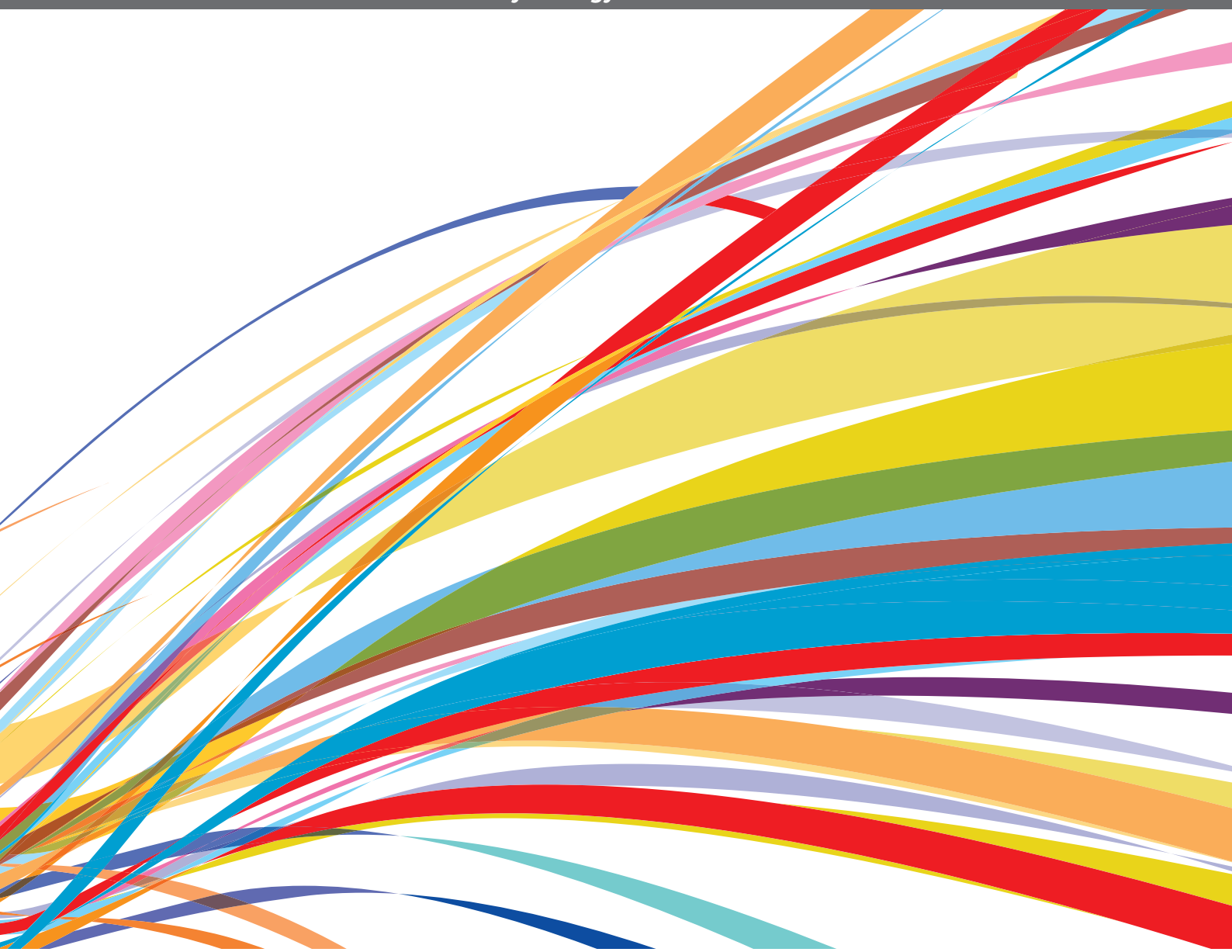
NOVEL DEVELOPMENTAL PERSPECTIVES ON THE LINK BETWEEN MORALITY AND SOCIAL OUTCOMES

EDITED BY: Simona Carla Silvia Caravita, Miriam Beauchamp and Robert Thornberg

PUBLISHED IN: Frontiers in Psychology



frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88974-970-6
DOI 10.3389/978-2-88974-970-6

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

NOVEL DEVELOPMENTAL PERSPECTIVES ON THE LINK BETWEEN MORALITY AND SOCIAL OUTCOMES

Topic Editors:
Simona Carla Silvia Caravita, University of Stavanger, Norway
Miriam Beauchamp, Université de Montréal, Canada
Robert Thornberg, Linköping University, Sweden

Citation: Caravita, S. C. S., Beauchamp, M., Thornberg, R., eds. (2022). Novel Developmental Perspectives on the Link Between Morality and Social Outcomes. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88974-970-6

Table of Contents

frontiers in Psychology

EDITORIAL
published: 30 March 2022
doi: 10.3389/fpsyg.2022.888373

Editorial: Novel Developmental Perspectives on the Link Between Morality and Social Outcomes

Simona C. S. Caravita[1,2*†], *Miriam H. Beauchamp*[3,4†] *and Robert Thornberg*[5†]

[1] Norwegian Centre for Learning Environment and Behavioural Research in Education, University of Stavanger, Stavanger, Norway, [2] Department of Psychology, Catholic University of the Sacred Heart, Milan, Italy, [3] Department of Psychology, University of Montreal, Montreal, QC, Canada, [4] Sainte-Justine Hospital Research Center, Montreal, QC, Canada, [5] Department of Behavioural Sciences, University of Linköping, Linköping, Sweden

Keywords: morality, moral development, social behavior, socialization, neuroscience, novel perspectives

Editorial on the Research Topic

Novel Developmental Perspectives on the Link Between Morality and Social Outcomes

Morality is a complex construct examined in research from a number of disciplinary perspectives. Often thought of as the ability *to decide* about rightness or wrongness in situations involving a person's well-being and in terms of concerns about justice, rights, caring and virtues, morality also refers to the ability *to regulate behaviors* affecting others. Although the constructs overlap, moral cognition, moral standards and moral actions (social outcomes) are not equivalent; their relation and gaps can be altered by emotions and social influences. The complex association between moral standards and actions has been investigated in psychological and educational sciences, neurosciences, and philosophy, often separately and from different angles. From a developmental point of view, some researchers (e.g., Bandura, 1986) emphasized the possibility that moral cognition and action originate from social learning and transactions within social contexts, while others (e.g., Haidt, 2001) suggested that morality stems primarily from human biological organization.

In recent years, the scientific debate on the roots of morality and the relationship between moral cognition and behavior has produced different, sometimes contradictory, theorizations and studies with equivocal results. Difficulties in disentangling the origins of morality and capturing to what extent moral cognition and standards translate into social behavior also reflects the complexity of developing assessment measures able to accurately quantify or qualify moral processes in relation to social dimensions and outcomes. This Research Topic aims to contribute to this interdisciplinary scientific debate about morality with nine papers providing novel contributions in terms of theorization, empirical data and the development of new measures, and focusing on the developmental span.

With regards to the origins of morality, in their theoretical contribution, Carpendale et al. argue in favor of overcoming a Cartesian-split-mechanistic view of morality as originating from culture or biology. They propose a novel process-relational perspective about knowledge and morality as constructed through social interaction and as a process of coordinating perspectives. Four other papers deepen the debate about the role played by moral processes related to self-justifying one's own transgressions (*moral disengagement*; Bandura, 1986) in the social phenomenon of bullying. The literature has provided consistent evidence that the proneness to morally disengage is linked with higher levels of bullying behaviors and lower proneness to defend bullied peers (e.g., Thornberg et al., 2015). Nevertheless, there is still debate about the extent to which this type of moral cognition is correlated with being a passive bystander in bullying,

OPEN ACCESS

Edited and reviewed by:
Katharina J. Rohlfing, University of Paderborn, Germany

***Correspondence:**
Simona C. S. Caravita
simona.c.caravita@uis.no

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

Received: *02 March 2022*
Accepted: *11 March 2022*
Published: *30 March 2022*

Citation:
Caravita SCS, Beauchamp MH and Thornberg R (2022) Editorial: Novel Developmental Perspectives on the Link Between Morality and Social Outcomes. Front. Psychol. 13:888373. doi: 10.3389/fpsyg.2022.888373

about the interplay of moral disengagement mechanisms and other moral dimensions in the explanation of behaviors in bullying, and about the role played by moral disengagement in relation to bias-based bullying (e.g., ethnic bullying). The four papers provide novel contributions to this debate. By means of longitudinal data, Falla et al. shed further light on the complex associations that exist between social behaviors and morality, providing evidence that perpetrating negative behaviors (bullying) can increase moral disengagement and decrease empathy, and that some moral disengagement mechanisms mediate the link between behavior and empathy. Caravita et al. for their part, contribute to clarifying that moral disengagement and other forms of moral cognition (comprehension of rules) are separate mechanisms, and when they are both taken into account, moral disengagement is the only moral cognitive dimension associated with the perpetration of bullying. Further highlighting the complexity of moral mechanisms and their associations with social outcomes, Iannello et al. show that moral disengagement mechanisms mediate the association between ethnic prejudice and perpetrating ethnic bullying. They also provide novel results suggesting that closeness to teachers (an emotional contextual factor) can help restrain morally disengaged children from perpetrating bullying. The study by Lo Cricchio et al. provides a novel measure to assess moral disengagement in situations of ethnic bullying, thus presenting moral disengagement in a situational perspective. These studies offer important insights also for a moral component in anti-bullying intervention.

At the crossroads of neuroscience and developmental psychology research, Bacchini et al. contribute to the limited literature on deontological vs. utilitarian moral reasoning in adolescence (e.g., Caravita et al., 2017) with an innovative study showing how these forms of moral reasoning are related to individual factors (callous-unemotional traits and moral disengagement proneness) and contextual experiences (perceived parental rejection and exposure to community violence). A further novel contribution to the research on the development of moral cognition comes from the study by Zhou and Wong, who compare the understanding of restorative vs. retributive justice among children between 5 and 8 years of age. They find that a higher preference for a restorative justice approach emerges with age.

The relevance of positive life experiences and relationships in early childhood to build positive socio-moral temperament is investigated in a cross-cultural study presented in the article by Narvaez et al. Results highlight the relevance of guaranteeing children's emotional wellbeing, in terms of happiness and thriving and low depression and anxiety, to promote their self-regulation and positive moral socialization outcomes through socio-moral temperament. Lastly, a novel methodological contribution to research on moral reasoning comes from the study by Zarglayoun et al. who developed the MorALERT serious videogame to assess and strengthen moral reasoning skills and competencies. This tool offers new possibilities for assessment, remediation and intervention research in this area.

Together, the nine papers enrich knowledge on moral processes, their development and how they are linked to social behaviors, and open new important avenues and lines of research on this topic.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

REFERENCES

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Hoboken: Prentice-Hall.

Caravita, S. C. S., De Silva, L. N., Banfi, P. V., and Antonietti, A. (2017). Age-related differences in contribution of rule-based thinking toward moral evaluations. *Front. Psychol.*, 8:597. doi: 10.3389/fpsyg.2017.00597

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychol. Rev.* 108, 814–834. doi: 10.1037/0033-295X.108.4.814

Thornberg, R., Pozzoli, T., Gini, G., and Jungert, T. (2015). Unique and interactive effects of moral emotions and moral disengagement on bullying and defending among school children. *Elementary School J.* 116, 322–337. doi: 10.1086/683985

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 03 September 2021
doi: 10.3389/fpsyg.2021.713081

Prejudice and Ethnic Bullying Among Children: The Role of Moral Disengagement and Student-Teacher Relationship

Nicolò Maria Iannello[1], Marina Camodeca[1]*, Carmen Gelati[2] and Noemi Papotti[3]*

[1]Department of Languages and Literatures, Communication, Education, and Society, University of Udine, Udine, Italy, [2]Department of Psychology, University of Milano – Bicocca, Milan, Italy, [3]Department of Psychology, Catholic University of the Sacred Heart, Milan, Italy

The identification of factors associated with ethnic bullying within multiethnic schools is a timely social issue. Up to now, ethnic prejudice has been found to facilitate aggression triggered by schoolmates' cultural background. Yet, there is still a dearth of research about the mechanisms underlying this relation among children. In order to fill this gap, by adopting a social-cognitive developmental perspective on prejudice and morality, this paper investigated the mediating role of moral disengagement in the association between ethnic prejudice and ethnic bullying, as well as the moderating role of closeness with the teacher. A mediation model and a moderated mediation model were applied to data collected from 552 primary school children aged 8–10 years. Ethnic prejudice, ethnic bullying, and moral disengagement were assessed through self-reported questionnaires, whereas a questionnaire was administered to teachers to assess the level of closeness with their pupils. Results indicated that ethnic prejudice was directly and positively related to ethnic bullying and that moral disengagement partially mediated this association. This indirect link was particularly strong for children with low levels of closeness with their teachers, whereas it resulted not significant for pupils with high levels of closeness, suggesting that closeness with the teacher might restrain morally disengaged children from enacting ethnic bullying. Implications for research and practice aimed at reducing prejudice and moral disengagement, as well as at promoting positive relationships among children and between pupils and teachers, are discussed.

Keywords: ethnic prejudice, ethnic bullying, moral disengagement, closeness with the teacher, primary school

OPEN ACCESS

Edited by:
Robert Thornberg,
Linköping University, Sweden

Reviewed by:
Dagmar Strohmeier,
University of Applied Sciences Upper Austria, Austria
Francesca Giovanna Maria Gastaldi,
University of Turin, Italy

****Correspondence:***
Nicolò Maria Iannello
nicolo.iannello@uniud.it
Marina Camodeca
marina.camodeca@uniud.it

Specialty section:
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

Received: *21 May 2021*
Accepted: *05 August 2021*
Published: *03 September 2021*

Citation:
Iannello NM, Camodeca M, Gelati C and Papotti N (2021) Prejudice and Ethnic Bullying Among Children: The Role of Moral Disengagement and Student-Teacher Relationship. Front. Psychol. 12:713081. doi: 10.3389/fpsyg.2021.713081

INTRODUCTION

The term ethnic bullying identifies an aggressive action perpetrated toward individuals on ground of their ethnic origins (Elamé, 2013). Similar to traditional bullying, this type of bias-based harassment is carried out, intentionally and repetitively, against children who are not able to defend themselves and is enacted through verbal attacks (e.g., name calling), physical

means (e.g., hitting), and relational/social aggression (e.g., exclusion; McKenney et al., 2006; Scherr and Larson, 2010; Elamé, 2013).

Several studies highlighted the incidence of bullying related to ethnicity among school-age children in Europe (Strohmeier et al., 2011; Xu et al., 2020). For instance, in Italy, primary and middle school students belonging to minority groups are likely to be easy targets of aggression by majority group members (Caravita et al., 2016). Moreover, in Britain schools, Hindu, Indian Muslim, and Pakistani children have been found to be bullied by other Asian pupils because of differences in skin color, faith, spoken language, and/or traditions (e.g., food and clothing; Eslea and Mukhtar, 2000).

Negative consequences of ethnic bullying on individuals' development are documented and include poor adjustment and internalizing and externalizing problems (McKenney et al., 2006). Yet, little is known about what might push or limit the attacks toward peers on ground of their ethnic background (Bayram Özdemir et al., 2015). A better understanding of ethnic bullying seems to be urgent in a country like Italy where non-Italian students, mainly belonging to a second generation of immigrants, are increasing in number and have a migratory background traceable in more than 200 countries (Italian Ministry of Education, 2020). Such a situation results in a mosaic of traditions and cultures requiring innovative practices aimed at promoting not only social inclusion of students with different ethnic background, but also positive relationships at school. From this perspective, it is useful identifying mechanisms and risk and protective factors that might encourage and/or restrain ethnic bullying.

