Social Psychology of Education (2025) 28:196
https://doi.org/10.1007/s11218-025-10156-z

Moral disengagement and defender self-efficacy and their associations with defending, unconcerned bystanding, and guilty bystanding in school bullying

Robert Thornberg[1] · **Björn Sjögren**[1] · **Gianluca Gini**[2] · **Tiziana Pozzoli**[2]

Received: 20 March 2025 / Accepted: 29 October 2025

Abstract
The aim of the present study was to investigate whether moral disengagement and defender self-efficacy were related to defending, unconcerned bystanding, and guilty bystanding in school bullying among pre-adolescent students. In this study, 1168 pre-adolescent students from 74 upper elementary classrooms in 31 schools in Sweden completed a series of self-report scales. Through structural equation modeling, the study found that defending was positively associated with defender self-efficacy and negatively associated with moral disengagement. Unconcerned bystanding was negatively associated with defender self-efficacy and positively associated with moral disengagement. Finally, guilty bystanding was negatively associated with defender self-efficacy and moral disengagement.

Keywords Bullying · Bystander · Moral disengagement · Defender self-efficacy

1 Introduction

In the context of school bullying, a *bystander* refers to any student who witnesses a bullying incident (Polanin et al., 2012). Although school bullying typically occurs in the presence of bystanders, these witnesses rarely intervene to help the victim (Craig et al., 2000; Hawkins et al., 2001). This study focuses on two specific types of bystander behavior; namely defending and passive bystanding. *Defending* refers to bystanders siding with victims. While defending can take many forms such as

✉ Robert Thornberg
robert.thornberg@liu.se

[1] Department of Behavioral Sciences and Learning, Linköping University, Linköping SE-58183, Sweden

[2] Department of Developmental Psychology and Socialization, University of Padua, Padua, Italy

Published online: 11 November 2025

confronting the bullies, reporting to authority, and comforting the victims afterwards (Garandeau et al., 2022; Lambe & Craig, 2020; Steinvik et al., 2025; Wang et al., 2023), in the current study, we delimited our focus on active defending operationalized as reporting to a teacher, telling the bullies to stop, and stepping between and trying to make them stop. *Passive bystanding*, on the other hand, involves bystanders refraining from involvement, avoiding engagement in the bullying situation, and remaining passive (Campbell et al., 2023; Gini et al., 2008). Obermann (2011) proposes a distinction between two types of passive bystanders: *unconcerned bystanders* (those who passively witness the bullying without feeling troubled and responsible), and *guilty bystanders* (those who witness the bullying and believe that they should help the victim, but fail to take action). With the exception of Obermann's (2011) study, these two forms of passive bystanding have not yet been investigated in the school bullying literature.

1.1 Social cognitive theory

Children judge peer aggression as morally wrong from early preschool years, often citing the harm or unfairness caused by these actions (Yoo & Smetana, 2022). However, according to social cognitive theory (Bandura, 1999, 2016), moral conceptions alone are insufficient to produce moral behavior. For moral agency to occur—such as defending a victim of school bullying—children must develop motivational and self-regulatory processes that translate moral understanding into action (Brüggemann et al., 2019; Bussey, 2023; Lambe et al., 2019).

1.1.1 Moral disengagement

Social cognitive theory (Bandura, 1999, 2016) introduces *moral disengagement* to explain the discrepancy between moral standards and behavior. It encompasses self-serving cognitive distortions that enable individuals to perceive immoral acts is as acceptable and thus not conflicting with moral standards. Moral disengagement includes eight mechanisms: moral justification, euphemistic labeling, advantageous comparison, displacement of responsibility, diffusion of responsibility, ignoring the harmful effects, dehumanization, and blaming the victim. Through these social cognitive processes, individuals may engage in antisocial behaviors (e.g., bullying) or refrain from prosocial actions (e.g., remaining a passive bystander instead of defending a victim of bullying) without evoking feelings of guilt or remorse. When such mechanisms are activated, immoral behaviors—whether enacted or witnessed—are interpreted as justified or acceptable.

Consistent with social cognitive theory, higher levels of moral disengagement have been linked to lower levels of defending (Campbell et al., 2023; Eijigu & Teketel, 2023; Gini et al., 2020, 2022; Jiang et al., 2022; Pozzoli et al., 2016; Shi et al., 2024; Sjögren et al., 2021a; Thornberg et al., 2022, 2023; Tolmatcheff et al., 2022a; Yang & Gao, 2023; although only marginal in Caravita et al., 2022 and Sjögren et al., 2024; exceptions in Sjögren et al., 2021b; Thornberg et al., 2020). This negative relationship has been supported in three meta-analyses, though with small effect size (Killer et al., 2019; Luo & Bussey, 2023; Ma et al., 2019).

Fewer studies have examined the link between moral disengagement and passive bystanding, and results have been more inconsistent. While some studies have reported a negative association (Caravita et al., 2019; Gini, 2006; Thornberg & Jungert, 2013; Shi et al., 2024), others have found a positive relationship (Eijigu & Teketel, 2023; Gini et al., 2015, 2020; Jiang et al., 2022; Sjögren et al., 2021a, b; Shi et al., 2024; Thornberg et al., 2017, 2022, 2023; Tolmatcheff et al., 2022a; Wang et al., 2024; Yang & Gao, 2023), and yet some have identified no significant association at all (Caravita et al., 2019, 2022; Mazzone et al., 2016). For example, Shi et al. (2024) found a weak but significant negative correlation between moral disengagement and passive bystanding. However, when this relationship was examined within a more complex model that included social support and classroom collective efficacy, the association shifted to weak yet significant positive association. In Caravita et al.'s (2019) study, a significant but weak negative correlation between moral disengagement and passive bystanding was observed among non-immigrant youth, whereas this relationship was non-significant among immigrant youth. Killer et al.'s (2019) meta-analysis found no significant relationship, whereas Luo and Bussey (2023) identified a small but positive relationship, moderated by age and context, in their meta-analysis. Thus, in addition to potential measurement differences contributing to inconsistencies across studies, Killer et al.'s (2019) reported that the association was not significant in American and European samples but was evident in Australian and Asian samples. Furthermore, positive correlations were somewhat stronger in samples exclusively comprising adolescents compared to those including both children and adolescents, whereas correlations became weakly negative when samples included both adolescents and adults. Considering that passive bystanders may act from different intentions, motives, and cognitions, such individual differences could further help explain the heterogeneity observed across studies.

Obermann's (2011) study might shed some light on these mixed findings by distinguishing between two subgroups of passive bystanders: unconcerned and guilty bystanders. The former scored significantly higher in moral disengagement suggesting that passive bystanders are not a homogenous group. The inconsistency in the literature may partly reflect whether passive bystanders are unconcerned about what is happening (greater moral disengagement) or feel guilty about their passivity (less moral disengagement). The continued passivity of guilty bystanders, despite low moral disengagement and awareness of victims' suffering, may be better explained through a cost–reward analysis (Dovidio et al., 2006). Such individuals may remain passive due to uncertainty about how to act, or fears of embarrassment, peer rejection, retaliation, or becoming targets themselves (Strindberg et al., 2020). These concerns likely reflect low self-efficacy in intervening.

