



# Minimizing responsibility in the aggressive dynamics of bullying and its impact on other strategies of moral disengagement: a longitudinal study

Daniel Falla<sup>1</sup> · Rosario Ortega-Ruiz<sup>1</sup> · Eva M. Romera<sup>1</sup>

Accepted: 29 December 2022 / Published online: 18 January 2023  
© The Author(s) 2023, corrected publication 2023

## Abstract

Not assuming responsibility, or minimizing it, after committing an offence is one of the four moral disengagement strategies linked to the phenomenon of bullying described by the Social Cognitive Theory. However, to date, there has been no research into the role of the agent's locus in this process and the mediating effect of bullying perpetration in the possible evolutionary sequencing of moral disengagement strategies. This study addresses both of these goals. A total of 1107 schoolchildren (54.7% girls;  $M_{\text{age}} = 14.49$ ;  $SD = 0.789$ ) were surveyed in a longitudinal study at three time points spaced six months apart. The results indicated that minimizing responsibility directly predicts both cognitive restructuring and distortion of consequences. They also confirmed that aggressive perpetration in bullying has a mediating effect on all three strategies. Nevertheless, this sequential dynamic does not include dehumanization, which was not directly linked to minimizing responsibility and was mediated by the perpetration of aggression in bullying. We discuss the extent to which minimizing responsibility is the first step in a temporal sequence of moral disengagement mechanisms that help maintain the aggressive dynamic in bullying, so that it stimulates the other mechanisms and incorporates the locus of the aggressive agent. These findings allow us to advance in our understanding of the ethical dimension (sensitivity and moral criteria) implicit in the phenomenon of unjustified aggressiveness known as bullying.

**Keywords** Minimizing responsibility · Bullying perpetrators · Moral disengagement strategies · Longitudinal study

## Introduction

Interpersonal relationships among schoolchildren comprise a framework for communication and joint activity that forms a prosperous, dynamic ecosystem which stimulates the socialization process and improves both the sensitivity and the social and moral adjustment of peer groups (Ortega-Ruiz, 2010; Veiga Simão et al., 2017). However, in the general context of education, not all interactions between peers constitute an expression of sensitivity and good judgment or respect the rules of fairness and moral mutuality. On the contrary, they often involve phenomena of unjustified aggressiveness and immoral behaviour, exercised by certain individuals or groups towards others who they consider

weaker or inferior. This is the case of the phenomenon of bullying, which involves disrespect, intimidation, aggression and mistreatment which not only emotionally and psychologically damages the victim, but also affects the sensitivity and moral criteria of all those involved (Falla et al., 2021). This type of phenomenon occurs in most schools worldwide, and prevalence rates show that around one in three schoolchildren is involved (Modecki et al., 2014).

The psychological and socioemotional damage that this type of violence causes the victims has been evidenced in different studies that point to an increase in depression and anxiety, suicidal ideation and loneliness, among others (Moore et al., 2017). Although the moral consequences have been less studied, a recent study concludes that victims can also become aggressors, which damages their moral judgment, thus producing the so-called cycle of violence (Falla et al., 2022). However, it is also important to understand the consequences and psychological states that this involvement has on aggressors, specifically in the moral sphere, as it has been shown that they are at greater risk of continuing their

✉ Eva M. Romera  
eva.romera@uco.es

<sup>1</sup> Department of Psychology, Universidad de Córdoba, Av San Alberto Magno, s/n, 14071 Cordoba, Spain

antisocial behaviour in the future (Nasaescu et al., 2020). In this way, violence prevention through the promotion of a positive moral development at an earlier age is essential.

Work on bullying and morality has been mainly based on Bandura's (1986) Social Cognitive Theory. However, much of the research has focused in particular on moral disengagement (MD) as a risk factor in bullies (Killer et al., 2019), while only a small number of works have explored whether bullying aggression may influence MD (Thornberg et al., 2019). However, there are aspects of the theory that have hardly been addressed and whose ideas can be equally valuable in anti-bullying programmes, for example, the selective and progressive character of MD strategies. The work of Bandura and later authors (e.g., Thornberg et al., 2020) recognises MD as a selective cognitive process with a progressive nature, in which the result of the behavioural evaluation can change depending on the situation. However, to date, no research has looked into whether some MD strategies may be related in the long run to the other strategies, thus favouring a progression in moral inhibition and, in turn, the immoral reasoning of bullies.

The present study aims to respond to this gap in the research by hypothesising that the minimisation of the agent's responsibility maybe the factor which activates the other MD strategies and aggressive behaviour. This hypothesis is based on the fact that minimizing responsibility tends to occur more often in first-time or sporadic aggression (Pozzoli et al., 2012) and the other strategies (cognitive restructuring, distortion of consequences and dehumanisation of the victim) seem to require a greater recognition of involvement in aggression, and therefore a higher level of complexity in order to reconceptualise a situation. Thus, some cross-sectional studies have pointed out how the minimisation of responsibility may be more marked in those who are less active in bullying, while the distortion of consequences or other mechanisms within cognitive restructuring, such as euphemistic labelling, are more often found in pure aggressors, who take part in bullying more frequently (Runions et al., 2019). In addition, we also hypothesized that participation in bullying episodes may mediate the activation of these more complex strategies, which may often begin with a simple minimisation of responsibility. However, to date, this has not been proven in a longitudinal study. If this process were confirmed, it would seem to support the progressive nature of the MD mechanisms explained by Bandura (2016), but not yet tested, within the process of moral self-regulation that the author described in a three-step model, where the role of the agent plays a prominent role, thus corroborating the selective nature of MD strategies (Thornberg et al., 2020). Exploring the sequential progress of moral disengagement strategies and bullying behaviour is of particular interest for school violence prevention programs. Although these practical programs already focus on

moral disengagement to prevent bullying (Cross et al., 2015; Wang & Goldberg, 2017), this study could provide useful information to understand how moral judgement can be eroded by the progressive increase in false reasoning based on moral self-deception, and how this, in turn, can support the idea of a progressive time sequence of aggression and the repetition of aggressive behaviour in bullying. Understanding this process may allow us to strengthen prevention and intervention programmes by focusing on certain strategies which tend to activate the rest.