Ethnic Prejudice and Bullying

Ethnic prejudice refers to the tendency to overgeneralize and simplify (mostly in a negative sense) information on other cultural groups and to have irrational preconceptions about them (van Dijk, 1984). In particular, it relates to beliefs and thoughts about ethnically different groups or individuals (cognitive component), to the emotional reactions (e.g., discomfort) associated with these groups and individuals (affective component), and to the actions carried out toward these targets (behavioral component; Rosenberg and Hovland, 1960; Duckitt, 2003). Despite their young age, even children seem to hold negative views on outgroup members (Levy et al., 2004). Particularly, they seem to progressively shift from a condition in which they have a mere preference for their in-group to a phase in which they might adversely appraise outgroups (Nesdale, 2010).

Although it has been highlighted that hostile predispositions toward those who are culturally different are likely to drive bias-based bullying at school (Dessel, 2010), research so far mainly focused on adolescents (Bayram Özdemir et al., 2015; Caravita et al., 2020). Interestingly, some works have stressed that emotions, more than beliefs, are at stake when individuals relate to outgroup members and that the emotional component of prejudice might be conductive of ethnicity-based bullying among youth (Papotti and Caravita, 2020). In the present study, we borrowed from this body of works on adolescents and explored the role of children's negative attitudes toward outgroups in fostering ethnic bullying. Particularly, given the relevant part of emotions in conditioning intergroup relations (Tropp and Pettigrew, 2005), the affective facet of children's ethnic prejudice was investigated and hypothesized to have a direct association with ethnic bullying. In other words, it was anticipated that children who experience negative feelings toward culturally different groups are more likely to attack their members.

Moral Disengagement, Ethnic Prejudice, and Bullying

In their pathway toward ethnicity-based bullying, children may also turn to specific moral cognitive distortions that would help them legitimate their reprehensible behaviors (Caravita et al., 2019). In particular, according to Bandura's (1991) social-cognitive theory of moral thought and action, individuals tend to deactivate moral control over their conduct through moral disengagement, a means by which people avoid self-sanctions and negative emotions (e.g., guilt and shame) that would prevent them from engaging in harmful acts. Literature agrees that, despite individuals judge bullying as wrong, they continue carrying out different forms of harassment by condoning their behaviors through moral cognitive processes (Hymel and Bonanno, 2014; Killer et al., 2019; Lo Cricchio et al., 2020). In light of this, it is reasonable to suppose that moral disengagement over personal actions against culturally different peers may be directly linked to ethnic bullying (Caravita et al., 2019; Bayram Özdemir et al., 2020).

Although the relationship between ethnic prejudice and moral disengagement is an understudied topic, some findings showed positive correlations between prejudice and the mechanism of dehumanization (Costello and Hodson, 2012), that might indicate that children who have negative views about ethnically different peers may also be inclined to justify discrimination through moral cognitive distortions. In turn, the belief that members of the outgroup lack human attributes may be conductive of negative behaviors, as found among adults (Vaes et al., 2003; Demoulin et al., 2004). On the base of these findings, it was supposed that moral disengagement might mediate the relation between affective ethnic prejudice and ethnic bullying. Specifically, a way in which children might turn their negative feelings and attitudes into reprehensible behaviors is by condoning their actions through moral disengagement mechanisms, which may make their conduct appear legitimate and, thus, facilitate ethnic bullying. This line of reasoning may be justified by the integrative social-cognitive developmental perspective on prejudice (Rutland et al., 2010), which posited that children consider together group-based criteria (e.g., group identity, in-group favoritism, and stereotyping) and morality (e.g., believing that it is fair/unfair to exclude someone) when they are about to reject groups and individuals. Given that moral disengagement processes might be influenced by situational dimensions (e.g., targets' immigrant status), in the

current study, a measure of moral disengagement that assesses the proneness to justify transgressive behaviors toward peers with a different cultural background was used (Caravita et al., 2019).

The Role of the Relationship With Teachers

Consistently with Bandura's (1986) social-cognitive theory, moral behaviors are the product of the interactions between individual and environmental factors. In light of this, we may suppose that the effect of ethnic prejudice and moral disengagement on ethnic bullying at school might be damped by specific contextual factors, such as the student-teacher relationship. Indeed, teachers spend much time with their pupils, represent relevant adults for children, and are likely to affect their development and behaviors. In particular, they often play a protective role in face of different risks (Sabol and Pianta, 2012). Generally, the student-teacher bond is evaluated on the base of its quality, that is, the extent to which the dyad is characterized by close or conflictual interactions (Fraire et al., 2008; Sabol and Pianta, 2012).

Following a social learning framework (Bandura, 1971, 1986), children learn and behave through modeling and imitation of others, as well as through vicarious experiences, particularly when the role models are socializing agents (e.g., teachers and parents) or are taken as positive models. Therefore, children and adolescents "with positive social relationships with parents, peers, and teachers benefit from these experiences and, therefore, are more likely to display better social, emotional and behavioral outcomes" (Wachs et al., 2020, p. 2). In addition, teachers who hold a positive relationship with their pupils may be more prone to promote an efficacious communication with them, encourage them, reinforce positive behaviors, and provide helpful feedbacks, which may foster pupils' self-efficacy and willingness to behave properly (Wachs et al., 2020).

Relationships with teachers might be viewed from an attachment perspective as well (Davis, 2003; Bouchard and Smith, 2017). Referring to Bowlby's (1969) theory, a secure relationship with a caregiver, being parent or teacher, is predictive of psychosocial and emotional adaptation. Teachers may contribute to children's working models of peer relationships, by fostering useful skills for self-regulation and child-to-child interactions, and by hindering aggressive interactions (Bouchard and Smith, 2017). A strong bond with significant others, such as teachers, may also increase students' sense of being valued and trusted and facilitate social sharing of experiences and feelings, which may inhibit bullying behaviors (Cho and Lee, 2018; van Aalst et al., 2021). As a matter of fact, children securely attached with teachers are less likely to be involved in bullying (Cho and Lee, 2018). This protective role of reorganizing relational schemas is particularly relevant for those children with insecure previous experiences of attachment or problem behaviors, such as aggression, and compensates for negative relationships with peers (Sabol and Pianta, 2012; van Aalst et al., 2021).

Whatever mechanisms are involved, several studies pointed out that a close relationship with teachers is associated with fewer bullying episodes and its negative outcomes, whereas a conflictual relationship with teachers seems to increase bullying involvement (Richard et al., 2011; Wang et al., 2015; Longobardi et al., 2018; Camodeca and Coppola, 2019). Teachers seem to play an important role also in orienting their pupils' morality and attitudes toward ethnic outgroups. In general, it has been shown that perceiving a positive school climate, that includes support from teachers, might weaken the impact of moral disengagement on students' bullying perpetration (Teng et al., 2020). In particular, it could be surmised that teachers sharing close connections with their students might help them consider the consequences of immoral conducts, act properly by monitoring their cognitions and attitudes, and, consequently, prevent them from bullying others with different origins. In addition, recent studies pointed out that close relationships with the teachers, as secondary attachment figures, might provide children with a sense of relational security that would help them be more open to and have positive attitudes toward ethnic outgroups (Geerlings et al., 2017). In sum, on ground of these theoretical considerations and empirical evidences, it could be hypothesized that the effect of both ethnic prejudice and moral disengagement on ethnic bullying might vary as a function of the extent to which teachers have a warm relationship with their pupils. It could also be possible that a positive relationship between pupils and teachers might inhibit the path from ethnic prejudice to ethnic bullying by deactivating moral disengagement mechanisms.

The Present Study

To the best of our knowledge, although both ethnic prejudice and moral disengagement have been found to be related to ethnic bullying (Bayram Özdemir et al., 2015; Caravita et al., 2019), and warm student-teacher interactions positively influence pupils' cognitions and behaviors (Davis, 2003; Sabol and Pianta, 2012), all these variables have never been investigated within a comprehensive conceptual model. In order to fill this gap, the present work sought to provide a better understanding of how affective ethnic prejudice[1], moral disengagement, and closeness with the teacher might jointly impact ethnicity-based bullying among primary Italian and immigrant school children. The links between these variables were tested within a moderated mediation model (**Figure 1**). Particularly, on the basis of the aforementioned discussion and literature, it was hypothesized that (H1) ethnic prejudice would be positively and directly associated with both ethnic bullying and moral disengagement; (H2) moral disengagement would be positively and directly associated with ethnic bullying; (H3) moral disengagement would mediate the relation between ethnic prejudice and ethnic bullying; and (H4) the quality of student-teacher relationship, considered as a protective factor, would moderate the direct and indirect associations between ethnic prejudice, moral disengagement, and ethnic

[1]For parsimony, throughout the paper, we often used only the term "ethnic prejudice" to refer to the affective component of ethnic prejudice, which was actually what we tested.

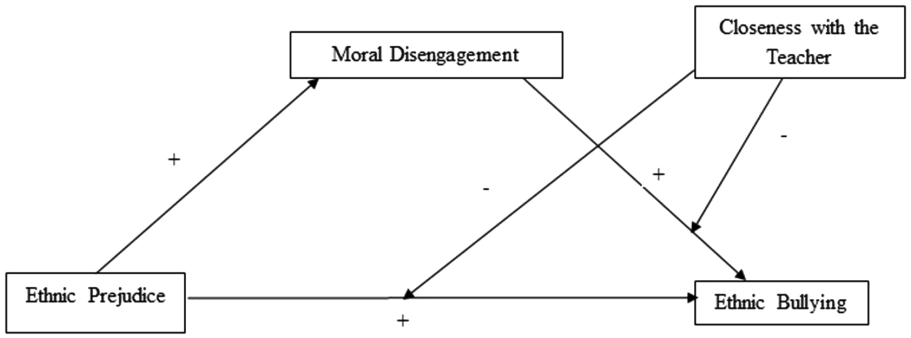


FIGURE 1 | Theoretical model in which ethnic prejudice is expected to affect ethnic bullying through moral disengagement and in which closeness with the teacher is expected to moderate direct and indirect associations.

bullying, which would be weaker for children with a positive relationship with their teachers.

Finally, gender, grade, and immigrant status were controlled for due to their role in affecting study variables, as indicated in the literature. Indeed, boys, compared to girls, have been found more prejudiced (Costello and Hodson, 2012), more prone to bully others on ground of their ethnicity (Bayram Özdemir and Özdemir, 2020), more morally disengaged (Thornberg and Jungert, 2014), and less close with their teachers (Murray and Murray, 2004). Younger children have been reported to be more prejudiced (Raabe and Beelmann, 2011) and to have warmer interactions with the teacher (Drugli, 2013) than older pupils. Students from minority groups, in comparison with the majority ones, have been found to show more problematic relationships with their teachers (Jerome et al., 2009) and higher levels of moral disengagement (Caravita et al., 2019), whereas findings about differences in ethnic bullying involvement between minority and majority groups are still scarce (Tolsma et al., 2013).

MATERIALS AND METHODS

Participants

The sample included 552 children aged 8–10 years ($M = 9.08$ years, $SD = 0.59$; 289 boys and 263 girls). Pupils attended the third (44.9%) and the fourth grade (55.1%) of 13 public primary schools in two different areas in Northern Italy. The sample was ethnically heterogeneous with children mostly Italians[2] (74.6%). Children from other countries were mainly born in Italy (82%), whereas their cultural background posed them as original from Eastern Europe (31%), North Africa (22%), Far East (14%), South America (11%), other African countries (10%), other European countries (9%), and other (3%). The sample included students from a wide range of sociocultural backgrounds (from low and working class to upper class), with a university degree obtained by 32% of fathers and 41% of mothers, whereas 20% of fathers and 14% of mothers did not obtained a high school degree.

[2]According to the Italian law, a person is considered a foreigner if both his/her parents have a non-Italian citizenship, even if he/she is born in Italy. Consequently, a child born from an Italian parent and a foreigner parent is considered Italian, similarly to a child born from two Italian parents.

Procedure

The present study is part of a large project aimed at investigating ethnic bullying and including many measures. The main objectives and the methodology of the study were introduced to school principals and teachers, who agreed to participate. Parents were sent a letter to explain the study and were asked to give their informed consent, which was granted for 84.41% of the original sample contacted. The instruments were administered to students in the class group, during school time, in two different days. Children were first explained what we meant with "ethnic" or "origin," saying that we referred to "those people (or their families) who talk a different language, or have the culture, the skin color, or the religion different from your own, or who come from different countries. For instance, we can think about ethnic groups such as Italians, Chinese, Albanians, Moroccans." Teachers filled their questionnaires within 1 week. Participants were assured about the confidentiality of all the information provided and that they could withdraw at any time. The Ethical Committee of the University of Udine approved this study, and all procedures were performed in accordance with the ethical principles for psychological research of the Italian Association of Psychology.

Measures

Sociodemographic Variables

Participants' sociodemographic variables were provided by their parents who were asked to indicate their children's gender, age, grade, and place of birth. Parents were also asked to indicate their own countries of origin, education level, and job.

Ethnic Prejudice

The affective component of ethnic prejudice was assessed through two items aimed at measuring the extent (from 0 = very happy to 4 = very annoyed) to which respondents felt happy

or annoyed to sit next to a classmate "from a different cultural background (e.g., with a different skin color or language) than your own" and "with a different religion from your own" (adapted; Buccoliero and Maggi, 2017).

Moral Disengagement

Children were proposed with the following hypothetical scenario, specifically designed to assess children' proneness to legitimate negative behaviors toward a newcomer immigrant student (Caravita et al., 2019); the gender of the protagonist matched participants' gender ("Hamir" in the male scenario and "Elissar" in the female scenario): *"Hamir/Elissar, a child from another country, is your new classmate; for some weeks you have both been back at school after the summer vacation. You started to call him/her "stupid" because he/she does not talk much; you also started to wait for him/her in the corridor before lessons and hit him/her to let him/her fall down. Sometimes you hide his/her school bag so that he/she could not find it and damage his/her books or copybooks. At the break, you do not talk with Hamir/Elissar and you do not want to play with him/her; you do not even want to invite him/her at your birthday party that you are organizing with your classmates. Hamir/Elissar in all these situations cannot defend him/herself."* In order to assess self-justifications processes for harassing culturally different peers, children were asked to answer eight questions on a 5-point Likert-type scale, ranging from totally false (1) to totally true (5; e.g., "If you misbehave towards Elissar/Hamir it is because he/she misbehaved towards you first"; "Damaging books is not really harmful"). The wording of scenario and items was simplified to make it more suitable for young children. Items were averaged to rate moral disengagement as an overall disposition to condone one's despicable behaviors; high scores indicate high moral disengagement.

Ethnic Bullying

An adaptation of the Florence Bullying and Victimization Scale (FBVS; Palladino et al., 2020) was used to evaluate ethnic bullying. A definition of bullying was presented prior to administering the scale *"Bullying happens when some children offend, ignore, kick, push, threaten, exclude other peers on purpose, or say bad things behind their back. It is also bullying when a child is teased repeatedly and in a nasty way. These episodes happen frequently, and it is difficult for the children who suffer from bullying to defend themselves. It is not bullying if two students of about the same strength quarrel or fight."* Children were asked to think how often they have been involved in bullying behaviors in the last 2 or 3 months. Despite the instrument assesses traditional bullying and victimization as well, for the purpose of the current study, four items concerning ethnic bullying perpetration were used, covering different forms of bullying (e.g., "I have hit/excluded/teased/spread rumors about/someone because of his/her origin, for instance for the color of the skin, the language, the religion"). Responses were on a 5-point Likert-type scale, from never (1) to several times a week (5).

The Quality of Student-Teacher Relationship

The Student-Teacher Relationship Scale (Pianta, 1994; Italian version by Fraire et al., 2008) was adopted to investigate how teachers perceived their relationship with each pupil. The instrument consists of three subscales: Closeness, Conflict, and Dependence. For the purpose of this study, only the Closeness subscale was used, which comprises eight items (e.g., "I share an affectionate, warm relationship with this child"). Teachers responded on a 5-point Likert-type scale, ranging from not applicable (1) to totally applicable (5).