1.1.2 Defender self-efficacy

According to social cognitive theory (Bandura, 1997), individuals' motivation to engage in a specific behavior depends on their self-efficacy in enacting that behavior. *Self-efficacy* is defined as the beliefs an individual has in their capabilities to successfully organize and execute the lines of action required to achieve a desired outcome. High self-efficacy motivates individuals to engage in the behavior if it aligns with

their personal goals, while low self-efficacy demotivates and inhibits the behavior (Bandura, 1997).

Social cognitive theory (Bandura, 1997) emphasizes that self-efficacy is a domain-specific construct and should be delimited to and examined in relation to the specific activity or behavior. Accordingly, this study investigated how students' defender self-efficacy relates to their different bystander behaviors in school bullying. *Defender self-efficacy* is defined as the belief in one's ability to successfully intervene in bullying or peer aggression to defend a victim (Thornberg et al., 2017). Most research on defender self-efficacy and bystander behaviors in bullying and peer aggression has focused on defending. These studies consistently demonstrate that students with high defender self-efficacy are more inclined to defend victims (Eijigu & Teketel, 2023; Gini et al., 2022; Peets et al., 2015; Pöyhönen et al., 2012; Sjögren et al., 2024b; Sjögren et al., 2020, 2021b, 2024a, b; Thornberg et al., 2017, 2020; van der Ploeg et al., 2017). The positive association between defender self-efficacy and defending has also been confirmed in a meta-analysis (Ma et al., 2019).

Fewer studies have examined the link between defender self-efficacy and passive bystanding. The available studies suggest that students with higher levels of defender self-efficacy are less likely to act as passive bystanders (Eijigu & Teketel, 2023; Sjögren et al., 2020, 2021b, 2024b; Thornberg et al., 2017; for an exception, see Pöyhönen et al., 2012). However, previous research has not distinguished between unconcerned and guilty bystanding when investigated the link between defender self-efficacy and passive bystanding. Therefore, clarifying the extent to which defender self-efficacy is related to unconcerned bystanding and guilty bystanding is critical to better understand the complexity of passive bystanding.

1.2 The present study

The aim of the present study was to investigate whether moral disengagement and defender self-efficacy were related to defending, unconcerned bystanding, and guilty bystanding in school bullying among pre-adolescent students in upper elementary schools. Based on the literature, we hypothesized that moral disengagement would be positively associated with unconcerned bystanding and negatively associated with both defending and guilty bystanding (Hypothesis 1). Furthermore, we hypothesized that defender self-efficacy would be positively associated with defending (Hypothesis 2). To the best of our knowledge, this study is the first to examine whether defender self-efficacy is linked to the two forms of passive bystanding. Considering that passive bystanding as a general construct has been linked to less defender self-efficacy, and that guilty bystanders think that they should intervene while remaining passive, it is plausible to hypothesize that defender self-efficacy is negatively associated with guilty bystanding (Hypothesis 3). Due to the novelty of this study and lack of previous empirical research on the possible link between defender self-efficacy and unconcerned bystanding, we could not derive a specific hypothesis from the literature. Instead, we investigated the relationship between defender self-efficacy and unconcerned bystanding in an exploratory manner. Finally, recent findings suggest that moral disengagement may alter how defender self-efficacy translates into bystander behavior. According to social-cognitive theory, moral and self-evaluative

mechanisms interact in guiding social action, and when moral self-regulation is deactivated, self-efficacy beliefs may exert weaker influence on behavior. Empirically, Sjögren et al. (2024a, b) found that defender self-efficacy predicted lower passive bystanding only among students low in moral disengagement, suggesting that moral disengagement can attenuate the behavioral expression of self-efficacy. Building on this reasoning, we also tested, as an exploratory extension, the interaction between moral disengagement and defender self-efficacy in predicting defending, guilty bystanding, and unconcerned bystanding.

Age and gender were included as control variables in the present study. Empirical findings generally indicate that defending decreases with age (Eijigu & Teketel, 2023; Gini et al., 2022; Pöyhönen et al., 2012; Sjögren et al., 2021b; Thornberg et al., 2023; for exceptions, see Mazzone et al., 2016; Pozzoli et al., 2016; Tolmatcheff et al., 2022). According to a meta-analysis by Ma et al. (2019), younger students are more likely than older students to defend, although the average effect size was small. Empirical findings regarding passive bystanding are more mixed: some research suggests that passive bystanding increases with age (Gini et al., 2022; Pöyhönen et al., 2012; Wang et al., 2024; Yang & Gao, 2023), while other studies find no significant association (Eijigu & Teketel, 2023; Gini et al., 2022; Mazzone et al., 2016; Pozzoli et al., 2016; Sjögren et al., 2021b; Thornberg et al., 2023; Tolmatcheff et al., 2022a). Only one study has examined whether age is associated with being an unconcerned or guilty bystander and found no significant relationship in either case (Obermann, 2011).

Regarding gender, research has shown that girls are more likely than boys to defend victims (Caravita et al., 2019, 2022; Gini et al., 2015, 2022; Mazzone et al., 2016; Sjögren et al., 2021a, b; Thornberg et al., 2022, 2023; Tolmatcheff et al., 2022a; van der Ploeg et al., 2017; Yang & Gao, 2023). Meta-analytic findings (Ma et al., 2019) support this pattern, but the effect was small. In contrast, boys have been found to be more likely than girls to engage in passive bystanding (Eijigu & Teketel, 2023; Sjögren et al., 2021a, b; Thornberg et al., 2023; Tolmatcheff et al., 2022a; Wang et al., 2024; Yang & Gao, 2023). Nevertheless, some studies report no significant association between gender and passive bystanding (Mazzone et al., 2016; Thornberg et al., 2022), and both Caravita et al. (2019) Pyhönen et al. (2012) found that girls were more likely to engage in both defending and passive bystanding. According to Obermann (2011), more boys than girls were classified as unconcerned and guilty bystanders, while more girls than boys were classified as defenders.

Based on the findings from Ma et al.'s (2019) meta-analysis, we hypothesized that girls and younger students would be more likely to engage in defending compared to boys and older students. Previous research on passive bystanding is more inconsistent but suggests that both age and gender may influence passive bystanding. Given the mixed findings and the fact that only one study has explored unconcerned and guilty bystanders (Obermann, 2011), we chose not to formulate specific hypotheses. Instead, age and gender were included as control variables in an exploratory manner regarding these forms of passive bystanding.

2 Method

2.1 Participants and procedure

Following the approval of the Regional Ethical Review Board at Linköping, students from grades 4–6 (ages 10–14 years) were recruited from 74 classrooms in 31 schools in Sweden. The original sample consisted of 1,527 students (755 girls [49%] and 772 [51%] boys). In the study, 268 students were excluded because their parents did not grant active consent; 82 students were absent at the data collection session or did not want to participate; 8 students were excluded due to participation difficulties (i.e., language, reading or cognitive difficulties); and 1 student was dropped because of full missing data on all scales included in the study, leading to the final sample of 1,168 students (53% girls, mean age = 11.71, *SD* = 0.86). Although socioeconomic status was not measured at the individual level, the sample of schools included a wide range of socio-geographic locations and socioeconomic backgrounds of students due to the purposeful sampling of schools. The vast majority of participants had a Swedish ethnic background, while a minority (14%) had foreign backgrounds; that is, either they had been born in another country, or both their parents were born in another country.