### The Social Cognitive Theory and the transformative power of moral disengagement

According to Bandura (1986), moral agency exerts an intentional, more or less rational but always challenging, influence on a person's civic behaviour, and this allows the individual to tell the difference between right and wrong. However, on many occasions, when human beings are subjected to strong peer pressure and/or different personal variables, they often resort to cognitive self-deception in order to evade the emotions or moral reproaches implicit in the act of transgressing the ethical norm, while at the same time accruing personal benefits from the act (Schwardmann & van der Weele, 2019). The Social Cognitive Theory describes how moral agency is influenced by a number of motivational, emotional and cognitive processes which invariably depend on the interaction between three elements: the socio-moral development which individuals undergo and is influenced by biological and hereditary factors, their behaviour when faced with certain environmental situations, and the knowledge with which the pervading culture and context influence this interpersonal dynamic. In particular, on an individual level, the moral agency theory underlines the existence of dual aspects of self-regulation: proactive and inhibitive. The proactive process leads the individual to comply with norms and conventions through self-sanctions, thus ensuring positive social behaviour. The inhibitive process acts by cancelling out the proactive one, by rejecting self-sanctions and triggering other lines of thought which evade self-reproach. The proactive process, which Bandura (2016) described as fear of public sanction (external control) or fear of emotions of shame and guilt (internal control), can therefore be partially neutralized by the inhibitive process.

The author identifies four MD strategies or mechanisms of self-incriminating thoughts based on the control locus they use: (1) In *cognitive restructuring*, the thoughts that inhibit self-sanctioning are directed at normalizing, rationalizing, or justifying the behaviour displayed. This group of strategies includes moral justification, euphemistic labelling, and advantageous comparison; (2) *Minimizing responsibility or the role of causal agent*, which involves the activation of thoughts which tend to evade the responsibility of guilt by

displacement or diffusion; (3) *Distortion of consequences*, which involves the decentralization of the outcomes or consequences of the behaviour displayed by the aggressor/agent, which are then either ignored or lessened; and finally, (4) *Dehumanization*, which labels the victim as guilty or deserving of the attacks, due to their idiosyncratic characteristics or behaviour. Bandura (1986) situated these four MD strategies in a three-step process, through which moral self-regulation develops: reprehensible behaviour, harmful effects, victim. The agent locus plays a prominent role, since it is not directly related to any of these three steps, unlike the rest, although it is linked to two of them (reprehensible behaviour and harmful effects) involving cognitive restructuring and consequence distortion strategies, which has led some authors to reconceptualize this theory, to give a more prominent role to the reconstruction of moral agency itself (Schaefer & Bouwmeester, 2021). However, following the scheme proposed by Bandura (1986, p. 376), it might be asked what role this locus plays in the onset of aggressive behaviour, which is directly related to the first two steps of the process of moral self-regulation. In other words, increasing the MD of the agent locus, which is related to two of the steps of the self-regulation process, could increase the other MDs linked to these steps. However, most of studies have only focused on the concept of MD as unidimensional and have not considered the sequential process of self-regulation contemplated in this study.

Bandura (2016) also asserts the gradual nature of the use of MD mechanisms and antisocial behaviour. In other words, MD does not instantly transform civic and moral people into cruel and amoral individuals, but rather, the inhibition of civic behaviour is involved in a modelling process that depends largely on the experiences of the individual in their social context. It can therefore be assumed that there is a progressive withdrawal of the proactive control over transgressive acts, through MD mechanisms that inhibit this behaviour for fear of external sanctions. This occurs over time, depending on the reinforcing or punishing response detected in the pervading environment, and it may end up evading reproach for the most serious transgressions (Bandura, 2016). At this extreme, antisocial behaviour can become dangerously routine.

### Minimizing responsibility in traditional bullying

Not assuming one's responsibility when causing someone harm prevents the activation of moral control and therefore avoids a fair and ethical argument (Bandura, 2016). This MD mechanism puts the moral agency locus beyond the moral control of the aggressor (Romera et al., 2021b). According to Bandura (2016), this MD strategy can be manifested in two ways: (a) *displacement of responsibility*, by attributing it to third parties, and (b) *diffusion of responsibility*, by spreading

the responsibility to other people. A number of qualitative studies show that schoolchildren frequently complain that the school authorities do not take any steps to curb the injustice of bullying, and that they are therefore partly responsible for it (Cunningham et al., 2016), and this often leads the rest to assume that bullying is accepted at the school (Thornberg & Delby, 2019). In addition, given that bullying is a complex phenomenon in which the various actors take on different roles (Salmivalli, 2010), diffusion of responsibility allows the aggressor to distribute the blame for instigating the events among several individuals, and also to diffuse the blame and responsibility by placing them outside their moral control. The established theory of deindividuation (Festinger et al., 1952) has already showed how people, when they see themselves as part of a group and not as autonomous individuals, are more likely to display transgressive behaviour. This can be especially dangerous, since studies on group violence show that groups can commit more serious offences than those committed individually (Lantz, 2021).

Thornberg et al. (2020), in a study using hypothetical situations with vignettes, showed that group conformity and a laughing audience (common scenarios of bullying) lead to a greater degree of displacement of responsibility. The authors concluded that bullies are more likely to experience diffusion of responsibility when many people are involved in the aggression, and an important role is attributed to the witnesses and bystanders (Saarento & Salmivalli, 2015). Thus, studies on minimizing responsibility point out that this strategy of moral disengagement is widely used, especially when displacement and diffusion of responsibility are analysed separately (Méndez et al., 2020), but also when the analysis is carried out jointly (Pozzoli et al., 2012). It remains to be seen whether the strategy of minimizing responsibility, through which the aggressor's moral agency uses the group or another person to avoid taking on responsibility for the transgression, can increase the use of other MD strategies over time. If the agent does not assume their responsibility and minimizes it, it follows that they might resort to new, more extensive disengagement mechanisms to continue avoiding guilt or shame.

### Bullying as a possible mediator in moral disengagement strategies

A number of meta-analyses of this aspect of bullying have pointed out that MD is common among schoolchildren involved in the dynamics of aggression. For example, in the meta-analysis by Killer et al. (2019) based on Bandura's (1986) Social Cognitive Theory, they examined the relationship of MD with the roles involved in traditional bullying. In this work, which included 47 studies, they found that there was a positive relationship with the roles of aggressor and victim, a negative relationship with the defender, and

no relationship at all with bystanders. Among the studies analysed, they found some in which a decrease in MD produced a lower level of victimisation. Another meta-analysis, de Oliveira et al. (2021), which included 15 articles, found a relationship between MD and bullying perpetration, and therefore emphasised the relevance of working on these components in bullying prevention and intervention programmes. In the meta-analysis by Gini et al. (2014), on a sample of 27 studies, they found that there was a positive relationship between MD and the different types of aggression in bullying, and that this was more marked in adolescents. However, most of studies analysed MD as a one-dimensional construct, and only few considered the differences between the four strategies, focusing mainly on the direction of MD as a risk factor of bullying, and not vice versa, despite its bidirectional nature (Bandura, 2016).