The choice of considering only the Closeness scale mirrors our aim of testing a possible protective factor, discarding, therefore, the Conflict scale, which is usually regarded as a risk factor. In addition, the Dependence scale resulted unreliable in previous studies and, although it may be considered a protective factor at a young age, it includes aspects which can be risky at the end of primary school (Camodeca and Coppola, 2019).

Data Analysis

Only data from students who were present on both administration days, or who had the opportunity to fill all questionnaires, were considered. Among these, missing data were very few in each item of study variables (range: 0.4–2.3%) and were handled using expectation maximization (EM) algorithm (Graham, 2009). As to categorical variables, only immigrant status had missing values ($n = 18$; 3.2%) and participants with missing values did not significantly differ from those with complete data on any study variable, except for moral disengagement ($t = 2.34$; $p < 0.05$; $M_{\text{Missing}} = 2.37$ and $M_{\text{Complete}} = 2.00$). Following literature, we employed complete case analysis, removing these 18 participants, which is considered reasonable when missing data are less than 5% (Graham, 2009; Jakobsen et al., 2017).

Reliabilities were calculated as the greater lower bound (glb) index, indicating the lowest value of the real reliability (which ranges from glb to 1; Sijtsma, 2009). In order to examine the relations among study variables, Pearson correlations were employed. *T*-tests were conducted to compare boys and girls, younger and older children, and immigrant and non-immigrant children on the study variables.

The possibility of conducting a multilevel analysis was taken into account, given the nested nature of data. The intra-class correlation index (ICC) was calculated, which was below 0.05. Following the literature that identifies ICC indexes higher than 0.05 as suitable for performing a multilevel analysis, we decided to implement other types of analysis (Koo and Li, 2016).

Mediation and moderated mediation were tested by using PROCESS macro (Hayes, 2013), which calculates a series of regressions and includes all predictors in one block. Particularly, Model 4 was selected to test whether moral disengagement mediated the link between ethnic prejudice and ethnic bullying. Model 15 was employed to test whether closeness with the teacher moderated the association between ethnic prejudice and ethnic bullying *via* the mediator and

TABLE 1 | Descriptive statistics and bivariate correlations among study variables.

	1	2	3	4
1. Ethnic Prejudice	–			
2. Moral Disengagement	0.23***	–		
3. Ethnic Bullying	0.15***	0.15***	–	
4. Closeness with teacher	−0.14**	−0.04	−0.10*	
Means (SD)	0.83 (0.93)	2.00 (0.65)	1.08 (0.28)	4.00 (0.77)
Reliabilities	0.73	0.65	0.83	0.93

*Reliabilities indicate the greater lower bound (glb) index *p<0.05; **p<0.01; ***p<0.001.*

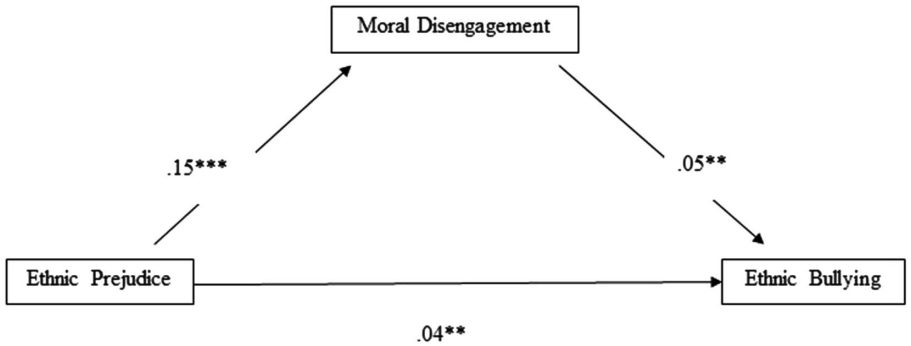


FIGURE 2 | Unstandardized regression coefficients of the mediation model. The covariates (gender, grade, and immigrant status) were not included in the figure. The indirect effect of ethnic prejudice on ethnic bullying *via* moral disengagement was also significant (*B* = 0.01; *SE* = 0.01, 95% *CI* = 0.0003, 0.0197). **p<0.01; ***p<0.001.

TABLE 2 | Unstandardized regression coefficients of moral disengagement and ethnic prejudice on ethnic bullying, and the moderating effect of closeness with the teacher.

Predictors (IV)	Model 1 (DV: MD)	Model 2 (DV: EB)
Ethnic Prejudice	0.15***	0.04**
Moral Disengagement		0.05**
Closeness		−0.02
Ethnic Prejudice X Closeness		0.03
Moral Disengagement X Closeness		−0.08**
Gender	−0.02	−0.00
Grade	0.04	0.00
Immigrant Status	−0.00	0.04**
R2	0.05***	0.08***
F	7.55	5.55

*IV, independent variable; DV, dependent variable; MD, moral disengagement; and EB, ethnic bullying. Gender is coded as: boys = −1 and girls = +1. Immigrant status is coded as: Italian = − 1 and immigrant = 1 **p<0.01; ***p<0.001.*

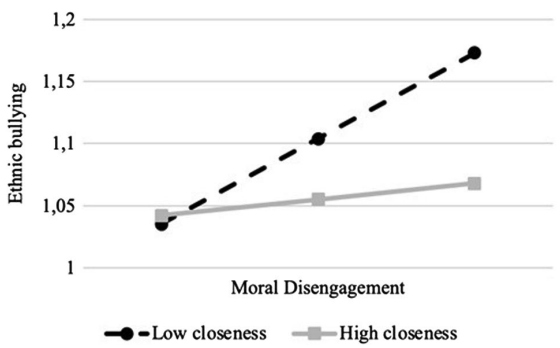


FIGURE 3 | Interaction between moral disengagement and closeness with the teacher on ethnic bullying.

whether the interactions terms between closeness with teachers and ethnic prejudice/moral disengagement affected ethnic bullying. Bootstrapping with 5,000 resamples to compute 95% confidence intervals was used to test the significance of the regression coefficients. If the confidence intervals did not contain zero, then statistics were significant. Gender, grade, and immigrant status were entered as covariates in all analyses.

RESULTS

Preliminary Analyses

Table 1 summarizes means, standard deviations, reliabilities, and the correlation matrix. Positive and significant correlations emerged between ethnic prejudice, moral disengagement, and ethnic bullying perpetration. Closeness with the teacher was negatively correlated with ethnic prejudice and bullying.

T-tests indicated that girls had warmer relationships with their teachers ($t = -3.55$; $p < 0.001$) and less prejudice ($t = 3.28$;

$p < 0.01$) than boys. Also, third graders had closer relationships with their teachers than fourth graders ($t = 3.72$; $p < 0.001$); moreover, students from a migratory background, compared to their Italian peers, were more prone to bully others on ground of their ethnicity ($t = -2.46$; $p < 0.05$) and were less close to their teachers ($t = 2.01$; $p < 0.05$).

Mediation Effect Analysis

Model 4 in PROCESS macro (Hayes, 2013) was used to test the mediating effect of moral disengagement on the link between ethnic prejudice and ethnic bullying perpetration, while controlling for gender, grade, and immigrant status. The results are shown in **Figure 2**. Ethnic prejudice was significantly and positively related to moral disengagement ($R^2 = 0.05$; $p < 0.001$) and both ethnic prejudice and moral disengagement had a significant positive association with ethnic bullying ($R^2 = 0.05$; $p < 0.001$). In addition, the indirect effect of ethnic prejudice on ethnic bullying through moral disengagement was also significant ($B = 0.01$; $SE = 0.01$; 95% $CI = 0.0003, 0.0197$). Gender and grade were not significantly related to any of the model variables, whereas immigrant status was positively associated with ethnic bullying ($B = 0.04$; $p < 0.01$), indicating that children with a migratory background were more involved in ethnic bullying than their Italian peers.

Moderated Mediation Effect Analysis

Closeness with the teacher was expected to moderate the associations between ethnic prejudice and ethnic bullying and between moral disengagement and ethnic bullying, as well as the indirect link. To test this hypothesis, while controlling for gender, grade, and immigrant status, closeness with the teacher was included and Hayes' s PROCESS macro (model 15) was used. As **Table 2** illustrates, with regard to covariates, as in the previous analysis, only immigrant status was significantly and positively associated with ethnic bullying, indicating a higher involvement for immigrant than Italian children. Closeness with the teacher was not associated with ethnic bullying nor was its interaction with ethnic prejudice. In contrast, the interaction term between moral disengagement and closeness with the teacher had a significant effect on ethnic bullying. A simple slope analysis (**Figure 3**) evidenced that moral disengagement was positively associated with ethnic bullying perpetration for children with a low level of closeness with the teacher ($B = 0.11$; $p < 0.001$), whereas this association was not significant for children with a high level of closeness with the teacher ($B = -0.01$; $p > 0.05$).

Results further indicated that the indirect effect of ethnic prejudice on ethnic bullying perpetration *via* moral disengagement was moderated by the quality of student-teacher relationship ($B = -0.01$; $SE = 0.01$; 95% $CI = -0.0274, -0.0008$). Specifically, for children with a low level of closeness with the teacher, ethnic prejudice had a positive effect on ethnic bullying *via* moral disengagement ($B = 0.02$; $SE = 0.01$; 95% $CI = 0.0017, 0.0376$), whereas this indirect effect was non-significant in case of high closeness ($B = 0.00$; $SE = 0.00$; 95% CI = −0.0094, 0.0064).[3]

DISCUSSION

In the present work, a moderated mediation model was proposed examining whether moral disengagement mediated the linkage between negative feelings and attitudes toward pupils from other countries and ethnic bullying perpetration, and whether closeness with the teacher impacted the direct and indirect associations between ethnic prejudice, moral disengagement, and ethnic bullying.

Results partially confirmed our hypotheses. Ethnic prejudice, in terms of its affective component, was positively related to ethnic bullying, and moral disengagement mediated this link. In addition, findings highlighted that closeness with the teacher moderated the association between moral disengagement and ethnic bullying, whereas it did not moderate the relation between ethnic prejudice and ethnic bullying. Finally, this study pointed out that the indirect effect of ethnic prejudice on ethnic bullying was stronger for pupils with low levels of closeness with their teacher. In the following paragraphs, these outcomes are discussed thoroughly.

Direct and Indirect Effects on Ethnic Bullying

The hypothesis that ethnic prejudice was directly and positively associated with both ethnic bullying perpetration and moral disengagement was supported (H1). Consistently with previous studies that considered adolescents samples (Papotti and Caravita, 2020), the current work highlighted that the affective component of ethnic prejudice, that is experiencing aversive feelings (e.g., annoyance) toward pupils from another country, might trigger engagement in bullying also among children. In compliance with the social categorization framework (Tajfel and Wilkes, 1963; Tajfel et al., 1971), it could be speculated that children might tend to classify their schoolmates on the base of their salient cultural (e.g., spoken language) or physical (e.g., skin color) traits. This process might easily lead to "Us" vs. "Them" construal, as well as to in-group favoritism and outgroup derogation (Tajfel et al., 1971). As research showed (Kawakami et al., 2017), the use of social categories might affect intergroup relations, resulting in negative responses to the outgroups. It could be the case of this form of selective bullying that

[3]Adopting an exploratory approach, we also observed whether the moderated mediation associations held for Italian and immigrant students in the same way or differently. Therefore, we added the immigrant status as another moderator of the mediation and employed the Model 17 in Process, which allows two moderators in the relations between independent variable and outcome and between mediator and outcome. Results indicated that the immigrant status neither moderated the mediation ($B = 0.01$; $SE = 0.01$; 95% $CI = -0.0137, 0.0325$) nor resulted in significant interactions with moral disengagement and prejudice ($B = 0.02$, $p > 0.05$, and $B = 0.02$, $p > 0.05$, respectively), highlighting that the processes involved worked in the same way for all children, being them Italians or immigrants.

targets individual on ground of their ethnical and cultural background.

As expected, more prejudiced children might also be more morally disengaged, meaning that individuals might not feel any sense of guilt or blame when they hold negative attitudes toward ethnic outgroups, who are seen as a threat to one's in-group. Thus, children who have hostile feelings toward outgroups might also suspend their moral principles and values that would prevent them from derogating members of the outgroup. In turn, moral disengagement was connected to bullying based on ethnicity (H2). This result is in line with works showing moral disengagement as conductive of traditional bullying behaviors (Hymel and Bonanno, 2014; Killer et al., 2019) and with those highlighting the pervasive role of moral cognitive processes in prompting ethnic bullying among children and adolescents (Caravita et al., 2019; Bayram Özdemir et al., 2020). Additionally, these findings confirmed that moral disengagement might be influenced by situational dimensions, such as the ethnicity of the target. Indeed, one of the main novelties of this study is that it assessed children' tendency to justify aggression toward specific peers characterized by different nationalities and cultural backgrounds (Caravita et al., 2019).

Evidence for H3 about the mediation of moral disengagement between ethnic prejudice and ethnic bullying was also found. Children might be able to put their aversive feelings into action also by adopting moral strategies that would help them justify their reprehensible conduct, such as affirming that harassing someone on ground of his/her ethnic origins is not severe, but legitimate and deserved. According to the integrative social-cognitive developmental perspective on prejudice (Rutland et al., 2010), in their pathway toward bullying others because of their cultural origins, children might ground their choice both on group-based criteria, such as preserving own in-group, and on moral reasoning, such as evaluating admissible to act unfavorably toward members of the outgroups.

The Moderating Effect of Closeness With the Teacher

In line with our expectations, closeness with the teacher moderated the relation between moral disengagement and ethnic bullying, as well as the indirect effect of ethnic prejudice on ethnic bullying *via* moral disengagement (H4). Although morally disengaged children tend to bully their peers independently of the relationship with teachers, we can surmise that a not close relationship may facilitate such association, whereas a harmonious one is likely to make moral disengagement ineffective in pushing pupils to bully outgroup members. It seems that moral disengagement does not necessarily lead to bullying because, when protective factors intervene, the antisocial final behavior may be avoided. In addition, a little close bond with the teacher might foster prejudiced children to turn their aversive feeling into action by facilitating their mechanisms of moral disengagement toward culturally different peers. Altogether, it is likely that, regardless of their prejudice and moral cognition, children take their teachers as an example to follow or imitate, or as a source of information to guide behavior (Bouchard and Smith, 2017), in line with the theories previously advanced and in the following discussed.

Bandura's (1991) social-cognitive theory of morality proposes that moral agency is learned also through the environment in which individuals are embedded. As suggested by Rose et al. (2016) with preschool children, teachers are often unaware of bullying behavior, and, if they respond, they use verbal reprimands. It is possible that teachers are more ready to react to bullying behaviors enacted by those children they feel closer and to employ effective strategies with them, such as encouraging empathy for the victims or condemning the aggressive behavior, and less likely to blame the bully, which has been found ineffective in predicting bullies' intention to stop their behavior (Garandeau et al., 2016). Teachers may also be more prone to help these children reframe the meaning of their immoral behaviors (e.g., by showing them the consequences of bad actions), or it could be that they better know how to deal with them.

Following an attachment point of view, a little close relationship with the teacher may reflect a sort of independence or avoidance on behalf of specific children, who may feel freer to behave in an undesired and antisocial way. It is possible that teachers holding a positive relationship with their students communicate them that they care and have expectations about their behavior, which may contribute to refrain from bullying (Yoon and Bauman, 2014). When bullies have a close relationship with their teachers, they may wish to comply with the rules and norms set by them, in order to appear nice at their eyes, not disappoint them, and maintain or improve such a lovely bond. Their attempt to avoid any damage to this relationship may also be associated with their worry of undermining their self-esteem, which is related to positive student-teacher relationships (van Aalst et al., 2021).

Contrary to our hypotheses, closeness with the teacher did not moderate the relation between ethnic prejudice and ethnic bullying perpetration. It is likely that children might have interiorized aversive prejudices, which may have deep roots within them and be hardly hindered by teachers (Nesdale, 2010; Geerlings et al., 2017). Also, it could be posited that rather than teachers' relational and emotional support, it is teachers' explicit and implicit views on cultural diversity to weaken the impact of children's ethnic prejudice on their proneness to bully outgroup members.