Students completed a questionnaire during regular school hours in their ordinary classrooms. Trained graduate students in psychology were present in the classrooms during the survey administration. They explained the study procedure, data confidentiality, and the possibility to withdraw from the study at any time without being penalized. They also assisted participants who needed help (e.g., by providing reading support and clarifying specific items or words in the questionnaire). Participants responded anonymously to the questionnaire. To ensure privacy, participants were instructed to move away from each other and separate their desks.

2.2 Measures

2.2.1 Moral disengagement

We used an 18-item moral disengagement in bullying scale (Thornberg & Jungert, 2014) to measure students' proneness to morally disengage in bullying situations (e.g., "It's okay to harm another person a couple of times a week if you do that to protect your friends", "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", "If you can't be like everybody else, you have to blame yourself if you get bullied"). Responses were rated on a 7-point scale (1 = "strongly disagree" to 7 = "strongly agree") and were averaged across the items to generate a composite score representing overall moral disengagement (Cronbach's alpha = 0.89). To validate the factor structure of the scale, a CFA was conducted. The model specified seven latent factors, with dehumanization and victim blaming combined into one factor (victim attribution) in line with previous studies. More specifically, the model specified seven latent factors, which were combined into a higher-order factor representing overall moral disengagement. The higher-order CFA was estimated using maximum likelihood estimation with robust standard errors (MLR) to account for the non-normal distribution

of the data (Mardia's skewness = 35,158.69, $p < .001$ and kurtosis = 350.87, $p < .001$; Shapiro–Wilk tests for all items, $p < .001$). The model provided support of the higher-order model: $\chi^2(128) = 311.28$, $p < .001$, CFI = 0.948, RMSEA = 0.051 [90% CI: 0.044, 0.059], and SRMR = 0.045. Thus, the scale can be utilized to measure moral disengagement either as a global construct or through its seven mechanisms.

2.2.2 Defender self-efficacy

Defender self-efficacy was measured using a 5-item Likert scale ranging from 1 = *strongly disagree* to 7 = *strongly agree* (Thornberg et al., 2017). The scale aimed to capture students' confidence in their ability to intervene effectively in bullying situations. Participants were asked to rate their confidence in their ability to intervene effectively in bullying situations (e.g., "I feel that I am very good at speaking up when students are mean toward another student"; "I feel that I am very good at helping students who are victims of bullying"). Responses to these items were averaged to form a composite score representing overall defender self-efficacy (Cronbach's alpha = 0.92). A CFA was conducted, using maximum likelihood estimation with robust standard errors (MLR) to account for the non-normal distribution of the data (Mardia's skewness = 402.14, $p < .001$ and kurtosis = 36.48, $p < .001$; Shapiro–Wilk tests for all items, $p < .001$), and confirmed the one-dimensional factor structure of the defender self-efficacy scale: $\chi^2(5) = 18.07$, $p = .003$, CFI = 0.994, RMSEA = 0.066 [90% CI: 0.035, 0.100], and SRMR = 0.012.

2.2.3 Bystander behaviors

Bystander behaviors were measured using two parallel 8-item, seven-point Likert scales, ranging from 1 = *strongly disagree* to 7 = *strongly agree*. The students were first asked about witnessing physical bullying, "If a student repeatedly gets punched, kicked, violently shoved or is held with force by students who are stronger, more popular, or more in charge in comparison to that student, what do you usually do?", followed by a set of items. Next, the students were asked about verbal bullying, "If a student is repeatedly teased or called names by students who are stronger, more popular, or more in charge in comparison to that student, what do you usually do?" The items on both scales were identical to ensure consistency in capturing different types of bystander behaviors across the two scenarios. Participants were asked to reflect on situations where they had witnessed victimization by peers who were bigger, stronger, more popular, or more powerful than the victimized student. After each scenario, students responded to items that measured unconcerned bystanding, guilty bystanding, and defending. *Unconcerned bystanding* was captured by the items, "I don't do anything because it's not my problem", "I don't care but just walk away", and "I don't care what they do but continue what I was doing". *Guilty bystanding* was captured by the items, "I feel bad because I don't help the student" and "I feel guilty because I don't do anything". *Defending* was captured by one reporting item ("I tell a teacher") and two confronting items ("I tell them to stop messing with the student", and "I step between and try to make them stop"), asked for both the physical and verbal scenarios. Preliminary item correlations indicated moderate links between report-

ing and confronting defending indicators (r=.36–0.50), stronger within-confronting associations (r=.70–0.83), and a high intercorrelation between the two reporting items (r=.78). Given these associations, we conducted sensitivity analyses estimating defending separately by confronting versus reporting indicators; results paralleled the main model (all significance patterns unchanged), although the association between defender self-efficacy and defending was attenuated in the reporting-only specification. Accordingly, we modeled defending as a single latent construct using all six indicators across scenarios. Responses for each scale were averaged across the two scenarios (physical and verbal) to create composite scores for each type of bystander behavior, resulting in three subscales: unconcerned bystanding, guilty bystanding, and defending.

To validate the dimensionality of the bystander behavior constructs, a confirmatory factor analysis (CFA) was conducted. The model specified three latent factors corresponding to defending, unconcerned bystanding, and guilty bystanding. Each factor was indicated by items from both the physical and verbal scenarios. Error covariances were specified between corresponding items across the two scenarios to account for shared method variance, as the items measured identical behaviors across two different types of bullying. Including these covariances was important to ensure that the similarity in the wording of items between physical and verbal scenarios did not artificially inflate the relationships among the latent factors, providing a more accurate representation of the constructs being measured. The CFA model was estimated using maximum likelihood estimation with robust standard errors (MLR) to account for the non-normal distribution of the data (as indicated by Mardia's test of multivariate skewness = 5,479.34, p <.001 and kurtosis = 108.03, p <.001, along with significant univariate Shapiro-Wilk tests for all items, p <.001) and provide more accurate parameter estimates and fit indices under such conditions. The model provided support for the three-factor structure: $\chi^2(93) = 246.41$, p <.001, CFI = 0.983, RMSEA = 0.045 [90% CI: 0.038; 0.052], and SRMR = 0.039. Furthermore, the scales displayed adequate internal consistency as indicated by Cronbach's alphas of 0.90 for defending, 0.89 for unconcerned bystanding, and 0.93 for guilty bystanding.

2.3 Statistical analyses

All analyses were carried out in R Studio using R version 4.4.2 (R Core Team, 2024). First, we analyzed descriptive statistics and correlations for all scales to assess their distributions and bivariate associations. Next, we used structural equation modeling (SEM) to investigate whether moral disengagement and defender self-efficacy were related to defending, unconcerned bystanding, and guilty bystanding in school bullying among pre-adolescents.