Additionally, most of the work analysing MD mechanisms or strategies to date has been cross-sectional (Bjärehed et al., 2020), avoiding the analysis of these relationships in their evolutionary dimension, which is precisely what can shed light on their socializing effect. Nevertheless, despite these weaknesses, most of the previous works have shown that the four strategies are related to aggression in bullying (Romera et al., 2021b). Research into the opposite tendency, which explores the impact of bullying aggressiveness on the tendency to MD, has been very limited, and the few published works show contradictory tendencies. Some works have found a link (Obermann, 2013), while others did not find any associations (Barchia & Bussey, 2011).

Although the (less numerous) longitudinal studies may provide a causal relationship between the variables, they have mainly focused on the variable of MD in a unidimensional way, pointing to a strong relationship between MD and aggression in bullying (Bjärehed et al., 2021), and also in the opposite direction (Romera et al., 2021a; Thornberg et al., 2019), although some studies found no such relationship (Wang et al., 2017). However, no work has explored the separate, individual relationships between the different mechanisms of MD, even though previous longitudinal studies that explored the mediating effect of MDs and bullying found different effects depending on the type of strategy (Falla et al., 2021, 2022). These results support the need to analyse from a longitudinal perspective the relationship between MD strategies and how bullying involvement may impact on the process of moral self-regulation.

## The present study

Although most of the studies that have linked MD and bullying have been based on Bandura's (1986) Social Cognitive Theory (see meta-analysis by Killer et al., 2019), so far there has been no deepening of the moral self-regulation process described by the author based on a temporal sequence of MD

strategies. As a result, it has not been not proven whether certain strategies can be linked to others in the long run. This information may be crucial to understand some of the characteristics of the three-step model that support a linear development of MD. It is essential to understand the process of moral self-regulation at this stage in order to prevent minor aggressions from escalating into violence and being sustained over time (Nasaescu et al., 2020). Indeed, this information is of fundamental importance for the practical design of prevention and intervention programmes against bullying, and can help to build a fairer and more sustainable moral criterion. Thus, the aims of this study are: first, to explore whether minimizing responsibility can be directly related to the other strategies of MDs, one year later; and secondly, to examine whether aggression in bullying at Time 2 is a mediator of the relation between minimizing responsibility and the other MD strategies.

These objectives are based on the following hypotheses:

Based on the three-step process of moral self-regulation explained by Bandura in Social Cognitive Theory (1986, p. 376), this study hypothesised that the arguments of minimizing responsibility in the agent locus, which in turn is the most frequent in schools (Pozzoli et al., 2012) and among less active schoolchildren in bullying perpetration (Runions et al., 2019), will be related to the strategies of cognitive restructuring and distortion of consequences one year later, following the scheme of Bandura's process of moral self-regulation and supporting the studies which claim that these mechanisms occur among the most ruthless bullies (Runions et al., 2019), and the progressive character of the Social Cognitive Theory (Bandura, 2016). However, we also hypothesised that it will not be directly related to dehumanisation, as these factors are not related in the three-step model of moral self-regulation, and there are no empirical studies to confirm it. Therefore, the first hypothesis will be:

H1. Minimizing responsibility at Time 1 will be related to cognitive restructuring, distortion of consequences and dehumanization at Time 3, if the dependent variables at Time 1 are controlled.

Considering that MD strategies have been linked to bullying perpetration (Romera et al., 2021b) and bullying perpetration has been associated with all four strategies (Falla et al., 2021), we consider that bullying perpetration may exert a mediating effect on the relationship established among the MD strategies, and so also hypothesised that:

H2. Aggression in bullying will mediate the relationship between minimizing responsibility and the other MD strategies (cognitive restructuring, distortion of consequences and dehumanization), if the dependent variables at Time 1 are controlled.

## Methods

### Participants

We carried out a longitudinal study out using non-probabilistic sampling for accessibility (Singleton & Straits, 2004), with 8th and 9th grade pupils from thirteen secondary schools. At Time 1 (T1), we surveyed a total of 1167 pupils aged between 12 and 16 (53% girls;  $M_{\text{age}} = 13.55$ ;  $SD = 0.770$ ). The second survey (T2) took place six months later, with a sample of 1140 pupils (53.4% girls;  $M_{\text{age}} = 14.04$ ;  $SD = 0.852$ ). Time 3 (T3) was one year after the first collection, and the final sample contained 1107 pupils (54.7% girls;  $M_{\text{age}} = 14.49$ ;  $SD = 0.789$ ) aged between 12 and 17, with a final retention rate of 94.86%. Logistic regression was used to verify the representativeness of the analytical longitudinal sample in relation to the total sample. The results show that there were no significant differences in the study variables between any of the three times ( $p > .05$ , in all cases).

### Instruments

Minimizing responsibility (MR) and the other moral disengagement strategies (cognitive restructuring, distortion of consequences and dehumanization) were measured using the *Mechanisms of Moral Disengagement Scale* (Caprara et al., 1996) and adapted to Spanish by Romera et al. (2023). This questionnaire is made up of 24 items which cover the four MD strategies, and each question has five Likert-type response options with values ranging from one to five (1 = strongly disagree; 2 = partly agree; 3 = I agree in general; 4 = strongly agree; 5 = totally agree). The MR strategy had 6 items (e.g., “If kids are living under bad conditions, they cannot be blamed for behaving aggressively.”). The rest of the strategies were grouped as follows: cognitive restructuring contained nine items (e.g., “To hit obnoxious classmates is just giving them ‘a lesson’”), distortion of consequences consisted of three items (e.g., “Insults among children do not hurt anyone”), and dehumanization was made up of six items (e.g., “Someone who is obnoxious does not deserve to be treated like a human being”). Internal consistency was measured using the omega coefficient (McDonald, 2013). The results showed acceptable values for MR at the three time points measured ( $\omega_{T1} = .67$ ,  $\omega_{T2} = .70$ ,  $\omega_{T3} = .72$ ). Similarly, the data indicated acceptable and good statistics for the other MD strategies: cognitive restructuring ( $\omega_{T1} = .80$ ,  $\omega_{T2} = .83$ ,  $\omega_{T3} = .82$ ); distortion of consequences ( $\omega_{T1} = .60$ ,  $\omega_{T2} = .64$ ,  $\omega_{T3} = .60$ ); and dehumanization ( $\omega_{T1} = .71$ ,  $\omega_{T2} = .77$ ,  $\omega_{T3} = .77$ ).

We used the European Bullying Intervention Project Questionnaire (EBIPQ) (Ortega-Ruiz et al., 2016) to measure the aggressiveness manifested in bullying. In this scale, the seven items refer to the last two months (for example, “I have threatened someone”) with a Likert scale ranging from 0 to 4 (0 = no; 1 = yes, once or twice; 2 = yes, once or twice a month; 3 = yes, about once a week; 4 = yes, more than once a week). Internal consistency was checked using omega coefficients. The values obtained for the three time points were good: ( $\omega_{T1} = 0.80$ ,  $\omega_{T2} = 0.80$ ,  $\omega_{T3} = 0.80$ ).