Finally, although the impact of grade, gender, and immigrant status on bullying was not among our aims, we just mention that neither grade nor gender were associated with bullying in the regressions. As for grade, it can be due to the similar age of the participants, whereas the fact of not distinguishing between direct and indirect forms of bullying may account for similarities between girls and boys. Children with a migratory background were more engaged in bullying episodes, compared to their majority peers, which is in line with some findings among adolescents (Fandrem et al., 2009; Larochette et al., 2010). The need for peer acceptance and affiliation, disadvantages associated with immigrant children's environment, and social stigma that non-natives face daily in host countries might be regarded as underlying motives

for immigrant pupils to initiate aggressive behaviors (Xu et al., 2020). Alternatively, it could be just a matter of probabilities, given that, in Italian classrooms, the possible victims with different origins from one's own are more numerous for immigrant children than for Italian children.

Limitations, Strength Points, and Implications for Practice

This study should be interpreted in light of several shortcomings. Its cross-sectional design hindered the possibility to establish the direction of the associations among the variables. Thus, future research may adopt a longitudinal approach which would be helpful to ascertain developmental or causal pathways. Then, student-teacher relationship was detected by using only teachers' perspective; including children's point of view may provide a clearer picture of the interactions between pupils and their teachers. Future works are encouraged to also take into account other relevant variables associated with ethnic bullying, such as those referring to intergroup relations (e.g., group identity, intergroup contact, and ethnocultural empathy) or to other aspects of prejudice beyond the affective one (e.g., cognitive and behavioral). Similarly, they could take advantage from employing less direct assessments, such as peer reports and observations.

Some strength points can be highlighted as well, such as the employment of a comprehensive model showing the interplay of individual (ethnic prejudice and moral disengagement) and contextual factors (student-teacher interactions) that might set the stage for ethnic bullying. In addition, a large sample size was recruited, and self-reports and teacher-reports were used, which reduce shared variance. Finally, we adopted a measure of moral disengagement that specifically addressed a target with a migratory background.

These outcomes, although correlational, might indicate some practical suggestions for teachers, educators, and practitioners and would be helpful in the implementation of anti-bullying programs in multicultural primary schools. As a possible underlying mechanism of ethnic bullying is experiencing negative feelings toward culturally different children, schools are encouraged to facilitate positive intergroup contact and to involve pupils in activities that improve their perception of and their empathy toward outgroups (Sklad and Park, 2017). Also, interventions are recommended to reduce moral disengagement, helping children reconstruct their beliefs about violence, be aware of negative consequences of their acts, and enhance their sense of personal responsibility for their conduct (Hymel and Bonanno, 2014).

Programs targeting teachers are relevant to underline their central role in children's wellbeing and to empower them. Albeit necessary, strengthening teachers' ability to intervene in bullying situations or changing their attitudes toward bullying may not be enough; it seems paramount that teachers develop socio-emotional skills and become "mindful of their relationships with students" (Bouchard and Smith, 2017, p. 117). Trainings for teachers, specifically aimed at reinforcing this type of competence and at promoting awareness, are therefore recommended, in particular if such education is missing or not systematic. Teachers who are equipped with good socio-emotional competencies can recognize their own and their pupils' emotions, are sensitive to their students' needs and desires, and can respond properly, especially to those children more easily involved in bullying behaviors. Indeed, a close teacher-child relationship may foster the development of a secure and supportive context in which violence and aggression are discouraged, and students' cooperative and relational skills promoted (Jennings and Greenberg, 2009; Bouchard and Smith, 2017).

DATA AVAILABILITY STATEMENT

The dataset presented in this article is not readily available to guarantee participants' privacy. Participants were ensured that data would have been disseminated only in an aggregated form, that is, at a group level. Requests to access the dataset should be directed to the authors.

ETHICS STATEMENT

The study involving human participants was reviewed and approved by the Institutional Review Board, Department of Languages and Literatures, Communication, Education, and Society, University of Udine, Italy (protocol N. CGPER-2019-12-09-05). Written informed consent to participate in this study was provided by teachers and by children's parents or legal guardians.

AUTHOR CONTRIBUTIONS

MMI was involved in designing the work, collecting and analyzing the data, interpreting the results, and writing. MC supervised the project, organized the data collection, and contributed in analyzing data and writing and revising the paper. CG supervised data collection and was involved in the design and interpretation of this work as well as in revising it. NP was involved in revising the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The present study was financed by the Italian Ministry of University, within the PRIN program (Relevant National Interest Research Projects), grant number 20173E3Z7W. Additionally, the open access publication was partly supported by a fund for basic research (FFABR 2017).

ACKNOWLEDGMENTS

The authors are thankful to all principals, teachers, children, and families for participating in the project.

REFERENCES

Bandura, A. (1971). *Social Learning Theory*. Morristown, NJ: General Learning Press.

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1991). "Social cognitive theory of moral thought and action," in *Handbook of Moral Behavior and Development: Theory. Vol. 1*. eds. W. M. Kurtines and J. L. Gewirtz (Hillsdale, NJ: Erlbaum), 45–103.

Bayram Özdemir, S., Giles, C., and Özdemir, M. (2020). Differences and similarities between perpetrators of ethnic and non-ethnicity-based victimization. *J. Youth Adolesc.* 49, 1805–1820. doi: 10.1007/s10964-020-01271-5

Bayram Özdemir, S. B., and Özdemir, M. (2020). The role of perceived inter-ethnic classroom climate in adolescents' engagement in ethnic victimization: For whom does it work? *J. Youth Adolesc.* 49, 1328–1340. doi: 10.1007/s10964-020-01228-8

Bayram Özdemir, S., Özdemir, M., and Stattin, H. (2015). What makes youth harass their immigrant peers? Understanding the risk factors. *J. Early Adolesc.* 36, 601–624. doi: 10.1177/0272431615574887

Bouchard, K. L., and Smith, J. D. (2017). Teacher–student relationship quality and children's bullying experiences with peers: reflecting on the mesosystem. *Educ. Forum* 81, 108–125. doi: 10.1080/00131725.2016.1243182

Bowlby, J. (1969). *Attachment and loss: Vol. 1, Attachment*. New York, NY: Basic Books.

Buccoliero, E., and Maggi, M. (2017). *Contrastare Il Bullismo, Il Cyberbullismo Dei Pericoli Della Rete: Manuale Operativo per Operatori e Docenti, Dalla Scuola Primaria Alla Secondaria di 2° Grado [Counteracting Bullying and Cyberbullying: Handbook for Educators and Teachers, from Primary to High School]*. Milano, Italia: Franco Angeli.

Camodeca, M., and Coppola, G. (2019). Participant roles in preschool bullying: The impact of emotion regulation, social preference, and quality of the teacher–child relationship. *Soc. Dev.* 28, 3–21. doi: 10.1111/sode.12320

Caravita, S. C., Donghi, E., Banfi, A., and Meneghini, F. (2016). Essere immigrati come fattore di rischio per la vittimizzazione nel bullismo: uno studio italiano su caratteristiche individuali e processi di gruppo [Being immigrants as a risk factor for bullying victimization: An Italian study on individual characteristics and group processes]. *Maltratt. Abus. Infanz.* 18, 59–87. doi: 10.3280/mal2016-001004

Caravita, S. C. S., Stefanelli, S., Mazzone, A., Cadei, L., Thornberg, R., and Ambrosini, B. (2020). When the bullied peer is native versus immigrant: A mixed-method study with a sample of native and immigrant adolescents. *Scand. J. Psychol.* 61, 97–107. doi: 10.1111/sjop.12565

Caravita, S. C. S., Strohmeier, D., Salmivalli, C., and Di Blasio, P. (2019). Bullying immigrant versus non-immigrant peers: moral disengagement and participant roles. *J. Sch. Psychol.* 75, 119–133. doi: 10.1016/j.jsp.2019.07.005

Cho, S., and Lee, J. M. (2018). Explaining physical, verbal, and social bullying among bullies, victims of bullying, and bully-victims: assessing the integrated approach between social control and lifestyles-routine activities theories. *Child Youth Serv. Rev.* 91, 372–382. doi: 10.1016/j.childyouth.2018.06.018

Costello, K., and Hodson, G. (2012). Explaining dehumanization among children: The interspecies model of prejudice. *Br. J. Soc. Psychol.* 53, 175–197. doi: 10.1111/bjso.12016

Davis, H. A. (2003). Conceptualizing the role and influence of student-teacher relationships on children's social and cognitive development. *Educ. Psychol.* 38, 207–234. doi: 10.1207/S15326985EP3804_2

Demoulin, S., Torres, R. R., Perez, A. R., Vaes, J., Paladino, M. P., Gaunt, R., et al. (2004). Emotional prejudice can lead to infra-humanisation. *Eur. Rev. Soc. Psychol.* 15, 259–296. doi: 10.1080/10463280440000044

Dessel, A. (2010). Prejudice in schools: promotion of an inclusive culture and climate. *Educ. Urban Soc.* 42, 407–429. doi: 10.1177/0013124510361852

Drugli, M. B. (2013). How are closeness and conflict in student-teacher relationships associated with demographic factors, school functioning and mental health in Norwegian schoolchildren aged 6-13? *Scand. J. Educ. Res.* 57, 217–225. doi: 10.1080/00313831.2012.656276

Duckitt, J. (2003). "Prejudice and intergroup hostility," in *Oxford Handbook of Political Psychology*. eds. D. O. Sears, L. Huddy and R. Jervis (New York: Oxford University Press), 559–600.

Elamé, E. (2013). *Discriminatory Bullying: A New Intercultural Challenge*. Milano: Springer-Verlag Italia.

Eslea, M., and Mukhtar, K. (2000). Bullying and racism among Asian schoolchildren in Britain. *Educ. Res.* 42, 207–217. doi: 10.1080/001318800363845

Fandrem, H., Strohmeier, D., and Roland, E. (2009). Bullying and victimization among native and immigrant adolescents in Norway. The role of proactive and reactive aggressiveness. *J. Early Adolesc.* 29, 898–923. doi: 10.1177/0272431609332935

Fraire, M., Longobardi, C., and Sclavo, E. (2008). Contribution to validation of the student-teacher relationship scale (STRS) in the Italian educational setting. *Eur. J. Educ. Psychol.* 1, 49–59. doi: 10.30552/ejep.v1i3.12

Garandeau, C. F., Vartio, A., Poskiparta, E., and Salmivalli, C. (2016). School bullies' intention to change behavior following teacher interventions: effects of empathy arousal, condemning of bullying, and blaming of the perpetrator. *Prev. Sci.* 17, 1034–1043. doi: 10.1007/s11121-016-0712-x

Geerlings, J., Thijs, J., and Verkuyten, M. (2017). Student-teacher relationships and ethnic outgroup attitudes among majority students. *J. Appl. Dev. Psychol.* 52, 69–79. doi: 10.1016/j.appdev.2017.07.002

Graham, J. W. (2009). Missing data analysis: making it work in the real world. *Annu. Rev. Psychol.* 60, 549–576. doi: 10.1146/annurev.psych.58.110405.085530

Hayes, A. F. (2013). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York: Guilford Press.

Hymel, S., and Bonanno, R. A. (2014). Moral disengagement processes in bullying. *Theory into Pract.* 53, 278–285. doi: 10.1080/00405841.2014.947219

Italian Ministry of Education (2020). *Gli alunni con cittadinanza non italiana. A.S.* 2018–2019 [Students with non-Italian citizenship. Scholastic year 2018–2019]. Document Available at: https://www.miur.gov.it/documents/20182/2447435/Notiziario+Alunni+con+Cittadinanza+non+italiana+A.S.+2018_2019.pdf/ad84f9fc-efe5-46bd-2aa4-091b81727197?version=1.0andt=1593701066178 (Accessed January 19, 2021).

Jakobsen, J. C., Gluud, C., Wertterslev, J., and Winkel, P. (2017). When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. *BMC Med. Res. Methodol.* 17:162. doi: 10.1186/s12874-017-0442-1

Jennings, P. A., and Greenberg, M. T. (2009). The prosocial classroom: teacher social and emotional competence in relation to student and classroom outcomes. *Rev. Educ. Res.* 79, 491–525. doi: 10.3102/0034654308325693

Jerome, E., Hamre, B. K., and Pianta, R. C. (2009). Teacher-child relationships from kindergarten to sixth grade: early childhood predictors of teacher-perceived conflict and closeness. *Soc. Dev.* 18, 915–945. doi: 10.1111/j.1467-9507.2008.00508.x

Kawakami, K., Amodio, D. M., and Hugenberg, K. (2017). "Intergroup perception and cognition: An integrative framework for understanding the causes and consequences of social categorization," in *Advances in Experimental Social Psychology Vol. 55*. ed. J. M. Olson (United States: Academic Press), 1–80.

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

Koo, T. K., and Li, M. Y. (2016). Cracking the code: providing insight into the fundamentals of research and evidence-based practice. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* 15, 155–163. doi: 10.1016/j.jcm.2016.02.012

Larochette, A., Murphy, A. N., and Craig, W. M. (2010). Racial bullying and victimization in Canadian school-aged children: individual and school level effects. *Sch. Psychol. Int.* 31, 389–408. doi: 10.1177/0143034310377150

Levy, S. R., West, T. L., Ramirez, L. F., and Pachankis, J. E. (2004). "Racial and ethnic prejudice among children," in *The Psychology of Prejudice and Discrimination: Racism in America. Vol. 1*. ed. J. L. Chin (Westport, CT: Praeger Pub), 37–60.

Lo Cricchio, M. G., García-Poole, C., te Brinke, W., Bianchi, D., and Menesini, E. (2020). Moral disengagement and cyberbullying involvement: A systematic review. *Eur. J. Dev. Psychol.* 18, 271–311. doi: 10.1080/17405629.2020.1782186

Longobardi, C., Iotti, N. O., Jungert, T., and Settanni, M. (2018). Student-teacher relationships and bullying: The role of student social status. *J. Adolesc.* 63, 1–10. doi: 10.1016/j.adolescence.2017.12.001

McKenney, K. S., Pepler, D., Craig, W., and Connolly, J. (2006). Peer victimization and psychological adjustment: The experiences of Canadian immigrant youth. *Electron. J. Res. Educ. Psychol.* 4, 239–264. doi: 10.25115/ejrep.v4i9.1189

Murray, C., and Murray, K. M. (2004). Child level correlates of teacher-student relationships: An examination of demographic characteristics, academic orientations, and behavioral orientations. *Psychol. Sch.* 41, 751–762. doi: 10.1002/pits.20015

Nesdale, D. (2010). "Peer group rejection and children's intergroup prejudice," in *Intergroup Attitudes and Relations in Childhood through Adulthood*. eds. S.R. Levy and M. Killen (Oxford, England: Oxford University Press), 32–46.