In the measurement part of the model, each latent variable was measured using multiple observed indicators: defending (six items), unconcerned bystanding (six items), guilty bystanding (four items), moral disengagement (18 items), and defender self-efficacy (five items). Given their conceptual overlap, we specified covariances between the three bystander behaviors. Additionally, residual covariances were specified between parallel items measuring these behaviors, as these items were nearly identical in wording, differing only in the scenario described. We evaluated model fit

using CFI, RMSEA, and SRMR, following common guidelines suggesting that CFI above approximately 0.95, RMSEA below 0.06, and SRMR below 0.08 indicate a good fit (Hu & Bentler, 1999). We used MLR (Maximum Likelihood with Robust Standard Errors) as the estimator, which provides standard errors robust to non-normality and heteroscedasticity. To handle missing data, we employed FIML (Full Information Maximum Likelihood) estimation, allowing all available information to contribute to parameter estimations without excluding participants with incomplete responses. As students were nested within classrooms, we also adjusted for clustering effects by specifying the class variable as a clustering factor in the model.

In the structural model, moral disengagement and defender self-efficacy were specified as predictors of the three bystander behaviors, while age and gender were included as covariates. In addition to this main-effects model, we tested an interaction model in which an interaction term between moral disengagement and defender self-efficacy was added as a predictor of the three bystander behaviors.

After an initial evaluation of the main-effects model, which indicated a slightly lower-than-expected CFI (0.91), we examined modification indices and residual covariances. Based on both statistical and theoretical considerations, we then specified additional correlated residuals among the moral disengagement items within three mechanisms: moral justification (i.e., MD1 ~~ MD12), displacement of responsibility (i.e., MD7 ~~ MD13, MD7 ~~ MD16, MD13 ~~ MD16), and victim attribution (i.e., MD3 ~~ MD4, MD3 ~~ MD17, MD4 ~~ MD17, see Thornberg & Jungert, 2014). These modifications were justified on theoretical grounds, as items within each mechanism not only measured the same underlying construct but also shared highly similar wording and structure, which could introduce systematic shared variance beyond what was accounted for by the latent construct. Additionally, some items within the same mechanism reflected highly specific aspects of the broader mechanism. For example, while there are various ways to justify harmful behavior, both items measuring moral justification in this study specifically framed the justification in terms of friendships—one emphasizing the protection of friends and the other emphasizing helping friends.

3 Results

3.1 Descriptive results and correlations

Table 1 presents descriptive statistics and bivariate correlations of the study variables. Participants indicated higher levels of defender self-efficacy (M=4.38, SD=1.65) compared to moral disengagement (M=1.78, SD=0.77). Among the bystander behaviors, defending was reported most frequently (M=4.42, SD=1.64), followed by guilty bystanding (M=3.41, SD=1.9), and unconcerned bystanding (M=2.34, SD=1.3). The correlation analyses showed that defender self-efficacy was positively related to defending (r=.64, p <.001) and negatively related to both unconcerned (r = –.30, p <.001) and guilty bystanding (r = –.28, p <.001). Conversely, moral disengagement was positively associated with unconcerned bystanding (r=.44, p <.001) and negatively associated with guilty bystanding (r = –.08, p <.01) and defending (r

Table 1 Correlations, means (M), and standard deviations (SD) of the study variables

Measure	1	2	3	4	5	6	7	M	SD
1. Moral Disengagement	–							1.78	0.77
2. Defender Self-Efficacy	0.02	–						4.38	1.65
3. Defending	–0.14***	0.64***	–					4.42	1.64
4. Unconcerned Bystanding	0.44***	–0.30***	–0.43***	–				2.34	1.30
5. Guilty Bystanding	–0.08**	–0.28***	–0.13***	0.01	–			3.41	1.90
6. Age	0.05	–0.17***	–0.20***	0.13***	0.01	–		11.71	0.86
7. Gender	–0.23***	0.05	0.10***	–0.25***	0.10**	–0.03	–	0.53	0.50

**p<.01, ***p<.001; gender: boy = 0, girl = 1

= –.14, p <.001). Regarding associations among the bystander behaviors, defending was negatively correlated with both guilty bystanding (r = –.13, p<.001) and unconcerned bystanding (r = –.43, p<.001), whereas guilty and unconcerned bystanding were not significantly correlated. Regarding the control variables, age was negatively correlated with defender self-efficacy (r = –.17, p <.001) and defending (r = –.20, p <.001), and positively correlated with unconcerned bystanding (r=.13, p <.001), while boys were more likely to score higher in moral disengagement (r = –.23, p <.001) and unconcerned bystanding (r = –.25, p <.001), and girls were more likely to score higher in defending (r=.10, p <.001) and guilty bystanding (r=.10, p <.01).

3.2 Main analysis

First, we evaluated the fit of the proposed main-effects model to the data. The model demonstrated a good fit based on the fit indices ($\chi^2(749) = 1722$, p=.001, CFI = 0.947, RMSEA = 0.040, 90% CI [0.037, 0.042], SRMR = 0.052), indicating that the model adequately represents the relationships among the study variables. Next, we examined the structural paths to test whether moral disengagement and defender self-efficacy were significant predictors of defending, unconcerned bystanding, and guilty bystanding (for an overview of the standardized coefficients, see Fig. 1).

Defending behavior was strongly positively associated with defender self-efficacy (β = 0.72, p<.001), indicating that students with higher confidence in their ability to defend were more likely to intervene when witnessing bullying. In contrast, moral disengagement was negatively, although weakly, associated with defending (β = –0.13, p=.001), suggesting that students who disengaged morally from the situation were less likely to act in defense of their peers. Neither age nor gender were significantly related to defending behavior.

Unconcerned bystanding was negatively associated with defender self-efficacy (β = –0.32, p<.001), meaning that students with greater confidence in their defending abilities were less likely to remain passive and unconcerned in bullying situations. Conversely, moral disengagement was strongly positively associated with unconcerned bystanding (β = 0.48, p<.001), indicating that students who justified or distanced themselves from the moral implications of bullying were more likely to ignore such incidents. Additionally, gender was significantly related to this form of passive bystanding (β = –0.15, p<.001), with boys reporting higher levels of uncon-