### Procedure

Our study forms part of a project approved by the Bioethics and Biosafety Committee at the authors’ university. The principal investigators contacted the schools, which, in turn, consulted the school council to allow the study to be carried out there. Written consent was given by the pupils’ families. The researchers came to the school on the agreed date, collected the authorizations from the families, and administered the questionnaires directly in the classrooms. Before the pupils began to fill in the questionnaires, they were informed that the data used in the study were voluntary, private and confidential, and they were instructed how to fill in the identification code, as this was a longitudinal study which was repeated on two further occasions, six months apart. Using this code, it was possible to identify the questionnaires while maintaining the pupils’ anonymity. The study was part of a broader project which included several questionnaires, in addition to those used in this work, and the approximate time allowed to fill them out was 45 min.

### Data analysis

The statistical and inferential analyses were performed using IBM SPSS Statistics version 26 (IBM, Armonk, NY, USA), which included means, standard deviations, and Student’s t-test to verify the existence of differences by sex. We also used Cohen’s d test to estimate the effect size. Sizes below 0.50 were considered as weak, between 0.50 and 0.80 moderate, and over 0.80 strong (Cohen, 2013). We used Pearson’s correlation to verify the link between the different MD strategies, aggressiveness in bullying and age at the three time points to find out if collinearity existed between the variables.

Next, the mediation analyses were carried out using the PROCESS v3.4 macro for SPSS (SPSS Inc., Chicago, IL, USA), following model 4 (Hayes, 2013) and the four-step procedure described by MacKinnon (2008). In all three models, MR was taken as a predictor variable in T1 and aggressiveness in bullying in T2 was taken as a mediator. The dependent variable for the first analysis was cognitive

restructuring at T3 (model 1), for the second analysis it was distortion of consequences at T3 (model 2), and for the third analysis, dehumanization at T3 (model 3). All the analyses checked for sex, age, and the dependent variable at T1, i.e., cognitive restructuring at T1 (in model 1), distortion of consequences at T1 (in model 2), and dehumanization at T1 (in model 3). In addition, we used the bootstrap method to find the indirect effects of the mediation analysis, assuming significance when the values did not include zero between the confidence intervals. Finally, the mediation effect was calculated through the relationship between the indirect effect and the total (Wen & Fan, 2015).

## Results

### Preliminary analysis

Table 1 shows the means and standard deviations for the four MD strategies and aggressiveness in bullying at the three time points. We found that MR was the MD strategy that presented the highest mean over the three time points. Pearson's correlation verified that there was no multicollinearity between the study variables.

Student's t-test indicated that boys reported significantly higher scores than girls in bullying for all three time points measured, although the effect sizes were weak. For MR, there were no significant differences between boys and girls at Times 1 and 3, but there were at Time 2, with a weak effect size. Also, in the rest of the MD strategies (cognitive restructuring, distortion of consequences and dehumanization), boys scored higher than girls at the three time points measured, with effect sizes ranging between weak and moderate (see Table 2).

### Cognitive restructuring model

According to the data, the model which examined the relationship between MR at T1 for cognitive restructuring at T3, mediated by aggressiveness in bullying at T2, was significant  $F(5, 964) = 119.03$ ;  $R^2 = 0.38$ ;  $p < .001$ . MR at T1 was positively related to cognitive restructuring at T3 ( $\beta = 0.09$ ,  $t = 3.17$ ,  $p < .01$ ). In addition, positive associations were found between the MR MD strategy at T1 and aggressiveness in bullying at T2 ( $\beta = 0.10$ ,  $t = 3.01$ ,  $p < .01$ ), and between the latter and cognitive restructuring at T3 ( $\beta = 0.20$ ,  $t = 7.11$ ,  $p < .001$ ). We also found that the total effect on the relationship between MR at T1 and cognitive restructuring at T3 was significant and direct ( $\beta = 0.11$ ,  $t = 3.78$ ,  $p < .001$ ) (see Fig. 1). Finally, we found an indirect effect of MR in T1 on cognitive restructuring in T3 through the channel of aggressiveness in bullying T2 ( $\beta = 0.02$ , IC

del 95% = [0.003, 0.042]), using the bias-corrected percentile bootstrap method, which therefore produced a median effect of 17.8%.

### Distortion of consequences model

The second of the mediation analyses allowed us to check the significance of the model which measured the relationship between MR in T1 on the distortion of consequences in T3, with aggressiveness in bullying in T2 as a mediating variable:  $F(5, 977) = 50.44$ ;  $R^2 = 0.21$ ;  $p < .001$ . The data showed a direct relationship between MR in T1 and distortion of consequences in T3 ( $\beta = 0.07$ ,  $t = 2.21$ ,  $p < .05$ ). We also found direct, significant relationships between MR in T1 and aggressiveness in bullying in T2 ( $\beta = 0.16$ ,  $t = 4.86$ ,  $p < .001$ ) and between aggressiveness and distortion of consequences in T3 ( $\beta = 0.21$ ,  $t = 6.83$ ,  $p < .001$ ). Next, we calculated the total effect in the relationship between MR in T1 and distortion of consequences in T3 ( $\beta = 0.11$ ,  $t = 3.23$ ,  $p < .01$ ) (see Fig. 2). The indirect effect was also verified in the relationship between MR in T1 and distortion of consequences in T3 through aggressiveness in bullying in T2 ( $\beta = 0.03$ , IC del 95% = [0.015, 0.058]), with the mediation effect being 32.2%.

### Dehumanization model

The third model examined the relationship between MR at T1 and dehumanization at T3, using aggressiveness in bullying at T2 as a mediator. The results found were significant:  $F(5, 965) = 80.61$ ;  $R^2 = 0.29$ ;  $p < .001$ . No direct relationship was found between MR at T1 and dehumanization at T3; instead, the analyses found a significant relationship between MR at T1 and aggressiveness in bullying at T2 ( $\beta = 0.12$ ,  $t = 3.60$ ,  $p < .001$ ) and between the latter and dehumanization at T3 ( $\beta = 0.16$ ,  $t = 5.32$ ,  $p < .001$ ). There was also a total effect in the relationship between MR at T1 and dehumanization at T3 ( $\beta = 0.08$ ,  $t = 2.42$ ,  $p < .05$ ) (see Fig. 3). Finally, an indirect effect was found in the relationship between MR at T1 and dehumanization at T3 via aggressiveness in bullying at T2 ( $\beta = 0.02$ , IC del 95% = [0.006, 0.040]). The mediation effect was 25.1%.