Palladino, B. E., Nappa, M. R., Zambuto, V., and Menesini, E. (2020). Ethnic bullying victimization in Italy: The role of acculturation orientation for ethnic minority adolescents with differing citizenship statuses. *Front. Psychol.* 11:499. doi: 10.3389/fpsyg.2020.00499

Papotti, N., and Caravita, S. C. S. (2020). Bullismo etnico: chi sono coloro che aggrediscono compagni con background migratorio? [ethnic bullying: who attacks those with migratory background?]. *Ric. Psicol.* 43, 327–350. doi: 10.3280/RIP2020-001015

Pianta, R. C. (1994). Patterns of relationships between children and kindergarten teachers. *J. Sch. Psychol.* 32, 15–31. doi: 10.1016/0022-4405(94)90026-4

Raabe, T., and Beelmann, A. (2011). Development of ethnic, racial, and national prejudice in childhood and adolescence: A multinational meta-analysis of ages differences. *Child Dev.* 82, 1715–1737. doi: 10.1111/j.1467-8624.2011.01668.x

Richard, J. F., Schneider, B. H., and Mallet, P. (2011). Revisiting the whole-school approach to bullying: really looking at the whole school. *Sch. Psychol. Int.* 33, 263–284. doi: 10.1177/0143034311415906

Rose, C. A., Richman, D. M., Fettig, K., Hayner, A., Slavin, C., and Preast, J. L. (2016). Peer reactions to early childhood aggression in a preschool setting: defenders, encouragers, or neutral bystander. *Dev. Neurorehabil.* 19, 246–254. doi: 10.3109/17518423.2014.979955

Rosenberg, M. J., and Hovland, C. I. (1960). "Cognitive, affective, and behavioral components of attitude," in *Attitude Organization and Change: An Analysis of Consistency among Attitude Components*. eds. M. J. Rosenberg, C. I. Hovland, W. J. McGuire, R. P. Abelson and J. W. Brehm (New Haven, CT: Yale University Press), 1–14.

Rutland, A., Killen, M., and Abrams, D. (2010). A new social-cognitive developmental perspective on prejudice: The interplay between morality and group identity. *Perspect. Psychol. Sci.* 5, 279–291. doi: 10.1177/1745691610369468

Sabol, T. J., and Pianta, R. C. (2012). Recent trends in research on teacher-child relationships. *Attach Hum. Dev.* 14, 213–231. doi: 10.1080/14616734.2012.672262

Scherr, T. G., and Larson, J. (2010). "Bullying dynamics associated with race, ethnicity, and immigration status," in *Handbook of Bullying in Schools. An International Perspective*. eds. S. R. Jimerson, S. M. Swearer and D. L. Espelage (New York: Routledge), 223–234.

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika* 74, 107–120. doi: 10.1007/s11336-008-9101-0

Sklad, M., and Park, E. (2017). Examining the potential role of education in the prevention of radicalization from the psychological perspective. *Peace Conflict: J. Peace Psychol.* 23, 432–437. doi: 10.1037/pac0000258

Strohmeier, D., Kärnä, A., and Salmivalli, C. (2011). Intrapersonal and interpersonal risk factors for peer victimization in immigrant youth in Finland. *Dev. Psychol.* 47, 248–258. doi: 10.1037/a0020785

Tajfel, H., Billig, M. G., Bundy, R. P., and Flament, C. (1971). Social categorization and intergroup behavior. *Eur. J. Soc. Psychol.* 1, 149–178. doi: 10.1002/ejsp.2420010202

Tajfel, H., and Wilkes, A. L. (1963). Classification and quantitative judgment. *Br. J. Psychol.* 54, 101–114. doi: 10.1111/j.2044-8295.1963.tb00865.x

Teng, Z., Bear, G. G., Yang, C., Nie, Q., and Guo, C. (2020). Moral disengagement and bullying perpetration: A longitudinal study of the moderating effect of school climate. *Sch. Psychol.* 35, 99–109. doi: 10.1037/spq0000348

Thornberg, R., and Jungert, T. (2014a). School bullying and the mechanisms of moral disengagement. *Aggress. Behav.* 40, 99–108. doi: 10.1002/ab.21509

Thornberg, R., and Jungert, T. (2014b). School bullying and the mechanism of moral disengagement. *Aggress. Behav.* 40, 99–108. doi: 10.1002/ab.21509

Tolsma, J., van Deurzen, I., Stark, T. H., and Veenstra, R. (2013). Who is bullying whom in ethnically diverse primary schools? Exploring links between bullying, ethnicity, and ethnic diversity in Dutch primary schools. *Soc. Netw.* 35, 51–61. doi: 10.1016/j.socnet.2012.12.002

Tropp, L. R., and Pettigrew, T. F. (2005). Differential relationships between intergroup contact and affective and cognitive dimensions of prejudice. *Personal. Soc. Psychol. Bull.* 31, 1145–1158. doi: 10.1177/0146167205274854

Vaes, J., Paladino, M. P., Castelli, L., Leyens, J. P., and Giovanazzi, A. (2003). On the behavioral consequences of infrahumanization: The implicit role of uniquely human emotions in intergroup relations. *J. Pers. Soc. Psychol.* 85, 1016–1034. doi: 10.1037/0022-3514.85.6.1016

van Aalst, D. A. E., Huitsing, G., Mainhard, T., Cillessen, A. H. N., and Veenstra, R. (2021). Testing how teachers' self-efficacy and student-teacher relationships moderate the association between bullying, victimization, and student self-esteem. *Eur. J. Dev. Psychol.* 1–20. doi: 10.1080/17405629.2021.1912728

van Dijk, T.A. (1984). *Prejudice in Discourse. An Analysis of Ethnic Prejudice in Cognition and Conversation*. Amsterdam: John Benjamin Publishing Company.

Wachs, S., Görzig, A., Wright, M. F., Schubarth, W., and Bilz, L. (2020). Associations among adolescents' relationships with parents, peers, and teachers, self-efficacy, and willingness to intervene in bullying: A social cognitive approach. *Intern. J. Environ. Res. Public Health.* 17:420. doi: 10.3390/ijerph17020420

Wang, C., Swearer, S. M., Lembeck, P. T., Collins, A., and Berry, B. (2015). Teachers matter: An examination of student-teacher relationships, attitudes toward bullying, and bullying behavior. *J. Appl. Sch. Psychol.* 31, 219–238. doi: 10.1080/15377903.2015.1056923

Xu, M., Macrynikola, N., Waseem, M., and Miranda, R. (2020). Racial and ethnic differences in bullying: review and implication for intervention. *Aggress. Violent Behav.* 50:101340. doi: 10.1016/j.avb.2019.101340

Yoon, J., and Bauman, S. (2014). Teachers: A critical but overlooked component of bullying prevention and intervention. *Theory Into. Pract.* 53, 308–314. doi: 10.1080/00405841.2014.947226

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 10 September 2021
doi: 10.3389/fpsyg.2021.703468

Aggression, Moral Disengagement and Empathy. A Longitudinal Study Within the Interpersonal Dynamics of Bullying

Daniel Falla, Eva M. Romera and Rosario Ortega-Ruiz*

Department of Psychology, Universidad de Cordoba, Cordoba, Spain

Moral disengagement and empathy have been linked to aggression in traditional bullying. A number of longitudinal studies have focused on how these variables predict aggressive behavior within the dynamics of bullying. However, no conclusive results have been produced as to whether aggressive behavior in bullying can predict lower levels of empathy, and to date, no studies have explored in depth the mediating role of moral disengagement strategies in this relationship, which is the aim of this study. A total of 1,810 students (51.0% girls; $M_{age} = 14.50$; $SD = 1.05$) completed a survey in three waves at 6-month intervals. The results showed that aggressive behavior in bullying at Time 1 was inversely related to affective and cognitive empathy at Time 3. Minimization of responsibility, distortion of consequences and dehumanizing mediated in the aggressive behavior exhibited by the bullying aggressors and in cognitive empathy, while cognitive restructuring and the distortion of consequences mediated in affective empathy. We discuss the impact on moral and emotional sensitivity of the continued aggression occurring in the interpersonal dynamics of bullying, as well as the relationship between certain strategies of moral disengagement and the different types of empathy. We also comment on the need to design intervention programs to address the lowering of moral criteria and empathy in young people and adolescents involved in traditional bullying.

Keywords: aggression, moral disengagement, empathy, bullying, longitudinal design

OPEN ACCESS

Edited by:
Simona Carla Silvia Caravita, University of Stavanger, Norway

Reviewed by:
Barbara Colombo, Champlain College Neuroscience Lab, United States
Frank Guangzhe Yuan, University of Macau, China

****Correspondence:***
Eva M. Romera
eva.romera@uco.es

Specialty section:
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

Received: *30 April 2021*
Accepted: *18 August 2021*
Published: *10 September 2021*

Citation:
Falla D, Romera EM and Ortega-Ruiz R (2021) Aggression, Moral Disengagement and Empathy. A Longitudinal Study Within the Interpersonal Dynamics of Bullying. Front. Psychol. 12:703468. doi: 10.3389/fpsyg.2021.703468

INTRODUCTION

Aggressive behavior is a pattern of conduct whose adaptive origin is based on neurophysiological conditions, which are in turn modulated by processes of socialization (Blair, 2010). As a result, in most cases, aggressive behavior is controlled through cognitive and socio-affective processes which are derived from the competences which our brain uses in a global and coordinated way in its normal functions (Preckel et al., 2018). In this context, neuroscience has highlighted the key role played by empathy in regulating the processes which control aggressive impulses. Empathy, defined as the competence to register, recognize and experience the feelings and emotions of others (Weisz and Cikara, 2020), is a human characteristic, which like all the basic components of social behavior, is modulated and in most cases optimized for cognitive, emotional and socio-moral development throughout childhood and adolescence (see the systematic review by Silke et al., 2018). It is unquestionable that empathy plays a major role in the lives of groups, fostering the establishment of warm, affectionate and civic social relationships which help to stabilize the

ecosystem of coexistence in which the development and learning processes take place (Ortega-Ruiz, 2020).

Studies in morality and aggressive behavior (Molchanov, 2014; Romera et al., 2019a) also point to the role of empathy in sensitivity and the recognition of aggression in bullying as an immoral act (Ortega-Ruiz, 2010). A growing number of studies have pointed out the relationship of the mechanisms of moral disengagement in acts of bullying, and that these cognitive processes can be activated in situations and interpersonal behavior where a moral judgment is required (see the meta-analysis by Killer et al., 2019 and the study into low levels of empathy by Haddock and Jimerson, 2017). Cross-sectional studies suggest that the experiences of being an aggressor toward one's peers, within the complex dynamics of bullying, lead to a greater indifference toward the victims' feelings compared to the moral sensitivity reported by those who are not aggressors toward their peers (Romera et al., 2019b).

Aggression in Bullying and Moral Disengagement

Bullying is considered an intentional phenomenon involving persistent, unjustified aggression, and it is clear that this repeated abuse both damages the victim and lessens the aggressor's sensitivity and moral criteria, to such an extent that they become aware that what they are doing is morally reprehensible (Ortega-Ruiz, 2020). Aggressors use physical or psychological superiority to intimidate, mistreat and ultimately physically attack their victims in different ways, ranging from insulting or hitting to more sophisticated, relational forms of bullying, such as social exclusion (Menesini and Salmivalli, 2017). The prevalence figures indicate that about 36% of boys and girls bully their schoolmates with some frequency (Modecki et al., 2014), with boys being the most commonly involved and the most frequent perpetrators, although this male preponderance decreases in early adolescence (Smith et al., 2019).

It is undeniable that the unjustified, intentional and repetitive aggression that occurs in the dynamics of bullying includes ethical elements: in order to justify its repetition, the aggressor's repeated behavior and the roles of aggressor, victim and spectator, it is necessary to take a cynical view and deny the evidence that this aggression harms the victim. In this context, the definition of moral disengagement mechanisms proposed by Bandura (2002) is a key construct to help us progress in our understanding of the morally complex dynamics of the phenomenon of bullying. Bandura distinguished eight disengagement mechanisms, which were grouped into four domains or strategies as follows: (a) cognitive restructuring, which allows offenders to interpret behavior that is clearly immoral as fair or reasonable; (b) minimization of responsibility, which consists of disregarding or transferring responsibility for one's own antisocial actions; (c) distortion of the consequence, which permits the person committing immoral acts not to fully consider the impact their actions have on others; and (d) dehumanization, which is used to reject, undervalue or even blame the victim for what is happening. This model of four strategies of moral disengagement is particularly apt for understanding the unethical dimension of the type of aggression which occurs in the dynamics of bullying. A large number of scientific works have used this model (see meta-analysis by Gini et al., 2014; Killer et al., 2019) and some studies have revealed the possible socio-cultural intricacies of these strategies. For instance, according to Pornari and Wood (2010) in the cultural context of the United Kingdom, one common strategy is to minimize responsibility and the mechanisms related to cognitive restructuring, and to utilize euphemistic language to justify the facts; Scandinavian studies have also observed attribution of blame to the victim in order to justify such conduct (Thornberg and Jungert, 2014; Bjärehed et al., 2020). In Poland, mechanisms related to cognitive restructuring (advantageous compassion, euphemistic labeling and moral justification) and distortion of consequences were associated with the perpetration of bullying (Zych and Llorent, 2019). In Australia, Runions et al. (2019) pointed out that bullies commonly made use of the mechanisms of minimizing responsibility, distorting consequences and euphemistic labeling. In Spain, Romera et al. (2020) found that the disengagement mechanisms associated with bullying were dehumanization, distortion of consequences and cognitive restructuring, and reported that this last strategy was the one most closely linked to the aggressor's behavior.

Previous studies, taking the mechanism of moral disengagement as a one-dimensional construct, have shown that higher scores in perpetration in bullying predict higher scores in moral disengagement (Obermann, 2013; Thornberg et al., 2019). However, to date, no longitudinal studies have explored the possible influence of aggressive behavior within bullying on the different mechanisms of moral disengagement.

Aggression and Empathy in the Bullying Phenomenon

As stated above, empathy has been defined as the human competence to recognize and experience the feelings and emotions of others. Traditionally, two types of empathy have been identified, depending on the importance of the more cognitive-rational or affective-emotional aspects in the process of putting oneself in another person's position. These types have been associated with two differential kinds of neural processing: cognitive and affective (Healey and Grossman, 2018), which do not act completely independently, but which can be differentiated behaviourally. The type known as affective empathy, which involves elements of emotional contagion, in which one is "infected" by another's emotions (Cuff et al., 2016), appears to involve subcortical structures such as the limbic lobe (Derntl et al., 2010), while cognitive empathy seems to stimulate the activation of the pre-frontal and ventromedial cortex (Decety, 2011), which permits a certain prevalence of reflective and perhaps rational thinking. Thus, cognitive empathy allows us to understand not only the emotions and feelings of others, in line with what has been called the Theory of Mind (Healey and Grossman, 2018), but also to realize that the other person is a human being similar to oneself and who, therefore, can be expected to think, feel and behave as one does, or decide not to. In the face of conflicts of interest and rivalries, empathy allows us, on the one hand, to sympathize with the feelings of

the sufferer and, on the other, to hypothesize to what extent the response of the other, in the heat of the conflict, will be aggressive or peaceful. Thus, it is to be expected that empathy which is both cognitive and affective, socialized and adjusted to socio-moral norms works as a control and modulation mechanism in conflict dynamics (Klimecki, 2019) which may include aggressive behavior (Tampke et al., 2020).

Research has revealed that interpersonal situational contexts significantly influence an individual's empathic processing (Cheng et al., 2017). In socially complex interpersonal dynamics such as bullying, where sustained hostility plays an important role, empathy and aggression seem to interact (van Noorden et al., 2015). In addition, studies based on the cycle of violence (Widom, 1989) have shown how basic social skills, fundamental to the development of empathy, are impaired in hostile contexts where there is exposure to ongoing violence and abuse (Heleniak and McLaughlin, 2020). Similarly, social cognitive theory points out how context can affect socio-cognitive reasoning and thus affective processing, including empathy (Bandura, 2002). Recent meta-analyses have shown that low empathy is related to a higher tendency toward aggression (Zych et al., 2019), although other meta-analyses (Vachon and Lynam, 2013) have produced conflicting results, albeit among adults. In longitudinal studies, some researchers have found that aggression in bullying influences cognitive empathy (Williford et al., 2016), while other studies found no direct relationship between bullying and empathy (Walters and Espelage, 2018) and others showed that, in bullying, empathy and aggression are bidirectionally related (Stavrinides et al., 2010). In short, although cross-sectional studies indicate that aggression and empathy are related, more longitudinal work remains to be done to test whether bullying may be related to lower affective and cognitive empathy scores in the medium to long term.