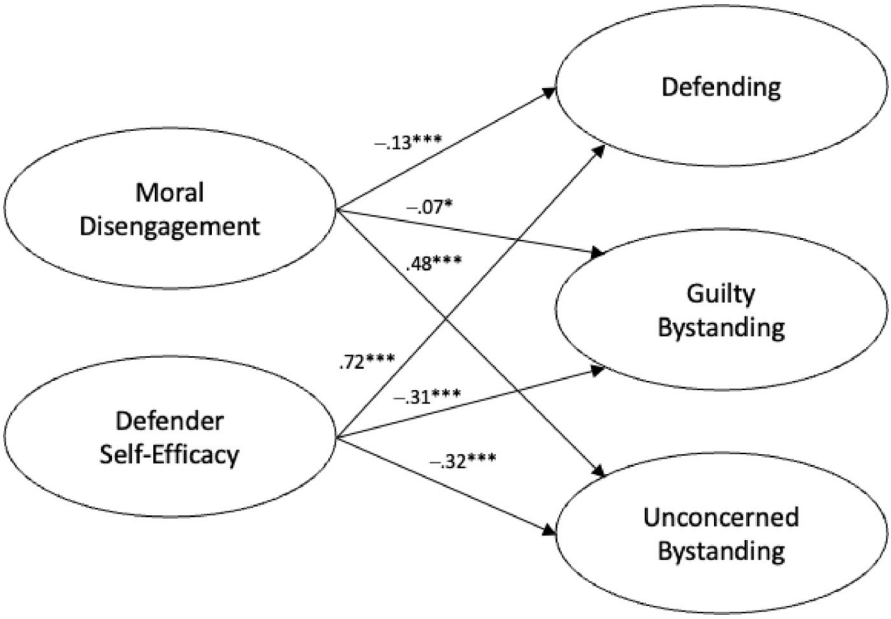

Fig. 1 Standardized coefficients of the SEM analysis. *p<.05, ***p<.001

cerned bystanding compared to girls, whereas age was not significantly associated with unconcerned bystanding.

Guilty bystanding was negatively associated with defender self-efficacy (β = –0.31, p<.001) and, to a lower degree, moral disengagement (β = –0.07, p=.033), suggesting that students who felt capable of defending or who reported higher levels of moral disengagement from the situation were less likely to passively witness bullying while experiencing guilt. Gender was significantly linked to this form of passive bystanding (β=0.10, p<.001), indicating that girls were more likely than boys to report guilty bystanding, whereas age was not significantly linked to guilty bystanding.

Finally, in addition to the main-effects model, we also tested an interaction model including the interaction between moral disengagement and defender self-efficacy. This model demonstrated poorer fit ($\chi^2(166)=1149$, p=.001, CFI=0.906, RMSEA=0.069, 90% CI [0.066, 0.072], SRMR=0.050). The interaction was not significant for defending (β = −0.01, p=.80) or unconcerned bystanding (β=0.03, p=.55), but was positive and significant for guilty bystanding (β=0.19, p<.001).

4 Discussion

The present study aimed to examine whether moral disengagement and defender self-efficacy were related to defending, unconcerned bystanding, and guilty bystanding in school bullying among pre-adolescent students. In line with our first hypothesis, we found that moral disengagement was positively associated with unconcerned

bystanding and negatively associated with defending and guilty bystanding. While these findings align with previous research demonstrating the negative relationship between moral disengagement and defending (Killer et al., 2019; Luo & Bussey, 2023; Ma et al., 2019), they also provide insights into the inconsistent findings regarding the link between moral disengagement and passive bystanding (Killer et al., 2019; Luo & Bussey, 2023).

Distinguishing between unconcerned and guilty bystanding (Obermann, 2011) appears to be an important step in better understanding passive bystanding within a social cognitive framework (Bandura, 1997, 2016). The present findings suggest that students who tend to exhibit unconcerned bystanding are more prone to morally disengage, whereas students who are more likely to engage in guilty bystanding are less prone to morally disengage. This pattern can be related to Obermann's (2011) study, which found that unconcerned bystanders scored significantly higher in moral disengagement than guilty bystanders. Like Obermann's (2011) study, and with reference to the social cognitive theory of moral disengagement (Bandura, 2016), this finding supports "the underlying theoretical assumption that individuals who feel guilty and responsible are less likely to disengage morally" (Obermann, 2011, p. 252) than those who feel unconcerned. At the same time, the present results showed that moral disengagement was more strongly related to unconcerned bystanding compared to the two other bystander behaviors, suggesting that it plays a more crucial role in unconcerned bystanding than in defending and guilty bystanding. Students who morally disengage are less likely to perceive the situation as an emergency—thereby failing to recognize its moral issue and the victim's need for help—or do not feel personally responsible for intervening (Bandura, 2016). High moral disengagement enables passive bystanders to remain indifferent and unconcerned about both the situation they witness and their own role within it, and thus, corrupts their moral agency.

In accordance with social cognitive theory (Bandura, 1997, 2016), our findings highlight that defender self-efficacy is a key component of bystander's moral agency, showing that defender self-efficacy was positively related to defending. Specifically, students who scored higher in defender self-efficacy were more likely to intervene in bullying incidents to help the victim. This result supports our second hypothesis and is consistent with previous research showing a positive relationship between defender self-efficacy and defending (e.g., Eijigu & Teketel, 2023; Gini et al., 2022; Sjögren et al., 2024; for a meta-analysis, see Ma et al., 2019). High defender self-efficacy helps bystanders overcome inhibitory factors such as hesitation caused by social status differences, self-doubt regarding their ability to intervene effectively, and fear of retaliation or social repercussions. These factors have been identified as common inhibitors of intervention in studies examining students' perspectives on bystander behavior (Forsberg et al., 2018; Strindberg et al., 2020; Thornberg et al., 2018; Williams et al., 2018). High self-efficacy facilitates them to transform their moral understanding of the bullying situation into moral action (Bandura, 2016); that is, it helps them stand up for the victim through confronting the bullies and reporting to authority.

Additionally, in line with our third hypothesis and prior research demonstrating a negative link between defender self-efficacy and passive bystanding as a global construct (Eijigu & Teketel, 2023; Sjögren et al., 2020, 2021b, 2024b; Thornberg et al., 2017), the present findings showed that defender self-efficacy was negatively

associated with guilty bystanding and—beyond our hypotheses—with unconcerned bystanding. According to social-cognitive theory (Bandura, 1997), mastery experiences are the most influential source of self-efficacy. These involve personal experiences of success in performing the behavior, task, or activity. Successful experiences increase self-efficacy beliefs. Thus, a possible explanation for the negative association between defender self-efficacy and both forms of passive bystanding in the present study is that students who frequently act as either unconcerned or guilty bystanders in bullying situations have fewer opportunities to develop their defending capacity. As a result, they experience fewer instances of successfully organizing and executing actions to defend a bullied peer and stop bullying. Consequently, their recurrent pattern of passive bystanding limits their opportunities to cultivate a high level of defender self-efficacy.

Complementing the main effects, we also tested whether moral disengagement conditions how defender self-efficacy translates into bystander behavior (see also Sjögren et al., 2024a, b). The interaction did not emerge for defending or unconcerned bystanding but was positive and significant for guilty bystanding, suggesting that the negative association between defender self-efficacy and guilty bystanding weakens as moral disengagement increases. This pattern is consistent with the view that deactivated moral self-regulation can blunt the behavioral expression of self-efficacy, particularly for morally ambivalent passivity. However, it should be noted that the interaction model displayed poorer fit than the main-effects model.