## Discussion

The cognitive and affective complexity of the processes of socialization, by which schoolchildren acquire sensitivity and moral criteria, especially when faced with the intentional and sustained bullying of many schoolchildren towards others, has inspired a large number of works that have shown that MD strategies are the cognitive processes used by bullying aggressors to evade self-reproach (Killer

**Table 1** Means, standard deviations and correlations among all study variables

	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
1. MR1	(1–5)	1.92	0.75															
2. BP T1	(0–4)	0.25	0.41	0.246**														
3. CR T1	(1–5)	1.57	0.57	0.437**	0.453**													
4. DC T1	(1–5)	1.41	0.58	0.409**	0.375**	0.581**												
5. DH T1	(1–5)	1.49	0.58	0.427**	0.399**	0.640**	0.524**											
6. MR T2	(1–5)	1.82	0.71	0.415**	0.175**	0.308**	0.265**	0.309**										
7. BP T2	(0–4)	0.28	0.41	0.203**	0.524**	0.334**	0.189**	0.278**	0.265**									
8. CR T2	(1–5)	1.58	0.59	0.321**	0.339**	0.567**	0.328**	0.450**	0.558**	0.485**								
9. DC T2	(1–5)	1.42	0.58	0.285**	0.298**	0.339**	0.370**	0.331**	0.509**	0.375**	0.612**							
10. DH T2	(1–5)	1.53	0.63	0.297**	0.287**	0.414**	0.278**	0.540**	0.542**	0.377**	0.734**	0.579**						
11. MR T3	(1–5)	1.74	0.69	0.382**	0.214**	0.263**	0.238**	0.278**	0.490**	0.184**	0.310**	0.297**	0.297**					
12. BP T3	(0–4)	0.20	0.36	0.135**	0.406**	0.280**	0.188**	0.211**	0.197**	0.469**	0.250**	0.229**	0.227**	0.253**				
13. CR T3	(1–5)	1.55	0.57	0.298**	0.344**	0.526**	0.302**	0.425**	0.348**	0.362**	0.621**	0.401**	0.465**	0.510**	0.397**			
14. DC T3	(1–5)	1.39	0.56	0.243**	0.317**	0.348**	0.362**	0.327**	0.317**	0.299**	0.406**	0.472**	0.351**	0.511**	0.354**	0.638**		
15. DH T3	(1–5)	1.49	0.58	0.261**	0.311**	0.373**	0.288**	0.498**	0.358**	0.310**	0.489**	0.358**	0.577**	0.540**	0.337**	0.704**	0.632**	
16. Age				-0.005	0.048	0.090**	0.018	0.063*	0.011	0.055	0.057	0.021	0.029	-0.005	0.031	0.138**	0.054	0.069*

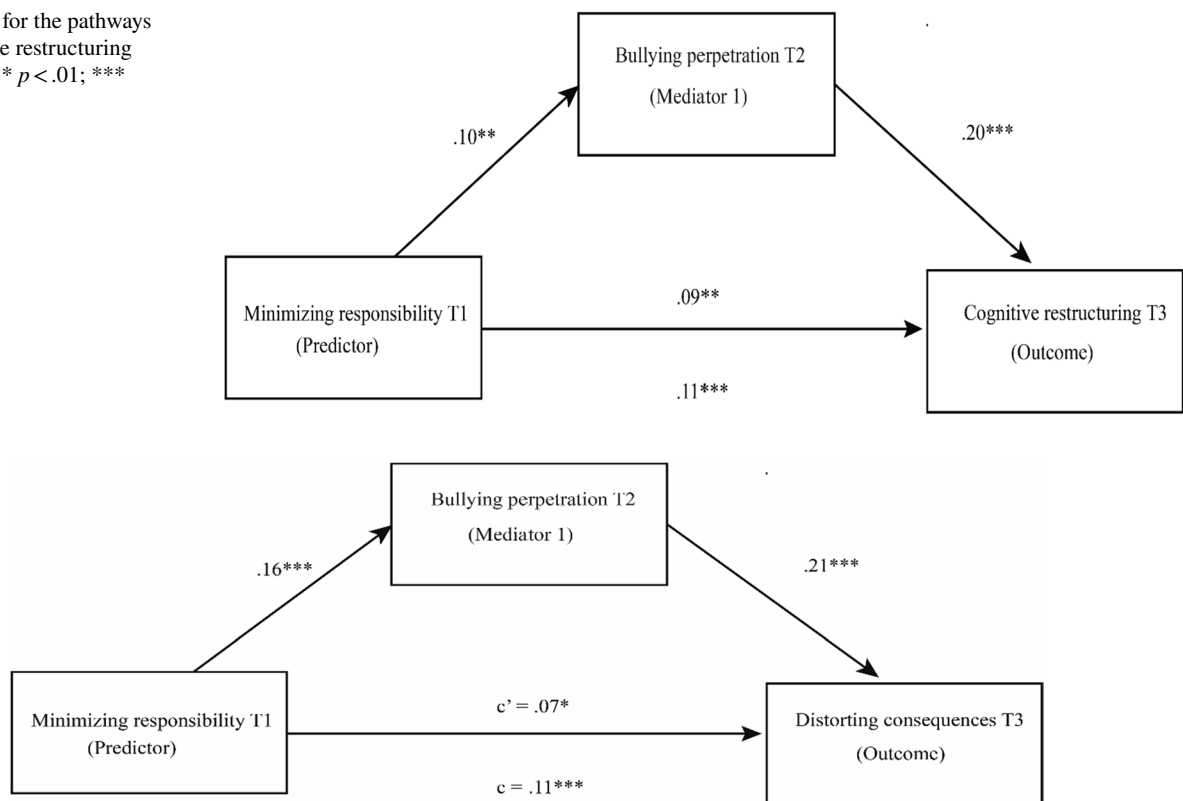
M = mean; SD = Standard deviation; MR = Minimizing responsibility; BP = Bullying perpetration; CR = Cognitive restructuring; DC = Distorting consequences; DH = Dehumanizing; \*  $p < .05$ ; \*\*  $p < .001$