Aggression in Bullying, Moral Disengagement and Empathy

According to the general aggression model (GAM) individuals behave aggressively due to the interaction of personal and situational factors, internal states and outcomes of evaluation and decision-making processes (Dewall et al., 2011). This multi-causal influence is also supported by social-cognitive theory that describes how there is also a bidirectional and reciprocal relationship between morality and aggression (Bandura, 2002). That is, hostile contexts can affect the moral judgement of individuals. Thus, aggression in bullying, as mentioned above, predicts higher scores on moral disengagement (Obermann, 2013; Thornberg et al., 2019). Considering the relationship described by Bandura (2002) between morality and affective processes such as empathy, the studies coming from neuroscience describing the existence of a social brain where morality and empathy are interconnected (Detert et al., 2008; Chen et al., 2018) and the evidence from developmental and educational psychology that supports an inverse relationship between moral disengagement and affective and cognitive empathy (Haddock and Jimerson, 2017), it would be plausible to hypothesize that aggression may also affect these socioemotional and socio-cognitive skills in the medium term. Thus, mechanisms of moral disengagement, in addition to preventing individuals from feeling negative emotions when committing transgressions (Mazzone et al., 2019), may lead to a decrease in affective and cognitive empathy. For example, in a previous study it was found that young people who had experiences as aggressors, tended to point to victims as indifferent to aggression (Romera et al., 2019b). This tendency may suggest that when as children engage in bullying, their ability or interest in taking the perspective of others (cognitive empathy) may be diminished (Haddock and Jimerson, 2017). There is also work indicating that a lack of emotional contagion or disconnection may occur among schoolchildren, as a type of adaptive response to avoid feeling negative emotions in maladaptive situations (Herrera-López et al., 2017).

Although there seem to be no studies examining which mechanisms of moral disengagement are most affected by aggression in bullying, some work has found a particularly important role played by cognitive restructuring in aggression in bullying (Falla et al., 2020). It therefore seems sensible to hypothesize that justifying or normalizing aggressive behavior, as well as inhibiting the negative emotions that transgressing the social norm would entail (Mazzone et al., 2019), could prevent recognition of the victim's emotions and also emotional contagion from occurring, even to the point of experiencing positive emotions for aggressing (Perren et al., 2012). Similarly, for some authors, ignoring or distorting consequences allows aggressors to disassociate themselves from the emotional harm of harmful actions, which may affect both empathies, so that the aggressor may infer that the victim accepts the aggressive behavior as a joke (Runions and Bak, 2015). Finally, attributing blame to the victim or dehumanizing the victim, in addition to holding the victim responsible for the behavior, leads to invalidating the victim's emotions and prevents emotional contagion. In this way, perpetrators become less likely to empathize with the victim and instead more motivated to hurt them (Haslam and Loughnan, 2014).

Aims of the Study

There is evidence of a relationship between aggressive behavior in bullying and the mechanisms of moral disengagement (Killer et al., 2019), as well as between these cognitive processes and empathy (Haddock and Jimerson, 2017). However, as yet, there is no proof whether moral disengagement strategies exert any mediating effect between aggression in bullying and cognitive and affective empathy, and whether this impact is sustained over time. The aims of this study are: (1) to explore whether there is a relationship between aggression in bullying in Time 1 and cognitive and affective empathy in Time 3; (2) to examine whether certain strategies of moral disengagement exert any mediating impact between aggressive behavior and the levels of cognitive and affective empathy over all three time measures. For this, we followed Mediation Model 4 (Hayes, 2013) and the following hypotheses were proposed:

H1. There will be a negative relationship between aggressive behavior in bullying and cognitive and affective empathy, which is sustained significantly over time.

H2. The strategies of moral disengagement, cognitive restructuring, distortion of consequences and dehumanization will mediate the relationship between aggressive behavior in bullying and cognitive and affective empathy.

METHODS

Participants

Thirteen schools in southern Spain (five urban and eight rural) were selected for accessibility using non-probabilistic sampling (Singleton and Straits, 2004). Although the data came from different classrooms and centers, they were taken as non-nested due to the statistical analyses used. The longitudinal study included three time waves, each six months apart, with a total period of one year between the first and the third. Time 1, between April and May 2018, involved 2,360 students (50.1% girls; $M_{age} = 13.58$; $SD = 1.13$) with the following distribution by school years: 7th (35.4%), 8th (33.6%) and 9th (31.0%). Time 2 took place between October and November 2018, with a retention rate of 86.06% ($N = 2031$) with (51.2% girls; $M_{age} = 13.97$; $SD = 1.04$), while the distribution by school years was: 7th (2.8%), 8th (37.2%), 9th (31.0%) and 10th (29.0%). In Time 3, the questionnaires were completed between April and May 2019, with a retention rate of 76.69% ($N = 1810$) with (51.0% girls; $M_{age} = 14.50$; $SD = 1.05$) and with the following distribution by school years: 7th (1.9%), 8th (37.8%), 9th (30.6%) and 10th (29.7%). The decrease in the total sample between waves was due to the fact that some schoolchildren did not attend on the day the survey was administered and others had changed school. Logistic regression was performed to check whether the analytical longitudinal sample was representative of the total sample, and no significant differences were found (all $ps > 0.05$) in the study variables between any of the three time periods.

Instruments

Empathy was measured using *The Basic Empathy Scale* (Jolliffe and Farrington, 2006). This scale contains 20 items, with a Likert scale from one to five (1 = *strongly disagree* to 5 = *strongly agree*) distributed in two factors: cognitive empathy (nine items) (e.g., "I can understand my friend's happiness when he/she does something well") and affective empathy (eleven items) (e.g., "After spending time with a friend who is upset about something, I usually feel sad"). The reliability analyses were acceptable for both cognitive empathy ($\omega_{T1} = 0.75$, $\omega_{T2} = 0.77$, $\omega_{T3} = 0.80$) and affective empathy ($\omega_{T1} = 0.77$, $\omega_{T2} = 0.77$, $\omega_{T3} = 0.78$).

Aggression in bullying was measured using the Spanish version of *the European Bullying Intervention Project Questionnaire* (EBIPQ) (Ortega-Ruiz et al., 2016). This scale is made up of 14 items, referring to the last two months and divided into two factors, and is scored on a Likert scale 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*). For the current study, only the "aggression" factor was used, which is made up of seven items (e.g., "I have excluded or ignored someone"). Omega coefficients were good for all three time periods ($\omega_{T1} = 0.81$, $\omega_{T2} = 0.81$, $\omega_{T3} = 0.78$).

The mechanisms of moral disengagement were measured using the *Mechanisms of Moral Disengagement Scale* (Caprara et al., 1996). The version used consisted of 24 items with five Likert-type response options, from 1 to 5 (1 = *strongly disagree*; 2 = *partially agree*; 3 = *generally agree*; 4 = *strongly agree*; 5 = *totally agree*), which were divided into four factors. The factorial structure of this instrument has been confirmed by Pozzoli et al. (2012). The domains were: cognitive restructuring (e.g., "It's okay to use force against a partner who insults your family"), minimizing responsibility (e.g., "You can't blame kids for swearing at their peers because most of their friends do it"), distorting the consequences (e.g., "Making fun of a classmate is not really hurting him") and dehumanizing (e.g., "There's nothing wrong with treating someone badly if they behave in a contemptible way"). The reliability analyses were acceptable: cognitive restructuring ($\omega_{T1} = 0.83$, $\omega_{T2} = 0.83$, $\omega_{T3} = 0.85$); minimization of responsibility ($\omega_{T1} = 0.70$, $\omega_{T2} = 0.73$, $\omega_{T3} = 0.75$); distortion of consequences ($\omega_{T1} = 0.59$, $\omega_{T2} = 0.60$, $\omega_{T3} = 0.66$); and dehumanization ($\omega_{T1} = 0.76$, $\omega_{T2} = 0.76$, $\omega_{T3} = 0.79$).

Procedure

The Ethics Committee (who remained anonymous) previously approved the project used to carry out this study. First, we contacted the secondary schools to explain the objectives of the study and request their participation. This was agreed by the schools' councils and, next, letters of consent were sent out to the families. Once permission had been obtained from the schools and families, the dates for conducting the survey were set.

The survey was administered in the classroom: one of the researchers explained the procedure and reminded the children of the anonymous, voluntary nature of the study. In addition, the researcher explained how to fill in the code required to be able to carry out the longitudinal study. Any children who did not want to fill in the questionnaires remained in the classroom, and the children were given approximately 30 min to complete the questionnaires.

Data Analysis

The descriptive analyses included means, standard deviations, bivariate correlations and Student's *t* and Cohen's *d tests* to determine gender differences and effect size, using IBM SPSS Statistics Version 26 (IBM, Armonk, NY, USA). The mediation analysis was performed using the PROCESS v3.4 macro for SPSS (SPSS Inc., Chicago, IL, USA), and all the variables used were standardized to be able to make comparisons of the effects. Model 4 was used following Hayes (2013), and the MacKinnon (2008) four-step procedure was followed. The variables used were aggression in bullying at Time 1 as a predictor variable, cognitive restructuring at Time 2 as the first mediator, minimization of responsibility as the second mediator, distortion of consequences as the third mediator, dehumanization as the

fourth mediator, cognitive empathy as a dependent variable for the first model and affective empathy as a dependent variable for the second model. Gender and age were used as covariates in all the analyses.

Indirect effects were tested with the bootstrapping method, in which the values were considered significant when the confidence intervals did not include zero. This method is optimal for linear hypotheses when the variables do not have a normal distribution (Chernick, 2008). The relationship between the independent and dependent variable enabled us to find the total effect, while the mediation effect was calculated between the indirect effect and the total effect (Wen and Fan, 2015).

RESULTS

Descriptive Results

The correlations between all the study variables were checked for the three time periods. A direct relationship was found between all three, except between age and the bullying-perpetration variables T2 and T3, and between minimization of responsibility T2 and T3 and gender (see **Supplementary Material**). The assumption of multicollinearity was not violated since the VIF was <2.42 in all variables. Similarly, the Student's *t tests* allowed us to verify the existence of significant differences between boys and girls for all the study variables. The girls scored higher in the two dimensions of empathy while boys obtained higher marks for the rest of the variables. The effect sizes were low to moderate (see **Table 1**).

TABLE 1 | Means, standard deviations and differences by gender for all variables.

	Sample		Boys		Girls			
	M	*SD*	*M*	*SD*	*M*	*SD*	*t*	*d*
BP T1	0.27	0.42	0.32	0.49	0.21	0.35	5.23**	0.26
CR T1	1.53	0.58	1.69	0.64	1.39	0.48	10.89**	0.53
MR T1	1.79	0.72	1.86	0.77	1.72	0.66	4.18**	0.20
DC T1	1.38	0.57	1.50	0.67	1.28	0.42	7.89**	0.40
DH T1	1.49	0.61	1.61	0.70	1.37	0.48	8.36**	0.42
CE T1	4.07	0.55	3.94	0.56	4.19	0.51	−9.69**	0.47
AE T1	3.55	0.65	3.32	0.66	3.76	0.57	−14.02**	0.72
BP T2	0.19	0.37	0.25	0.45	0.14	0.28	5.98**	0.30
CR T2	1.51	0.57	1.67	0.65	1.37	0.45	11.27**	0.54
MR T2	1.72	0.71	1.77	0.75	1.68	0.66	2.68*	0.13
DC T2	1.36	0.56	1.45	0.64	1.28	0.46	6.54**	0.31
DH T2	1.46	0.59	1.58	0.68	1.36	0.47	7.57**	0.38
CE T2	4.08	0.56	3.94	0.60	4.21	0.50	−9.92**	0.49
AE T2	3.55	0.64	3.30	0.63	3.76	0.57	−15.26**	0.77
BP T3	0.20	0.36	0.25	0.42	0.16	0.29	5.03**	0.25
CR T3	1.53	0.60	1.69	0.67	1.39	0.48	10.52**	0.52
MR T3	1.71	0.71	1.76	0.76	1.67	0.65	2.76*	0.13
DC T3	1.35	0.56	1.46	0.65	1.25	0.43	8.07**	0.38
DH T3	1.45	0.60	1.58	0.69	1.34	0.47	8.13**	0.41
CE T3	4.09	0.56	3.96	0.57	4.21	0.51	−9.50**	0.46
AE T3	3.56	0.64	3.32	0.63	3.78	0.56	−15.57**	0.77

*M, mean; SD, standard deviation; t, student's t; d, cohen's d; AG, bullying perpetration; CR, cognitive restructuring; MR, minimizing responsibility; DC, distorting consequences; DH, dehumanizing; CE, cognitive empathy, AE, affective empathy; *p < 0.05; **p < 0.001.*

Mediation Analysis for Cognitive Empathy Model

The mediation analyses were carried out using Model 4 (Hayes, 2013), and proved that the effect of bullying-perpetration in T1 (predictor variable) on cognitive empathy in T3 (dependent variable) was mediated by cognitive restructuring (mediator 1), minimization of responsibility (mediator 2), distortion of consequences (mediator 3) and dehumanization (mediator 4) was significant: $F_{(7,1469)} = 22.78$; $R^2 = 0.10$; $p < 0.001$. The data indicated a direct and negative relationship of bullying-perpetration in T1 on cognitive empathy in T3 ($\beta = -0.16$, $t = -6.09$, $p < 0.01$). In step 2, there was a direct, positive association of bullying perpetration with cognitive restructuring ($\beta = 0.30$, $t = 12.91$, $p < 0.01$), minimization of responsibility ($\beta = 0.15$, $t = 5.61$, $p < 0.01$), distortion of the consequences ($\beta = 0.23$, $t = 8.92$, $p < 0.01$) and dehumanization ($\beta = 0.27$, $t = 10.35$, $p < 0.01$). In step 3, a direct and positive relationship was found for minimization of responsibility in T2 with cognitive empathy in T3 ($\beta = 0.09$, $t = 2.66$, $p < 0.01$), together with a direct, negative relationship of the distortion of the consequences ($\beta = -0.09$, $t = -2.60$, $p < 0.01$) and dehumanization ($\beta = -0.11$, $t = -2.74$, $p < 0.01$) on cognitive empathy T3. In addition, a negative relationship was found between bullying perpetration and cognitive empathy ($\beta = -0.13$, $t = -4.44$, $p < 0.01$) (see **Figure 1**).

The percentile bootstrap method with bias correction indicated a positive relationship of the indirect effect of bullying perpetration T1 on cognitive empathy in T3 through the minimization of responsibility ($\beta = 0.013$, 95% CI = [0.003, 0.03]) and a negative effect by the pathways of distortion of consequences ($\beta = -0.02$, 95% CI = [−0.04, −0.004]) and dehumanization ($\beta = -0.03$, 95% CI = [−0.05, −0.004]). Thus, the mediation effect was 7.93% in minimization of responsibility, 12.93% in distortion of the consequences and 17.32% in dehumanization.