With respect to the control variables, the observed bivariate correlations in this study aligned with our hypothesis and previous research in that defending declined with age and was more prevalent among girls (e.g., Ma et al., 2019; Gini et al., 2022; Thornberg et al., 2023). However, our findings did not reveal any significant association between either age or gender and defending once moral disengagement and defender self-efficacy were accounted for. These results suggest that individual differences in moral disengagement and defender self-efficacy may play a more central role in shaping defending than age or gender in pre-adolescence. In terms of passive bystanding, gender significantly predicted both forms: boys were more likely to engage in unconcerned bystanding, whereas girls were more likely to report guilty bystanding. However, age was not significantly associated with either unconcerned or guilty bystanding when moral disengagement and defender self-efficacy were included as independent variables in the structural equation model. Taken together, these findings highlight the importance of considering gender-specific socialization experiences in understanding different passive bystanding in bullying, while also suggesting that age-related trends may be less pronounced during the upper elementary years when accounting for moral disengagement and defender self-efficacy.

4.1 Limitations

Our study employed a cross-sectional design, which limits our ability to establish causality or determine the direction of effects. The social cognitive theory posits bidirectional relationships between cognition and behavior, suggesting a complex interplay of environmental, individual, and behavioral influences (Bandura, 1997, 2016). Future research should therefore adopt longitudinal approaches to examine whether

moral disengagement and defender self-efficacy predict defending, guilty bystanding, and unconcerned bystanding over time, and whether these three bystander behaviors predict moral disengagement and defender self-efficacy over time.

Moreover, intervention studies with an experimental design (e.g., randomized control trials) aimed at decreasing moral disengagement and increasing defender self-efficacy among students should be conducted to test possible causal associations between these social-cognitive factors and the three bystander behaviors. This was partly demonstrated by Tolmatcheff et al. (2022), who showed that a decrease in moral disengagement following a targeted intervention was significantly associated with an increase in defending and a reduction in passive bystanding. Furthermore, the reliance on self-reported data may have inflated variable associations due to shared methods variance. Self-reporting is also susceptible to biases, including social desirability bias, recall bias, perception bias, and intentionally exaggerated responses. These factors could potentially affect the accuracy and reliability of the findings. Future research should therefore include multiple data sources, such as peer nominations and teacher reported data, in addition to self-reported student data.

Another limitation is that the present study treated and measured defending as a unidimensional construct, while researchers more recently have suggested that defending can be considered as a multidimensional construct including assertive defending, aggressive defending, comforting victims, reporting to authority (Lambe & Craig, 2020), and tactical defending (Wang et al., 2023). Thus, future studies should examine in what degree moral disengagement and defender self-efficacy are linked to different forms of peer defending, using more comprehensive scales that capture these different subtypes of defending. Future research may also consider examining the potential interplay between defender self-efficacy and moral disengagement in predicting subtypes of defending and passive bystanding. Additionally, the current study focused on students within a particular age range and from specific areas in Sweden. The generalizability of our findings to other demographic groups or geographic regions should therefore be considered with caution. Thus, the present findings need to be replicated in diverse cultural contexts and with various age groups to establish their broader applicability and robustness.

4.2 Practical implications

The findings of this study offer valuable insights for practitioners and policymakers interested in school-based intervention programs aiming to reduce bullying and promote prosocial bystander behavior. First, given the strong positive relationship between defender self-efficacy and defending behavior, interventions should prioritize strategies that enhance students' confidence in their ability to intervene in bullying situations. Self-efficacy can be strengthened through mastery experiences, social modeling, and verbal persuasion (Bandura, 1997). School-based programs could therefore incorporate role-playing exercises, peer mentoring, and scenario-based training to help students practice defending behaviors in a controlled and supportive environment (Forsberg et al., 2018; Gini et al., 2022; Thornberg et al., 2018). Moreover, training sessions that equip students with concrete strategies for effectively intervening may reduce the fear of retaliation or social repercussions (Forsberg et al.,

2018; Thornberg et al., 2018). Additionally, encouraging positive social reinforcement from teachers and peers may further strengthen students' confidence in their ability to act as defenders (Williams et al., 2018).

Second, the study highlights the role of moral disengagement in reinforcing passive bystanding, particularly unconcerned bystanding. Higher moral disengagement levels allow students to justify inaction, reducing their sense of personal responsibility (Bandura, 2016; Killer et al., 2019). To address this, educators should implement moral reasoning training, where students are encouraged to critically reflect on their justifications for non-intervention. Classroom discussions on empathy, responsibility, and the real-life consequences of bullying and moral dilemma discussions can help students develop a stronger sense of moral accountability and recognize the ethical implications of their bystander behavior, fostering the sense of a moral obligation to intervene (Gini et al., 2020; Luo & Bussey, 2023). Moreover, debiasing techniques could be used to foster students' critical thinking and teach them to recognize and counteract morally disengaged reasoning (Horton, 2004), by making the consequences of such biases visible, contrasting diffusion of responsibility, and avoiding the use of palliative language. Examples of these techniques include generating counter-arguments, articulating reasons in support of opposing beliefs, expressing the line of thinking that leads one to hold particular beliefs (Horton, 2004).

Third, the distinction between unconcerned and guilty bystanders offers a more nuanced understanding of passive bystanding (Obermann, 2011). Interventions should tailor strategies to these two groups. For instance, unconcerned bystanders may benefit from programs that increase their emotional engagement with victims, such as perspective-taking exercises or storytelling interventions where they listen to personal accounts from victims (Brüggemann et al., 2019; Strindberg et al., 2020). Guilty bystanders, on the other hand, may already recognize the need to intervene but lack the confidence or social support to do so. For them, assertiveness training and structured peer support systems could provide the encouragement needed to transition from passive to active defending (Dovidio et al., 2006; Thornberg et al., 2020).

Lastly, schools should create environments that encourage students to take an active stance against bullying. Implementing school-wide anti-bullying policies, fostering a culture of collective responsibility, and involving students in decision-making processes can reinforce a sense of agency and responsibility among bystanders (Salmivalli, 1999; Strindberg et al., 2020). By addressing both the cognitive and emotional barriers to bystander intervention, schools can create an environment where defending victims becomes the norm, ultimately contributing to a more supportive and inclusive school climate.

Acknowledgements This research was partially supported by a grant awarded to Robert Thornberg from The Swedish Research Council [grant number D0775301].

Funding Open access funding provided by Linköping University.

References

Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman and Company.

Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, *3*(3), 193–209. https://doi.org/10.1207/s15327957pspr0303_

Bandura, A. (2016). *Moral disengagement: How people do harm and live with themselves*. Worth.