**Table 2** Differences by gender for all variables

		Boys		Girls		<i>t</i>	<i>d</i>
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
MR T1	(1–5)	1.96	0.74	1.88	0.75	1.721	0.11
BP T1	(0–4)	0.32	0.49	0.20	0.32	4.666***	0.30
CR T1	(1–5)	1.75	0.63	1.42	0.46	9.611***	0.61
DC T1	(1–5)	1.52	0.65	1.31	0.50	5.081***	0.37
DH T1	(1–5)	1.64	0.67	1.36	0.47	7.638***	0.49
MR T2	(1–5)	1.89	0.75	1.75	0.66	3.142**	0.20
BP T2	(0–4)	0.34	0.47	0.23	0.35	4.299***	0.27
CR T2	(1–5)	1.76	0.65	1.43	0.50	9.113***	0.58
DC T2	(1–5)	1.55	0.69	1.31	0.44	6.700***	0.42
DH T2	(1–5)	1.69	0.74	1.40	0.49	7.350***	0.47
MR T3	(1–5)	1.79	0.72	1.71	0.66	1.856	0.12
BP T3	(0–4)	0.24	0.42	0.16	0.29	3.852***	0.23
CR T3	(1–5)	1.72	0.64	1.40	0.45	9.379***	0.59
DC T3	(1–5)	1.50	0.65	1.30	0.45	5.800***	0.36
DH T3	(1–5)	1.61	0.67	1.39	0.48	6.053***	0.38

*M*=mean; *SD*=Standard deviation; *t*=Student’s *t*; *d*=Cohen’s *d*; MR=Minimizing responsibility; BP=Bullying perpetration; CR=Cognitive restructuring; DC=Distorting consequences; DH=Dehumanizing; \**p*<.05; \*\**p*<.01; \*\*\* *p*<.001

**Fig. 1** Results for the pathways of the cognitive restructuring model. Note: \*\* *p*<.01; \*\*\* *p*<.001



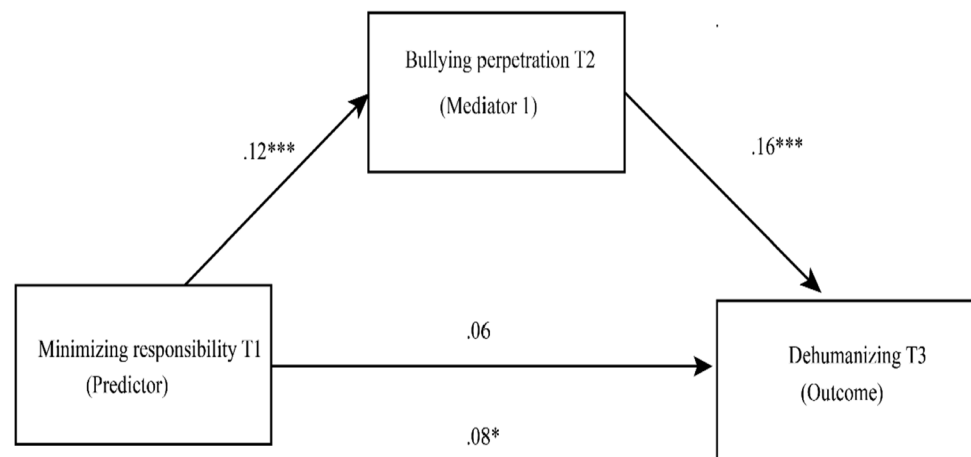
**Fig. 2** Results for the pathways of the distortion of consequences model. Note: \*\* *p*<.01; \*\*\* *p*<.001

et al., 2019). However, more research is needed into how these MD strategies work, to find an explanation for the interaction between these mechanisms and the equally

complex social dynamics of aggression in bullying. The aim of this work was, therefore, to explore minimizing responsibility, a strategy used by aggressors to avoid



**Fig. 3** Results for the pathways of the dehumanization model.  
 Note: \*\*  $p < .01$ ; \*\*\*  $p < .001$



assuming direct responsibility, which is shared with or transferred to others, and to verify whether this affects the other MD mechanisms mediated by aggressive behaviour in bullying.

The results show that the schoolchildren obtained higher mean scores in minimizing responsibility compared to the other MD strategies, as shown in previous studies (Romera et al., 2021b). Different works have indicated that from childhood onwards, children use the tactic of blaming others to evade external sanctions (Banerjee et al., 2020) or may even find it hard to assume their own responsibility for immoral, aggressive actions (Recchia et al., 2012). In line with these results, the present study indicates that minimizing responsibility occurs more often in children than the other MD strategies. In other words, as Bussey (2020) pointed out, it seems that it is more common, in evolutionary terms, for children to start using minimizing responsibility as a cognitive strategy to escape the assumption that they have broken the rules and to evade feelings of guilt or shame. However, it remains to be proved whether this MD strategy, which Bandura (1986) situated as the first step of moral self-regulation and outside the other steps of the moral self-regulation sequence (reprehensible behaviour, harmful effects, victim), could be a direct predictor, one year later, of the three other strategies (cognitive restructuring, distortion of consequences and dehumanization) and thereby prove that the process of moral self-regulation is progressive and sequential.

The results found in our study confirmed this idea, and suggest that not taking responsibility for an immoral act can be associated, over time, with MD strategies related to behaviour, which lead to aggressor to believe that this act is reasonable or fair, or to ignore or distort the harmful effects they have on the victim. In other words, over time, children who use this strategy tend to make more use of other moral disengagement mechanisms linked to the process of moral self-regulation. This may be because they become more aware of their involvement in these acts, and require

other strategies to enable them to normalize their behaviour and prevent them from seeing the harmful consequences it has on the victims. In contrast, in this study, we found no direct relationship with the locus centred on the role of the victim. Bandura (1986) also found this result in the representation of the self-regulation process, and so the theoretical model can be empirically confirmed. In addition, although this process of aggressive behaviour remains to be verified within the dynamics of bullying, it was used as a mediator in hypothesis two.

The second objective explored whether the relationship between MD strategies could be increased with the mediating effect of aggressiveness in bullying. Here we found an indirect effect in the case of cognitive restructuring and distortion of consequences. Therefore, when minimizing responsibility led to an increase in aggressiveness in bullying, this, in turn, caused an increase in distortion of consequences and cognitive restructuring. Interestingly, when dehumanization was not directly related to minimizing responsibility, an effect occurred which some authors have termed the Lazarus effect (Walters & Espelage, 2018): here, the indirect effect exerted by aggressiveness on bullying caused two MD strategies which were not related to each other to become related. These data confirm what Bandura (2016) called the transformative power of DM, which is progressive, selective and dependent on the learning context in which social and moral behaviour develops.

The study adds relevant information about how the displacement or diffusion of responsibility can be associated with other self-incriminating strategies of moral harm and how an aggressive context such as bullying can make these strategies increase or even lead to the use of self-deception, which has not previously been used, to continue with the attacks. In other words, greater minimization of responsibility in school contexts can increase bullying aggression, and the intentional, prolonged nature of bullying can cause an increase in the other MD strategies, leading the aggressor to continue transgressing or even to commit even more serious

aggressions, in which the victims are often blamed. These results therefore warn us of the risk of hostile contexts in which there is continued aggressiveness: in these cases, the group itself can construct its own moral criteria, leading to a greater tolerance for unethical arguments in interpersonal relationships between members of the group, and eventually affecting the moral sensitivity and civil behaviour that should prevail in social development at these ages.