Analysis of Mediation Effect for the Affective Empathy Model

Model 4 (Hayes, 2013) was also significant: $F_{(7,1457)} = 49.12$; $R^2 = 0.19$; $p < 0.001$ for the T1 perpetration bullying model on affective empathy in T3 mediated by cognitive restructuring, minimization of responsibility, distortion of consequences and dehumanization. Firstly, the bullying perpetration variable in T1 exerted a direct, negative relationship on cognitive empathy in T3 ($\beta = -0.17$, $t = -6.22$, $p < 0.01$). Secondly, the predictor variable had a direct, positive relationship with the four mediators: cognitive restructuring ($\beta = 0.31$, $t = 12.62$, $p < 0.01$), minimization of responsibility ($\beta = 0.16$, $t = 5.58$, $p < 0.01$), distortion of consequences ($\beta = 0.24$, $t = 8.99$, $p < 0.01$) and dehumanization ($\beta = 0.29$, $t = 10.71$, $p < 0.01$). Thirdly, the data showed a direct relationship for cognitive restructuring ($\beta = -0.14$, $t = -3.45$, $p < 0.01$) and the distortion

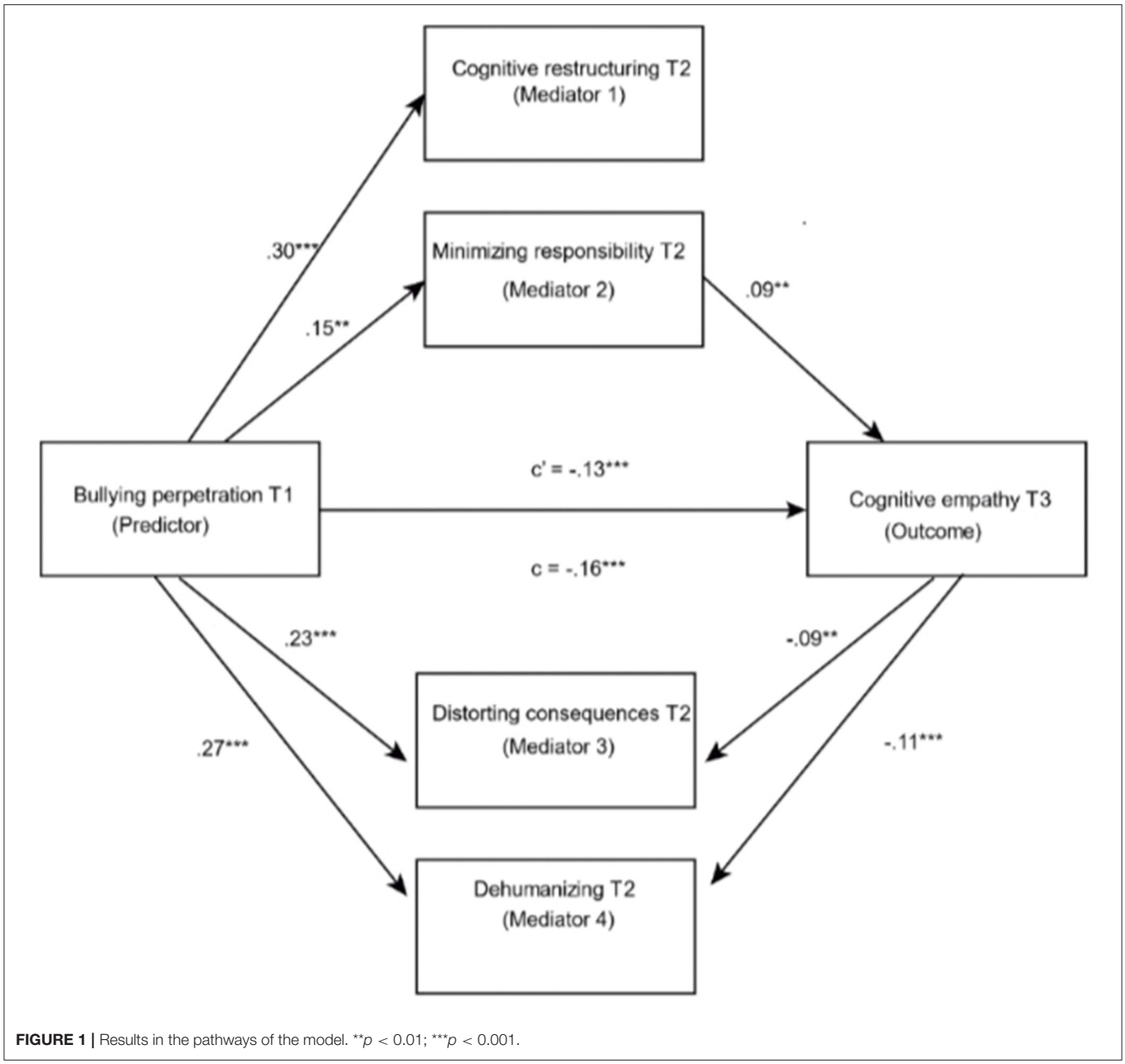


FIGURE 1 | Results in the pathways of the model. $^{**}p < 0.01$; $^{***}p < 0.001$.

of consequences ($\beta = -0.07$, $t = -2.14$, $p < 0.01$) with affective empathy, while bullying perpetration also correlated negatively with affective empathy ($\beta = -0.10$, $t = -3.57$, $p < 0.01$) (see **Figure 2**).

The analyses using the percentile bootstrap method with correction for bias showed a negative relationship for the indirect effect of bullying perpetration T1 on affective empathy in T3 via the cognitive restructuring pathways ($\beta = -0.04$, 95% CI $= [-0.07, -0.02]$) and distortion of consequences ($\beta = -0.02$, 95% CI $= [-0.04, -0.001]$). In this way, cognitive restructuring accounted for 25.21% of the mediation effect, while distortion of consequences contributed 10.47%.

DISCUSSION

The prolonged aggression that occurs among young people and adolescents, in which the bullying aggressor dominates in a cruel, unjustified way, producing a number of negative effects on the victim's moral integrity, constitutes immoral behavior, as it contradicts the general tendency toward socialization, which aims to find ways of avoiding violence following the ethical principles of society. In turn, the process of socialization appears to rely on human competence in order to be sensitive to the feelings of others, through a cognitive and affective process known as empathy. Empathy has especially attracted

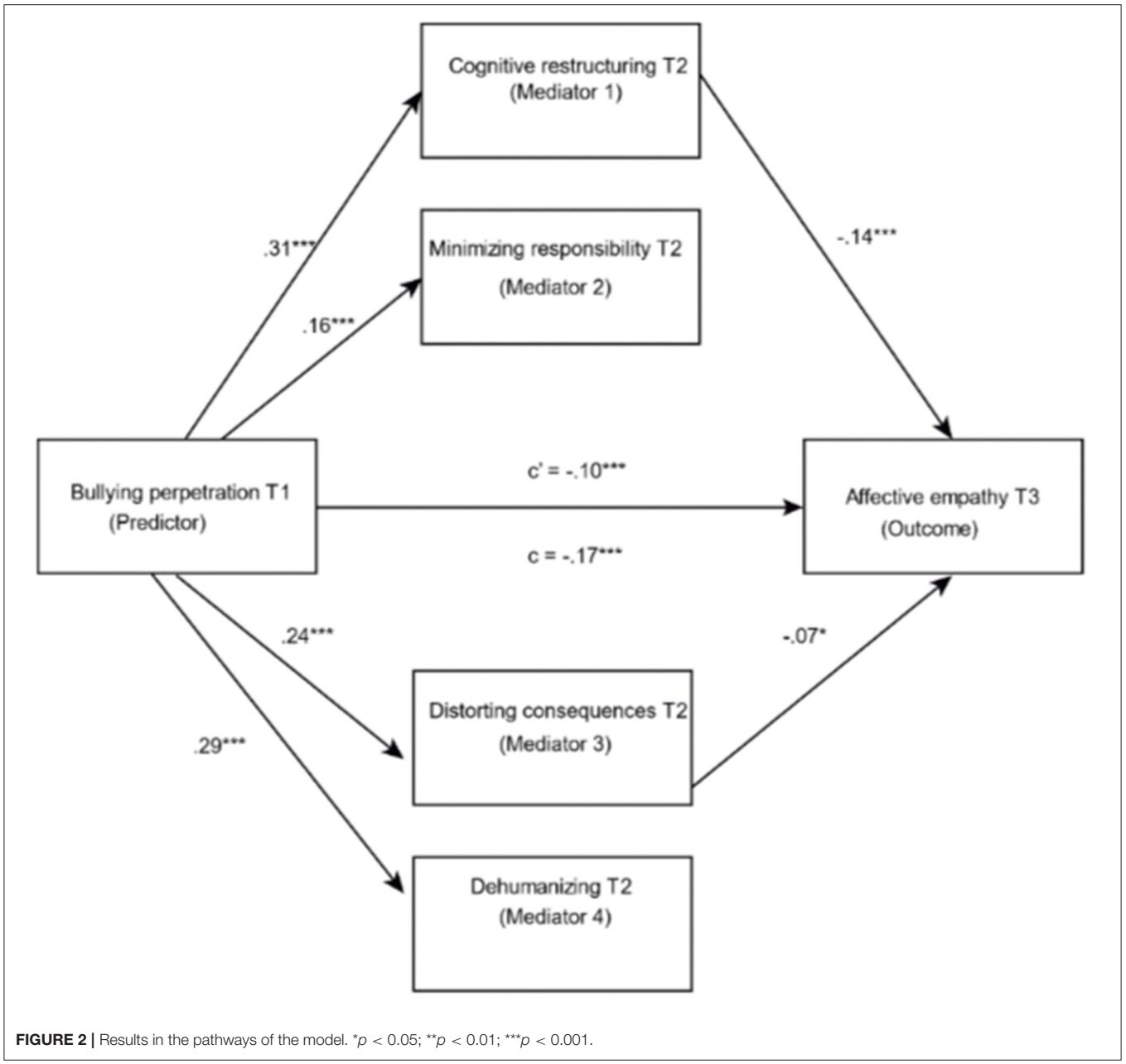


FIGURE 2 | Results in the pathways of the model. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

the interest of researchers into bullying as a protective factor against aggressive behavior (Garandeau et al., 2021). In addition, neuroscientific studies using neuroimaging inform us that moral judgments and empathic processing depend on past experiences (see meta-analysis by Bzdok et al., 2012). Similarly, it has been found that empathy is susceptible to the situational social context (Cheng et al., 2017) and that empathy, moral development and aggressive behavior are closely related (Blair, 2010). Based on these neurophysiological and psychological theoretical foundations, therefore, we hypothesized in this work that in the hostile, sustained context of bullying, the perpetrators will increase their mechanisms of moral disengagement and this will be related to lower levels of affective and cognitive empathy.

The first of the hypotheses was confirmed. It was shown that the aggressive behavior that occurs in bullying at Time 1 is related to low cognitive and affective empathy at Time 3. We know from research in the field of neurosciences that humans are sometimes able to avoid the emotional contagion which occurs through affective empathy to protect themselves from negative emotions such as guilt, pain or anguish (Lamm et al., 2007; Bensalah et al., 2016). This could happen in the case of the perpetrators in bullying, who end up dissociating themselves from the emotional contagion derived from affective empathy. However, these aggressors may also end up having less understanding of other people's thoughts and feelings, as Williford et al. (2016) found in their longitudinal study, and as

found by cross-sectional studies in elementary school students who had participated in bullying as aggressors indicated (when represented in vignettes showing aggressor-victim dynamics), these experiences were not related to the suffering they were made to feel (Romera et al., 2019b). In other words, these findings suggest that the continuous aggression over time found in this type of bullying can lead to the aggressors to lower both types of empathy toward the victim, which clouds their judgment about the suffering that they are causing the other person. This lack of emotional sensitivity, which has been studied as emotional disengagement in other works on empathy and normative adjustment of young people and adolescents (Herrera-López et al., 2017), as well as in research on cognitive neuroscience (Decety and Svetlova, 2012), could be a key factor to explain both the sustained repeated aggression toward the victim and the activation of a mechanism to disengage from the critical moral judgment about the behavior itself: that is, a mechanism of moral disengagement.

The second hypothesis was partly confirmed. The distortion of consequences, whose relationship with aggressive behavior in bullying has already been proved in cross-sectional studies (Runions et al., 2019; Romera et al., 2020), was the only mechanism of moral disengagement which mediated between the sustained aggression and the levels of cognitive and affective empathy, throughout the three periods of this longitudinal study. It is known that moral disengagement strategies are used to evade uncomfortable feelings of guilt and moral responsibility in aggressive behavior. It may be that, in an attempt to avoid these feelings, they elude the emotional contagion and solidarity implicit in affective and cognitive empathy: to do this, the consequences of harm have to be distorted. The strategies of minimizing responsibility and dehumanizing the victim also mediated in the cognitive empathy scores. In other words, the act of avoiding responsibility for the aggression or attributing the blame to the characteristics of the victim end up influencing the interpretation made by the aggressors about the feelings and thoughts of the victims themselves. However, in the case of responsibility minimization, it is a mechanism that requires recognition of the harm done to another person and the attribution of responsibility to others in order to alleviate feelings of guilt, hence its possible positive relationship with mechanisms that involve knowing how another person feels. However, the relationship is weak and should be further explored in future research. Whereas, in the case of dehumanization, the relationship is negative, which could explain certain phenomena such as discriminatory bullying (Rodríguez-Hidalgo et al., 2019), in which the aggressors harass the victims for reasons of gender, disability, race or cultural ethnicity, and this attribution of dehumanization and guilt about their defects could reinforce the moral impunity of the aggressors. On the other hand, cognitive restructuring and the distortion of consequences mediated in the scores for affective empathy. It seems logical to assume that these two strategies of moral disengagement, which involve a cold, external attribution, allow the bully to avoid feeling sympathy and empathy with those who suffer at their hands and avoid the emotional contagion and solidarity that could result from it; however, at the same time, it uses moral cynicism to reinforce the behavior of repeated, sustained aggression over time. These findings are consistent with previous studies which point to the importance of moral emotions in motivating pro-social behavior and an individual's moral self-concept (Christner et al., 2020).

Limitations and Practical Implications

This work has certain limitations which must be taken into account. The study sample was taken from a single country, so studies including samples from different countries could, firstly, confirm whether these processes are universal and, secondly, provide valuable information about the cultural element in processes which have an implicit moral criterion and shared values present in all interpersonal dynamics and, particularly, in bullying (Ortega-Ruiz, 2020). On the other hand, the use of self-reports as the sole source of data may lead to response bias, which could be resolved by conducting experimental or qualitative studies. Similarly, as this was an exploratory study and its purpose was to find out the medium-term interactions between aggressiveness in bullying, mechanisms of moral disengagement and affective and cognitive empathy, as well as the indirect effects of the mediating variables, a mediation analysis with the PROCESS macro was used. However, running path analyses using other software to deal with the non-normality of the variables and including both criterion variables in the same model, or even nesting the sample, could provide more robust results on these interactions. Also, the criterion variables, affective and cognitive empathy, were not controlled for at time 1. This is an important limitation that would have diminished the strength of the association between the variables, so we recommend that these limitations be addressed in future studies.

Despite the limitations, this study is a first step to consider the longitudinal interplay between aggression, moral disengagement strategies and empathy and provides relevant findings which further our knowledge in the complex interpersonal dynamics that take place in bullying, which, as stated above, is a clearly immoral, unfair and repetitive type of aggression. Although empathy has already been studied as a relevant protective factor in bullying (Garandeau et al., 2021), this work broadens this knowledge and argues that the low affective and cognitive empathy of the perpetrators of bullying possibly results from the continued experience of engaging in aggressive behavior which is clearly unfair and which infringes the general principles of socialization, which stress the importance of fair, respectful treatment toward others. If, as the data seem to show, the aggressors in bullying make use of mechanisms of moral disengagement without interruption over an extended period of time (remember here that bullying is not one specific event, but a persistent, repeated action), this will clearly lead to a lack of empathy. In the particular case of the mechanism for distorting the consequences for the victim, it is evident that an important cynical bias is at work in the moral criterion which mediates the main aspects of critical judgment. The findings of this work therefore show that the immorality and the deterioration of empathy experienced by the aggressors should also be addressed in specific programs as a consequence of the continuous aggression.

This work, in line with the contributions of neuroscientific research, highlights the close relationship between empathic processes and moral judgments, which can be especially useful in preventive and palliative intervention programs. For instance, mechanisms of moral disengagement focused on normalizing behavior and reducing consequences are linked to low affective empathy, so intervention programs could specifically focus on working on these aspects together in aggressors who show low affective empathy. On the other hand, cognitive empathy shows that specific work needs to be done to recognize the humanization of the victims and to encourage the self-recognition of responsibility for the harm caused to others.