Brüggemann, A. J., Forsberg, C., Colnerud, G., Wijma, B., & Thornberg, R. (2019). Bystander passivity in health care and school settings: Moral disengagement, moral distress, and opportunities for moral education. *Journal of Moral Education, 48*(2), 199–213. https://doi.org/10.1080/03057240.2018.1471391

Bussey, K. (2023). The contribution of social cognitive theory to school bullying research and practice. *Theory Into Practice, 62*(3), 293–305. https://doi.org/10.1080/00405841.2023.2226549

Campbell, M., Hand, K., Shaw, T., Runions, K., Burns, S., Lester, L., & Cross, D. (2023). Adolescent proactive bystanding versus passive bystanding responses to school bullying: The role of peer and moral predictors. *International Journal of Bullying Prevention, 5*(4), 296–305. https://doi.org/10.1007/s42380-020-00075-2

Cappadocia, M. C., Pepler, D., Cummings, J. G., & Craig, W. (2012). Individual motivations and characteristics associated with bystander intervention during bullying episodes among children and youth. *Canadian Journal of School Psychology, 27*(3), 201–216. https://doi.org/10.1177/0829573512450567

Caravita, S. C. S., Strohmeier, D., Salmivalli, C., & Di Blasio, P. (2019). Bullying immigrant versus non-immigrant peers: Moral disengagement and participant roles. *Journal of School Psychology, 75*, 119–133. https://doi.org/10.1016/j.jsp.2019.07.005

Caravita, S. C. S., Finne, J. N., & Fandrem, H. (2022). Two dimensions of moral cognition as correlates of different forms of participation in bullying. *Frontiers in Psychology, 12*, Article 768503. https://doi.org/10.3389/fpsyg.2021.768503

R Core Team (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/

Craig, W. M., Pepler, D., & Atlas, R. (2000). Observations of bullying in the playground and in the classroom. *School Psychology International, 21*(1), 22–36. https://doi.org/10.1177/0143034300211002

Dovidio, J. F., Piliavin, J. A., Schroeder, D. A., & Penner, L. A. (2006). *The social psychology of prosocial behavior*. Erlbaum.

Eijigu, T. D., & Teketel, S. Z. (2023). Students' defending and passive bystanding behaviours when witnessing school bullying in North Western ethiopia: Roles of personal and situational factors. *Journal of Community & Applied Social Psychology*, *33*(2), 202–220. https://doi.org/10.1002/casp.2676

Fischer, P., Krueger, J. I., Greitemeyer, T., Vogrincic, C., Kastenmüller, A., Frey, D., Heene, M., Wicher, M., & Kainbacher, M. (2011). The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin, 137*(4), 517–537. https://doi.org/10.1037/a0023304

Forsberg, C., Wood, L., Smith, J., Varjas, K., Meyers, J., Jungert, T., & Thornberg, R. (2018). Students' views on factors affecting their bystander acts in bullying situations: A cross-collaborative conceptual qualitative analysis. *Research Papers in Education*, *33*(1), 127–142. https://doi.org/10.1080/02671522.2016.1271001

Garandeau, C. F., Vermande, M. M., Reijntjes, A. H. A., & Aarts, E. (2022). Classroom bullying norms and peer status: Effects on victim-oriented and bully-oriented defending. *International Journal of Behavioral Development, 46*(5), 401–410. https://doi.org/10.1177/0165025419894722

Gini, G. (2006). Social cognition and moral cognition in bullying: What's wrong? *Aggressive Behavior, 32*(6), 528–539. https://doi.org/10.1002/ab.20153

Gini, G., Albiero, P., Benelli, B., & Altoè, G. (2008). Determinants of adolescents' active defending and passive bystanding behavior in bullying. *Journal of Adolescence, 31*(1), 93–105. https://doi.org/10.1016/j.adolescence.2007.05.002

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. https://doi.org/10.1007/s10802-014-9920-7

Gini, G., Thornberg, R., & Pozzoli, T. (2020). Individual moral disengagement and bystander behavior in bullying: The role of moral distress and collective moral disengagement. *Psychology of Violence, 10*(1), 38–47. https://doi.org/10.1037/vio0000223

Gini, G., Pozzoli, T., Angelini, F., Thornberg, R., & Demaray, M. K. (2022). Longitudinal associations of social-cognitive and moral correlates with defending in bullying. *Journal of School Psychology, 91*, 146–159. https://doi.org/10.1016/j.jsp.2022.01.005

Hawkins, D. L., Pepler, D. J., & Craig, W. M. (2001). Naturalistic observations of peer interventions in bullying. *Social Development, 10*(4), 512–527. https://doi.org/10.1111/1467-9507.00178

Horton, K. (2004). Aid and bias. *Inquiry, 47*, 545–61. https://doi.org/10.1080/00201740410004287

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jiang, S., Liu, R. D., Ding, Y., Jiang, R., Fu, X., & Hong, W. (2022). Why the victims of bullying are more likely to avoid involvement when witnessing bullying situations: The role of bullying sensitivity and moral disengagement. *Journal of Interpersonal Violence*, *37*(5–6), NP3062–NP3083. https://doi.org/10.1177/08862605209481

Killer, B., Bussey, K., Hawes, D. J., & Hunt, C. (2019). A meta-analysis of the relationship between moral disengagement and bullying roles in youth. *Aggressive Behavior, 45*(6), 450–462. https://doi.org/10.1002/ab.21833

Lambe, L. J., & Craig, W. M. (2020). Peer defending as a multidimensional behavior: Development and validation of the defending behaviors scale. *Journal of School Psychology, 78*, 38–53. https://doi.org/10.1016/j.jsp.2019.12.001

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

Luo, A., & Bussey, K. (2023). Moral disengagement in youth: A meta-analytic review. *Developmental Review*, *70*, 101101. https://doi.org/10.1016/j.dr.2023.101101

Ma, T. L., Meter, D. J., Chen, W. T., & Lee, Y. (2019). Defending behavior of peer victimization in school and cyber context during childhood and adolescence: A meta-analytic review of individual and peer-relational characteristics. *Psychological Bulletin, 145*(9), 891–928. https://doi.org/10.1037/bul0000205

Mazzone, A., Camodeca, M., & Salmivalli, C. (2016). Interactive effects of guilt and moral disengagement on bullying, defending and outsider behavior. *Journal of Moral Education, 45*(4), 419–432. https://doi.org/10.1080/03057240.2016.1216399

Obermann, M. L. (2011). Moral disengagement among bystanders to school bullying. *Journal of School Violence, 10*(3), 239–257. https://doi.org/10.1080/15388220.2011.578276

Peets, K., Pöyhönen, V., Juvonen, J., & Salmivalli, C. (2015). Classroom norms of bullying alter the degree to which children defend in response to their affective empathy and power. *Developmental Psychology*, *51*(7), 913–920. https://doi.org/10.1037/a0039287

Polanin, J. R., Espelage, D. L., & Pigott, T. D. (2012). A meta-analysis of school-based bullying prevention programs' effects on bystander intervention behavior. *School Psychology Review, 41*(1), 47–65. https://doi.org/10.1080/02796015.2012.12087375

Pöyhönen, V., Juvonen, J., & Salmivalli, C. (2012). Standing up for the victim, siding with the bully or standing by? Bystander responses in bullying situations. *Social Development, 21*(4), 722–741. https://doi.org/10.1111/j.1467-9507.2012.00662.x

Pozzoli, T., Gini, G., & Thornberg, R. (2016). Bullying and defending behavior: The role of explicit and implicit moral cognition. *Journal of School Psychology, 59*, 67–81. https://doi.org/10.1016/j.jsp.2016.09.005

Salmivalli, C. (1999). Participant role approach to school bullying: Implications for interventions. *Journal of Adolescence, 22*(4), 453–459. https://doi.org/10.1006/jado.1999.0239

Shi, L., Wu, Y., & Zhou, Y. (2024). The role of moral disengagement and classroom collective efficacy in social support and bystander behavior of elementary school students: A multilevel moderated mediation analysis. *School Psychology*. https://doi.org/10.1037/spq0000639. Advance online publication.