### Limitations and future lines of research

Despite being longitudinal, our study only spans one year between the first and the third time points. A study covering a longer period, starting from earlier ages such as late childhood, could provide more robust results about how socio-cognitive processes of inhibition of moral behaviour develop in children. Similarly, expanding the sample to cover other parts of the world could provide important results regarding the cultural component in the construction of moral criteria. However, despite these limitations, this work has explored for the first time the relationship between minimizing responsibility and the other MD strategies and the effect that aggressiveness in bullying can have on this relationship, exploring further the process of moral self-regulation described by Bandura (1986) to explain the phenomenon of traditional bullying. Our results show that this mechanism, which may be more common in school contexts, and which also appears in phenomena of group violence, could lead to an increase in the use of the other MD mechanisms.

### Conclusion

On balance, the findings of this work support the sequential, selective nature and transformative power of progressive MD, as put forward by Bandura (2016). The novel approach of our study has made it possible to verify how an increase in minimizing responsibility, which may occur more commonly among schoolchildren, leads to an increased use of the other MD strategies, mediated by aggressiveness in bullying. These findings justify the need to carry out future research which may allow us to further our understanding of the selective use and sequences of moral disengagement strategies in situations of abuse, exclusion or violence between peers. Similarly, the data show the importance of working on the subject of moral responsibility in schools, with the aim of fostering children's moral development and constructing a common group ethic.

**Author contributions** All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by [Daniel Falla] and [Eva M. Romera]. The first draft of the manuscript was written by [Daniel Falla] and [Rosario Ortega-Ruiz] and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

**Funding** Funding for open access publishing: Universidad de Córdoba/CBUA. This study was supported by MCIN/AEI/<https://doi.org/10.13039/501100011033> and European Union "NextGenerationEU"/PRTR (PDC2021-121741-I00).

**Data availability** The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

### Declarations

**Conflicting interests** The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

### References

- Bandura, A. (1986). *Social foundations of thought and action: a social cognitive theory*. Prentice-Hall.
- Bandura, A. (2016). *Moral disengagement: How people do harm and live with themselves* (Worth Publishers (ed.)). Retrieved April 11, 2022, from, <https://psycnet.apa.org/record/2015-43532-000>
- Banerjee, R., Heyman, G. D., & Lee, K. (2020). The development of moral self-presentation. In L. A. Jensen (Ed.), *The Oxford Handbook of Moral Development* (pp. 91–109). Oxford University Press.
- Barchia, K., & Bussey, K. (2011). Individual and collective social cognitive influences on peer aggression: exploring the contribution of aggression efficacy, moral disengagement, and collective efficacy. *Aggressive Behavior*, 37(2), 107–120. <https://doi.org/10.1002/ab.20375>
- Bjärehed, M., Thornberg, R., Wänström, L., & Gini, G. (2020). Mechanisms of moral disengagement and their associations with indirect bullying, direct bullying, and pro-aggressive bystander behavior. *Journal of Early Adolescence*, 40(1), 28–55. <https://doi.org/10.1177/0272431618824745>
- Bjärehed, M., Thornberg, R., Wänström, L., & Gini, G. (2021). Moral disengagement and verbal bullying in early adolescence: a three-year longitudinal study. *Journal of School Psychology*, 84, 63–73. <https://doi.org/10.1016/j.jsp.2020.08.006>

