	-0.06	0.10	.540	0.05	0.08	.490	-0.08	0.08	.317	0.18	0.08	.027	0.05	0.07	.468
B / V	1.14	0.19	.000	1.07	0.22	.000	1.20	0.26	.000	1.04	0.33	.002	1.06	0.33	.001	1.18	0.25	.000
Bully	1.16	0.14	.000	1.23	0.22	.000	1.09	0.28	.000	0.96	0.15	.000	1.23	0.19	.000	1.27	0.16	.000
Reference: Pure Victim																		
B / V	0.99	0.20	.000	1.15	0.26	.000	1.12	0.26	.000	1.08	0.36	.003	0.89	0.37	.017	1.22	0.19	.000
Bully	1.01	0.17	.000	1.28	0.23	.000	1.04	0.31	.001	1.07	0.18	.000	1.08	0.22	.000	1.15	0.27	.000
Reference: Pure Bully																		
B / V	0.00	0.28	.999	-0.11	0.35	.761	0.22	0.53	.674	0.00	0.40	.992	-0.12	0.47	.806	-0.04	0.37	.470

NB. All models adjusted for gender, grade level, consent status

MORAL DISENGAGEMENT IN PURE BULLIES AND BULLY/VICTIMS

Table 3. *Structural Equation Model of Bullying Perpetration and Victimization as Predictors of Moral Disengagement Mechanisms: Liberal Bullying Participation Cut-Offs*

(n = 1802)

	Moral Justification			Distortion of Consequences			Diffusion of Responsibility			Euphemistic Labelling			Victim Blame			Overall Moral Disengagement		
	B	SE	p	B	SE	p	B	SE	p	B	SE	p	B	SE	p	B	SE	p
Reference: Uninvolved																		
Victim	0.15	0.06	.008	-0.03	0.03	.299	0.07	0.05	.176	0.01	0.04	.716	0.11	0.05	.044	0.03	0.03	.251
B / V	0.57	0.09	.000	0.47	0.07	.000	0.45	0.12	.000	0.44	0.07	.002	0.56	0.09	.001	0.54	0.08	.000
Bully	0.75	0.11	.000	0.76	0.09	.000	0.87	0.14	.000	0.82	0.06	.000	0.65	0.11	.000	0.85	0.07	.000
Reference: Pure Victim																		
B / V	0.42	0.07	.000	0.51	0.09	.000	0.40	0.10	.000	0.41	0.06	.000	0.44	0.10	.000	0.79	0.09	.000
Bully	0.60	0.10	.000	0.78	0.10	.000	0.83	0.12	.000	0.81	0.08	.000	0.56	0.13	.000	0.49	0.08	.000
Reference: Pure Bully																		
B / V	-0.18	0.13	.173	-0.28	0.06	.006	-0.40	0.14	.006	-0.40	0.08	.000	-0.11	0.16	.494	-0.33	0.08	.000

NB. All models adjusted for gender, grade level, and consent status

Acknowledgements

Author's Contributions

KR participated in the design and coordination of the study and drafted the manuscript. TS participated in in the design and coordination of the study, performed the statistical analysis and helped to draft the manuscript. KB, RT, and CS helped to draft the manuscript. DC conceived of the study, participated in its design and coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

Data Sharing Declaration

This manuscript's data will not be deposited, per request by the relevant Department of Education.

Conflicts of Interest

The authors report no conflict of interests.

Compliance with Ethical Standards

Funding

This study was conducted as part of the *Beyond Bullying: Positive Change for All* study, which was funded in whole by the Australian National Health and Medical Research Council (NHMRC Project 1046086). DC's contribution to this article was supported by NHMRC Research Fellowship GNT 1119339.

Ethical Approval

The study was reviewed and approved by the University of Western Australia Human Ethics review board (RA/4/1/6934), and by the Department of Education of Western Australia, and Catholic Education Western Australia. The research was conducted in compliance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments, and in compliance with the Australian National Statement on Ethical Conduct in Human Research (2007).

Informed Consent

All participants were provided with information documentation and consent was required and was obtained for all participants in the study.