Sjögren, B., Thornberg, R., Wänström, L., & Gini, G. (2020). Associations between individual and collective efficacy beliefs and students' bystander behavior. *Psychology in the Schools, 57*(11), 1710–1723. https://doi.org/10.1002/pits.22412

Sjögren, B., Thornberg, R., Wänström, L., & Gini, G. (2021a). Associations between students' bystander behavior and individual and classroom collective moral disengagement. *Educational Psychology, 41*(3), 264–281. https://doi.org/10.1080/01443410.2020.1828832

Sjögren, B., Thornberg, R., Wänström, L., & Gini, G. (2021b). Bystander behaviour in peer victimisation: Moral disengagement, defender self-efficacy and student-teacher relationship quality. *Research Papers in Education, 36*(5), 588–610. https://doi.org/10.1080/02671522.2020.1723679

Sjögren, B., Thornberg, R., Kim, J., Hong, J. S., & Kloo, M. (2024a). Basic moral sensitivity, moral disengagement, and defender self-efficacy as predictors of students' self-reported bystander behaviors over a school year: A growth curve analysis. *Frontiers in Psychology, 15*, Article 1378755. https://doi.org/10.3389/fpsyg.2024.1378755

Sjögren, B., Thornberg, R., & Pozzoli, T. (2024b). Reciprocal longitudinal associations of defender self-efficacy with defending and passive bystanding in peer victimization. *Psychology in the Schools, 61*(5), 1766–1788. https://doi.org/10.1002/pits.23134

Snijders, T. A., & Bosker, R. (2011). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*.

Steinvik, H. R., Duffy, A. L., & Zimmer-Gemback, M. J. (2025). Adolescents' compassion is distinctively associated with more prosocial and less aggressive defending against bullying when considering empathic emotions and costs. *Journal of Adolescence Advance Online Publication*. https://doi.org/10.1002/jad.12513

Strindberg, J., Horton, P., & Thornberg, R. (2020). The fear of being singled out: Pupils' perspectives on victimisation and bystanding in bullying situations. *British Journal of Sociology of Education*, *41*(7), 942–957. https://doi.org/10.1080/01425692.2020.1789846

Thornberg, R., & Jungert, T. (2013). Bystander behavior in bullying situations: Basic moral sensitivity, moral disengagement and defender self-efficacy. *Journal of Adolescence, 36*(3), 475–483.

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

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

Thornberg, R., Landgren, L., & Wiman, E. (2018). It depends: A qualitative study on how adolescents explain bystander intervention and non-intervention in bullying situation. *School Psychology International*, *39*(4), 400–415. https://doi.org/10.1177/0143034318779225

Thornberg, R., Wänström, L., Elmelid, R., Johansson, A., & Mellander, E. (2020). Standing up for the victim or supporting the bully? Bystander responses and their associations with moral disengagement, defender self-efficacy, and collective efficacy. *Social Psychology of Education, 23*(3), 563–581. https://doi.org/10.1007/s11218-020-09549-z

Thornberg, R., Pozzoli, T., & Gini, G. (2022). Defending or remaining passive as a bystander of school bullying in Sweden: The role of moral disengagement and antibullying class norms. *Journal of Interpersonal Violence, 37*(19–20), NP18666–NP18689. https://doi.org/10.1177/08862605211037427

Thornberg, R., Jungert, T., & Hong, J. S. (2023). The indirect association between moral disengagement and bystander behaviors in school bullying through motivation: Structural equation modelling and mediation analysis. *Social Psychology of Education, 26*, 533–556. https://doi.org/10.1007/s11218-022-09754-y

Tolmatcheff, C., Galand, B., & Roskam, I. (2022a). Validation of the French version of the moral disengagement in bullying scale: Testing bandura's conceptual model. *Journal of School Psychology, 91*, 81–96. https://doi.org/10.1016/j.jsp.2022.01.002

Tolmatcheff, C., Galand, B., Roskam, I., & Veenstra, R. (2022b). The effectiveness of moral disengagement and social norms as anti-bullying components: A randomized controlled trial. *Child Development*, *93*(6), 1873–1888. https://doi.org/10.1111/cdev.13828

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

Wang, Z., Laninga-Wijnen, L., Garandeau, C. F., & Liu, J. (2023). Development and validation of the adolescent defending behaviors questionnaire among Chinese early adolescents. *Assessment, 30*(7), 2258–2275. https://doi.org/10.1177/10731911221149082

Wang, X., Wei, H., & Wang, P. (2024). Adolescents high in callous-unemotional traits are prone to be bystanders: The roles of moral disengagement, moral identity, and perceived social support. *Child Psychiatry*. https://doi.org/10.1007/s10578-024-01709-y. & *Human Development.* Advance online publication

Williams, A. J., Courtney, S. B., & Blake, J. J. (2018). High school bystanders motivation and response during bias-based bullying. *Psychology in the Schools*, *55*(10), 1259–1273. https://doi.org/10.1002/pits.22186

Yang, L., & Gao, T. (2023). Defending or not? The role of peer status, social self-efficacy, and moral disengagement on Chinese adolescents' bystander behaviors in bullying situations. *Current Psychology, 42*, 29616–29627. https://doi.org/10.1007/s12144-022-04039-1

Yoo, H. N., & Smetana, J. G. (2022). Distinctions between moral and conventional judgments from early to middle childhood: A meta-analysis of social domain theory research. *Developmental Psychology, 58*(5), 874–889. https://doi.org/10.1037/dev0001330

Zych, I., Gómez-Ortiz, O., Fernández Touceda, L., Nasaescu, E., & Llorent, V. J. (2020). Moral disengagement induction as a predictor of bullying and cyberbullying: Mediation by children's moral disengagement, moral emotions, and validation of a questionnaire. *Child Indicators Research*, *13*(3), 1065–1083. https://doi.org/10.1007/s12187-019-09670-2

Robert Thornberg is a Professor in the Department of Behavioural Sciences and Learning at Linköping University, Sweden. His primary research interests focus on cognitive, social, and moral processes underlying bullying, peer victimization, and bystander behavior. Additional areas of interest include hate speech, sexual harassment, school safety, and academic engagement among children and adolescents, as well as teacher efficacy and experiences of emotional distress among student teachers and novice teachers.

Björn Sjögren is a Senior Lecturer at the Department of Behavioural Sciences and Learning at Linköping University, Sweden. His research interests include cognitive and social processes related to children's social interactions with a specific focus on bullying, peer victimization, and bystander behaviors.

Gianluca Gini is Professor in the Department of Developmental and Social Psychology, University of Padua, Italy. His major research interests include the interplay of individual and group processes associated with different roles of participation in bullying, the health and psychosocial consequences of peer victimization, and the development of adolescents' relationships with peers on social media.

Tiziana Pozzoli is PhD in developmental psychology. She is currently Associate Professor in the Department of Developmental and Social Psychology, University of Padova, Italy. Her research interests concern individual and contextual variables associated with different behavior during bullying episodes, focusing especially on bystanders' behavior.