- Bussey, K. (2020). Development of Moral Disengagement: learning to make wrong right. In L. A. Jensen (Ed.), *The Oxford Handbook of Moral Development* (pp. 305–326). Oxford University Press.
- Caprara, G. V., Barbaranelli, C., Vicino, S., & Bandura, A. (1996). La misura del disimpegno morale in età evolutiva. *Età Evolutiva*, 51, 18–29.
- Cohen, J. (2013). Statistical power analysis for the behavioral Sciences. *Statistical power analysis for the behavioral Sciences*. Routledge.
- Cross, D., Barnes, A., Papageorgiou, A., Hadwen, K., Hearn, L., & Lester, L. (2015). A social-ecological framework for understanding and reducing cyberbullying behaviours. *Research Outputs 2014 to 2021*, 23, 109–117. <https://doi.org/10.1016/j.avb.2015.05.016>
- Cunningham, C. E., Mapp, C., Rimas, H., Cunningham, L., Mielko, S., Vaillancourt, T., & Marcus, M. (2016). What limits the effectiveness of antibullying programs? A thematic analysis of the perspective of students. *Psychology of Violence*, 6(4), 596. <https://doi.org/10.1037/A0039984>
- de Oliveira, W. A., da Silva, J. L., Risk, E. N., Komatsu, A. V., Silva, M. A. L., & dos Santos, M. A. (2021). Bullying and mechanisms of moral disengagement: systematic literature review with meta-analysis. *Psicologia Escolar e Educacional*, 25, 1–9. <https://doi.org/10.1590/2175-35392021223346>
- Falla, D., Romera, E. M., & Ortega-Ruiz, R. (2021). Aggression, moral disengagement and empathy. A longitudinal study within the interpersonal dynamics of bullying. *Frontiers in Psychology*, 0, 3981. <https://doi.org/10.3389/fpsyg.2021.703468>
- Falla, D., Ortega-Ruiz, R., Runions, K., & Romera, E. M. (2022). Why do victims become perpetrators of peer bullying? Moral disengagement in the cycle of violence. *Youth & Society*, 54(3), 397–418. <https://doi.org/10.1177/0044118X20973702>
- Festinger, L., Pepitone, A., & Newcomb, T. (1952). Some consequences of de-individualisation in a group. *The Journal of Abnormal and Social Psychology*, 47(2), 382–389. <https://doi.org/10.1037/h0057906>
- Gini, G., Pozzoli, T., & Hymel, S. (2014). Moral disengagement among children and youth: a meta-analytic review of links to aggressive behavior. *Aggressive Behavior*, 40(1), 56–68. <https://doi.org/10.1002/ab.21502>
- Hayes, A. F. (2013). *Introduction to Mediation, Moderation, and conditional process analysis*. Guilford.
- 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(4), 450–462. <https://doi.org/10.1002/ab.21833>
- Lantz, B. (2021). The consequences of crime in company: co-offending, victim–offender relationship, and quality of violence. *Journal of Interpersonal Violence*, 36(7–8), NP4363–NP4388. <https://doi.org/10.1177/0886260518786497>
- MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. Taylor & Francis Group.
- McDonald, R. P. (2013). Test theory: a unified treatment. *Test Theory*. Psychology Press. <https://doi.org/10.4324/9781410601087>
- Méndez, I., Liccardi, G., & Ruiz-Esteban, C. (2020). Mechanisms of Moral Disengagement Used to Justify School Violence in Sicilian Primary School. *European Journal of Investigation in Health, Psychology and Education*, 10, 682–690. <https://doi.org/10.3390/EJIHPE10030050>
- 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. <https://doi.org/10.1016/j.jadohealth.2014.06.007>
- Moore, S. E., Norman, R. E., Suetani, S., Thomas, H. J., Sly, P. D., & Scott, J. G. (2017). Consequences of bullying victimization in childhood and adolescence: a systematic review and meta-analysis. *World Journal of Psychiatry*, 7(1), 60. <https://doi.org/10.5498/wjp.v7.i1.60>
- Nasaescu, E., Zych, I., Ortega-Ruiz, R., Farrington, D. P., & Llorent, V. J. (2020). Longitudinal patterns of antisocial behaviors in early adolescence: a latent class and latent transition analysis. *European Journal of Psychology Applied to Legal Context*, 12(2), 85–92. <https://doi.org/10.5093/EJPALC2020A10>
- Obermann, M. L. (2013). Temporal aspects of moral disengagement in school bullying: crystallization or escalation? *Journal of School Violence*, 12(2), 193–210. <https://doi.org/10.1080/15388220.2013.766133>
- Ortega-Ruiz, R. (2010). *Agresividad injustificada, bullying y violencia escolar*. Alianza Editorial.
- Ortega-Ruiz, R., Del Rey, R., & Casas, J. A. (2016). Assessing bullying and cyberbullying: spanish validation of EBIPQ and ECIQP. *Psicologia Educativa*, 22(1), 71–79. <https://doi.org/10.1016/j.pse.2016.01.004>
- Pozzoli, T., Gini, G., & Vieno, A. (2012). Individual and class moral disengagement in bullying among elementary school children. *Aggressive Behavior*, 38(5), 378–388. <https://doi.org/10.1002/ab.21442>
- Recchia, H. E., Brehl, B. A., & Wainryb, C. (2012). Children’s and adolescents’ reasons for socially excluding others. *Cognitive Development*, 2(27), 195–203. <https://doi.org/10.1016/J.COGDEV.2012.02.005>
- Romera, E. M., Ortega-Ruiz, R., Runions, K., & Camacho, A. (2021a). Bullying perpetration, moral disengagement and need for popularity: examining reciprocal associations in adolescence. *Journal of Youth and Adolescence* 2021, 50:10(10), 2021–2035. <https://doi.org/10.1007/S10964-021-01482-4>
- Romera, E. M., Ortega-Ruiz, R., Runions, K., & Falla, D. (2021b). Moral disengagement strategies in online and offline bullying. *Psychosocial Intervention*, 30(2), 85–93. <https://doi.org/10.5093/pi2020a21>
- Romera, E. M., Herrera-López, M., Ortega-Ruiz, R., & Camacho, A. (2023). The moral disengagement scale-24: Factorial structure and cross-cultural comparison in Spanish and Colombian adolescents. *Psychology of Violence*, 13(1), 13–22. <https://doi.org/10.1037/VIO0000428>
- Runions, K. C., Shaw, T., Bussey, K., Thornberg, R., Salmivalli, C., & Cross, D. S. (2019). Moral disengagement of pure bullies and bully/victims: shared and distinct mechanisms. *Journal of Youth and Adolescence*, 48(9), 1835–1848. <https://doi.org/10.1007/s10964-019-01067-2>
- Saarento, S., & Salmivalli, C. (2015). The role of classroom peer ecology and bystanders’ responses in bullying. *Child Development Perspectives*, 9(4), 201–205. <https://doi.org/10.1111/CDEP.12140>
- Salmivalli, C. (2010). Bullying and the peer group: a review. *Aggression and Violent Behavior*, 15(2), 112–120. <https://doi.org/10.1016/j.avb.2009.08.007>
- Schaefer, U., & Bouwmeester, O. (2021). Reconceptualizing moral disengagement as a process: transcending overly liberal and overly conservative practice in the field. *Journal of Business Ethics*, 172(3), 525–543. <https://doi.org/10.1007/S10551-020-04520-6/FIGURES/1>
- Schwardmann, P., & van der Weele, J. (2019). Deception and self-deception. *Nature Human Behaviour* 2019, 3:10(10), 1055–1061. <https://doi.org/10.1038/s41562-019-0666-7>
- Singleton, R., & Straits, B. (2004). *Approaches to social research*. Oxford University Press.

- Thornberg, R., & Delby, H. (2019). How do secondary school students explain bullying? *Educational Research*, *61*(2), 142–160. <https://doi.org/10.1080/00131881.2019.1600376>
- Thornberg, R., Wänström, L., Pozzoli, T., & Hong, J. S. (2019). Moral disengagement and school bullying perpetration in middle childhood: a short-term longitudinal study in Sweden. *Journal of School Violence*, *18*(4), 585–596. <https://doi.org/10.1080/15388220.2019.1636383>
- Thornberg, R., Daremark, E., Gottfridsson, J., & Gini, G. (2020). Situationally selective activation of moral disengagement mechanisms in school bullying: a repeated within-subjects experimental study. *Frontiers in Psychology*, *0*, 1101. <https://doi.org/10.3389/FPSYG.2020.01101>
- Veiga Simão, A. M., Ferreira, P. C., Freire, I., Caetano, A. P., Martins, M. J., & Vieira, C. (2017). Adolescent cybervictimization – who they turn to and their perceived school climate. *Journal of Adolescence*, *58*, 12–23. <https://doi.org/10.1016/J.ADOLESCENCE.2017.04.009>
- Walters, G. D., & Espelage, D. L. (2018). Resurrecting the empathy–bullying relationship with a pro-bullying attitudes mediator: the lazarus effect in mediation research. *Journal of Abnormal Child Psychology*, *46*(6), 1229–1239. <https://doi.org/10.1007/s10802-017-0355-9>
- Wang, C., & Goldberg, T. S. (2017). Using children’s literature to decrease moral disengagement and victimization among elementary school students. *Psychology in the Schools*, *54*(9), 918–931. <https://doi.org/10.1002/pits.22042>
- Wang, C., Ryoo, J. H., Swearer, S. M., Turner, R., & Goldberg, T. S. (2017). Longitudinal relationships between bullying and moral disengagement among adolescents. *Journal of Youth and Adolescence*, *46*(6), 1304–1317. <https://doi.org/10.1007/s10964-016-0577-0>
- Wen, Z., & Fan, X. (2015). Monotonicity of effect sizes: questioning kappa-squared as mediation effect size measure. *Psychological Methods*, *20*(2), 193–203. <https://doi.org/10.1037/met0000029>

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.