CONCLUSIONS

The results of this work reinforce the close ties between cognitive and affective empathy and moral disengagement strategies, and show that the deterioration of empathy and high levels of moral disengagement strategies may be the consequence of the repeated use of these strategies and the deterioration of empathic sensitivity. The combined action of both processes reinforcing each other could account for the fact that the profile of the aggressor in bullying is dangerously far from the expectations of socialization and therefore of the control over their own behavior which moral principles dictate. In a nutshell, encouraging young people and adolescents to develop more critical, ethical thinking which is more supportive toward others requires a major effort of emotional and moral sensitivity to generate motivation to repair the damage caused, which seems an unlikely outcome if the strategies of moral disengagement are stubbornly perpetuated, or are increased. Dialogue, the peaceful resolution of conflicts and the beneficial effect of good friendships within the framework of interpersonal relationships are some of the elements of the social context that can have a palliative effect. However, as shown in this work, continuous, uninterrupted aggression which is not controlled by the context seems to reinforce in the aggressor biased moral judgments which are disengaged from the sensitivity and empathy toward others that civilized socialization requires, and intervention programs should focus their preventive and palliative work on this area.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The Ethics Committee of Bioethics and Biosafety at University of Córdoba. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

All authors contributed to the interpretation of data, helped to draft, revise the manuscript, read, and approved the final manuscript.

FUNDING

ER and RO-R were supported by project PSI2016-74871-R and DF, ER, and RO-R by project PSI2020-113911-RB-I00 (National Research Agency, Ministerio de Ciencia e Innovación, Spain).

ACKNOWLEDGMENTS

To the schools and children who participated in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.703468/full#supplementary-material

REFERENCES

Bandura, A. (2002). Selective moral disengagement in the exercise of moral agency. *J. Moral Educ.* 31, 101–119. doi: 10.1080/0305724022014322

Bensalah, L., Stefaniak, N., Carre, A., and Besche-Richard, C. (2016). The basic empathy scale adapted to French middle childhood: structure and development of empathy. *Behav. Res. Methods* 48, 1410–1420. doi: 10.3758/s13428-015-0650-8

Bjärehed, M., Thornberg, R., Wänström, L., and Gini, G. (2020). Mechanisms of moral disengagement and their associations with indirect bullying, direct bullying, and pro-aggressive bystander behavior. *J. Early Adolesc.* 40, 28–55. doi: 10.1177/0272431618824745

Blair, R. J. R. (2010). "Empathy, moral development, and aggression: a cognitive neuroscience perspective," in *Emotions, Aggression, and Morality in Children: Bridging Development and Psychopathology* (Worcester, MA: American Psychological Association), 97–114. doi: 10.1037/12129-005

Bzdok, D., Schilbach, L., Vogeley, K., Schneider, K., Laird, A. R., Langner, R., et al. (2012). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Struct. Funct.* 217, 783–796. doi: 10.1007/s00429-012-0380-y

Caprara, G. V., Barbaranelli, C., Vicino, S., and Bandura, A. (1996). La misura del disimpegno morale in età evolutiva. *Età Evolutiva* 51, 18–29.

Chen, C., Martínez, R. M., and Cheng, Y. (2018). The developmental origins of the social brain: empathy, morality, and justice. *Front. Psychol.* 9:2584. doi: 10.3389/fpsyg.2018.02584

Cheng, Y., Chen, C., and Decety, J. (2017). How situational context impacts empathic responses and brain activation patterns. *Front. Behav. Neurosci.* 11:165. doi: 10.3389/fnbeh.2017.00165

Chernick, M. R. (2008). *Bootstrap Methods: A Guide for Practitioners and Researchers, 2nd Edn*. Hoboken, NJ: Wiley.

Christner, N., Pletti, C., and Paulus, M. (2020). Emotion understanding and the moral self-concept as motivators of prosocial behavior in middle childhood. *Cogn. Dev.* 55:100893. doi: 10.1016/j.cogdev.2020.100893

Cuff, B. M. P., Brown, S. J., Taylor, L., and Howat, D. J. (2016). Empathy: a review of the concept. *Emot. Rev.* 8, 144–153. doi: 10.1177/1754073914558466

Decety, J. (2011). Dissecting the neural mechanisms mediating empathy. *Emot. Rev.* 3, 92–108. doi: 10.1177/1754073910374662

Decety, J., and Svetlova, M. (2012). Putting together phylogenetic and ontogenetic perspectives on empathy. *Dev. Cogn. Neurosci.* 2, 1–24. doi: 10.1016/j.dcn.2011.05.003

Derntl, B., Finkelmeyer, A., Eickhoff, S., Kellermann, T., Falkenberg, D. I., Schneider, F., et al. (2010). Multidimensional assessment of empathic abilities: neural correlates and gender differences. *Psychoneuroendocrinology* 35, 67–82. doi: 10.1016/j.psyneuen.2009.10.006

Detert, J. R., Treviño, L. K., and Sweitzer, V. L. (2008). Moral disengagement in ethical decision making: a study of antecedents and outcomes. *J. Appl. Psychol.* 93, 374–391. doi: 10.1037/0021-9010.93.2.374

Dewall, C. N., Anderson, C. A., and Bushman, B. J. (2011). The general aggression model: theoretical extensions to violence. *Psychol. Violence* 1, 245–258. doi: 10.1037/a0023842

Falla, D., Ortega-Ruiz, R., Runions, K., and Romera, E. M. (2020). Why do victims become perpetrators of peer bullying? Moral disengagement in the cycle of violence. *Youth Society*. doi: 10.1177/0044118X20973702

Garandeau, C. F., Laninga-Wijnen, L., and Salmivalli, C. (2021). Effects of the KiVa anti-bullying program on affective and cognitive empathy in children and adolescents. *J. Clin. Child Adolesc. Psychol.* 1–15. doi: 10.1080/15374416.2020.1846541

Gini, G., Pozzoli, T., and Hymel, S. (2014). Moral disengagement among children and youth: a meta-analytic review of links to aggressive behavior. *Aggress. Behav.* 40, 56–68. doi: 10.1002/ab.21502

Haddock, A. D., and Jimerson, S. R. (2017). An examination of differences in moral disengagement and empathy among bullying participant groups. *J. Relat. Res.* 8, 1–15. doi: 10.1017/jrr.2017.15

Haslam, N., and Loughnan, S. (2014). Dehumanization and infrahumanization. *Annu. Rev. Psychol.* 65, 399–423. doi: 10.1146/annurev-psych-010213-115045

Hayes, A. F. (2013). *Introduction to Mediation, Moderation, and Conditional Process Analysis*. New York, NY: Guilford.

Healey, M. L., and Grossman, M. (2018). Cognitive and affective perspective-taking: evidence for shared and dissociable anatomical substrates. *Front. Neurol.* 9:491. doi: 10.3389/fneur.2018.00491

Heleniak, C., and McLaughlin, K. A. (2020). Social-cognitive mechanisms in the cycle of violence: cognitive and affective theory of mind, and externalizing psychopathology in children and adolescents. *Dev. Psychopathol.* 32, 735–750. doi: 10.1017/S0954579419000725

Herrera-López, M., Gómez-Ortiz, O., Ortega-Ruiz, R., Jolliffe, D., and Romera, E. M. (2017). Suitability of a three-dimensional model to measure empathy and its relationship with social and normative adjustment in Spanish adolescents: a cross-sectional study. *BMJ Open* 7:e015347. doi: 10.1136/bmjopen-2016-015347

Jolliffe, D., and Farrington, D. P. (2006). Development and validation of the basic empathy scale. *J. Adolesc.* 29, 589–611. doi: 10.1016/j.adolescence.2005.08.010

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

Klimecki, O. M. (2019). The role of empathy and compassion in conflict resolution. *Emot. Rev.* 11, 310–325. doi: 10.1177/1754073919838609

Lamm, C., Batson, C. D., and Decety, J. (2007). The neural substrate of human empathy: effects of perspective-taking and cognitive appraisal. *J. Cogn. Neurosci.* 19, 42–58. doi: 10.1162/jocn.2007.19.1.42

MacKinnon, D. P. (2008). *Introduction to Statistical Mediation Analysis*. New York, NY: Taylor & Francis Group.

Mazzone, A., Yanagida, T., Caravita, S. C. S., and Strohmeier, D. (2019). Moral emotions and moral disengagement: Concurrent and longitudinal associations with aggressive behavior among early adolescents. *J. Early Adolesc.* 39, 839–863. doi: 10.1177%2F0272431618791276

Menesini, E., and Salmivalli, C. (2017). Bullying in schools: the state of knowledge and effective interventions. *Psychol. Health Med.* 22, 240–253. doi: 10.1080/13548506.2017.1279740

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

Molchanov, S. V. (2014). Empathy as the factor of moral dilemma solving in adolescence. *Procedia Soc. Behav. Sci.* 146, 89–93. doi: 10.1016/j.sbspro.2014.08.091

Obermann, M. L. (2013). Temporal aspects of moral disengagement in school bullying: crystallization or escalation? *J. Sch. Violence* 12, 193–210. doi: 10.1080/15388220.2013.766133

Ortega-Ruiz, R. (2010). *Agresividad injustificada, bullying y violencia escolar*. Madrid: Alianza Editorial.

Ortega-Ruiz, R. (2020). Educación para el Desarrollo Sostenible: del proyecto cosmopolita a la ciberconvivencia. *Investigación En La Escuela* 100, 11–22. doi: 10.12795/IE.2020.i100.02

Ortega-Ruiz, R., Del Rey, R., and Casas, J. A. (2016). Assessing bullying and cyberbullying: Spanish validation of EBIPQ and ECIPQ. *Psicologia Educativa* 22, 71–79. doi: 10.1016/j.pse.2016.01.004

Perren, S., Gutzwiller-Helfenfinger, E., Malti, T., and Hymel, S. (2012). Moral reasoning and emotion attributions of adolescent bullies, victims, and bully-victims. *Br. J. Dev. Psychol.* 30, 511–530. doi: 10.1111/j.2044-835X.2011.02059.x

Pornari, C. D., and Wood, J. (2010). Peer and cyber aggression in secondary school students: the role of moral disengagement, hostile attribution bias, and outcome expectancies. *Aggress. Behav.* 36, 81–94. doi: 10.1002/ab.20336

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

Preckel, K., Kanske, P., and Singer, T. (2018). On the interaction of social affect and cognition: empathy, compassion and theory of mind. *Curr. Opin. Behav. Sci.* 19, 1–6. doi: 10.1016/j.cobeha.2017.07.010

Rodríguez-Hidalgo, A. J., Calmaestra, J., Casas, J. A., and Ortega-Ruiz, R. (2019). Ethnic-cultural bullying versus personal bullying: specificity and measurement of discriminatory aggression and victimization among adolescents. *Front. Psychol.* 10:46. doi: 10.3389/fpsyg.2019.00046

Romera, E.va M., Ortega-Ruiz, R., Runions, K., and Falla, D. (2020). Moral disengagement strategies in online and offline bullying. *Psychosoc. Interv.* 30, 85–93. doi: 10.5093/pi2020a21

Romera, E. M., Casas, J. A., Gómez-Ortiz, O., and Ortega-Ruiz, R. (2019a). Moral domain as a risk and protective factor against bullying. An integrating perspective review on the complexity of morality. *Aggress. Violent Behav.* 45, 75–82. doi: 10.1016/j.avb.2018.07.005

Romera, E. M., Ortega-Ruiz, R., Rodríguez-Barbero, S., and Falla, D. (2019b). How do you think the victims of bullying feel? A study of moral emotions in primary school. *Front. Psychol.* 10, 1–11. doi: 10.3389/fpsyg.2019.01753

Runions, K. C., and Bak, M. (2015). Online moral disengagement, cyberbullying, and cyber-aggression. *Cyberpsychol. Behav. Soc. Netw.* 18, 400–405. doi: 10.1089/cyber.2014.0670

Runions, K. C., Shaw, T., Bussey, K., Thornberg, R., Salmivalli, C., and Cross, D. S. (2019). Moral disengagement of pure bullies and bully/victims: shared and distinct mechanisms. *J. Youth Adolesc.* 48, 1835–1848. doi: 10.1007/s10964-019-01067-2

Silke, C., Brady, B., Boylan, C., and Dolan, P. (2018). Factors influencing the development of empathy and pro-social behaviour among adolescents: a systematic review. *Child. Youth Serv. Rev.* 94, 421–436. doi: 10.1016/j.childyouth.2018.07.027

Singleton, R., and Straits, B. (2004). *Approaches to Social Research*. Oxford: Oxford University Press.

Smith, P. K., López-Castro, L., Robinson, S., and Görzig, A. (2019). Consistency of gender differences in bullying in cross-cultural surveys. *Aggress. Violent Behav.* 45, 33–40. doi: 10.1016/j.avb.2018.04.006

Stavrinides, P., Georgiou, S., and Theofanous, V. (2010). Bullying and empathy: a short-term longitudinal investigation. *Educ. Psychol.* 30, 793–802. doi: 10.1080/01443410.2010.506004

Tampke, E. C., Fite, P. J., and Cooley, J. L. (2020). Bidirectional associations between affective empathy and proactive and reactive aggression. *Aggress. Behav.* 46, 317–326. doi: 10.1002/ab.21891

Thornberg, R., and Jungert, T. (2014). School bullying and the mechanisms of moral disengagement. *Aggress. Behav.* 40, 99–108. doi: 10.1002/ab.21509

Thornberg, R., Wänström, L., Pozzoli, T., and Hong, J. S. (2019). Moral disengagement and school bullying perpetration in middle childhood: a

short-term longitudinal study in Sweden. *J. Sch. Violence* 18, 585–596. doi: 10.1080/15388220.2019.1636383

Vachon, D. D., and Lynam, D. R. (2013). The (Non)relation between empathy and aggression: surprising results from a meta-analysis five factor machiavellianism inventory view project hunting for the true TriPM view project. *Psychol. Bull.* 140, 751–773. doi: 10.1037/a0035236

van Noorden, T. H. J., Haselager, G. J. T., Cillessen, A. H. N., and Bukowski, W. M. (2015). Empathy and involvement in bullying in children and adolescents: a systematic review. *J. Youth Adolesc.* 44, 637–657. doi: 10.1007/s10964-014-0135-6

Walters, G. D., and Espelage, D. L. (2018). Resurrecting the empathy–bullying relationship with a pro-bullying attitudes mediator: the lazarus effect in mediation research. *J. Abnorm. Child Psychol.* 46, 1229–1239. doi: 10.1007/s10802-017-0355-9

Weisz, E., and Cikara, M. (2020). Strategic regulation of empathy. *Trends Cogn. Sci.* 25, 213–227. doi: 10.1016/j.tics.2020.12.002

Wen, Z., and Fan, X. (2015). Monotonicity of effect sizes: questioning kappa-squared as mediation effect size measure. *Psychol. Methods* 20, 193–203. doi: 10.1037/met0000029

Widom, C. S. (1989). The cycle of violence. *Science* 244, 160–166. doi: 10.1126/science.2704995

Williford, A., Boulton, A. J., Forrest-Bank, S. S., Bender, K. A., Dieterich, W. A., and Jenson, J. M. (2016). The effect of bullying and victimization on cognitive empathy development during the transition to middle school. *Child Youth Care Forum* 45, 525–541. doi: 10.1007/s10566-015-9343-9

Zych, I., and Llorent, V. J. (2019). Affective empathy and moral disengagement related to late adolescent bullying perpetration. *Ethics Behav.* 29, 547–556. doi: 10.1080/10508422.2018.1521282

Zych, I., Ttofi, M. M., and Farrington, D. P. (2019). Empathy and callous–unemotional traits in different bullying roles: a systematic review and meta-analysis. *Trauma Violence Abuse* 20, 3–21. doi: 10.1177/1524838016683456

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Psychology

ORIGINAL RESEARCH
published: 28 September 2021
doi: 10.3389/fpsyg.2021.715279

Young Children's Understanding of Restorative Justice

Zheng Zhou[1]* *and Wan-chi Wong*[2]

[1] The Research Institute of Social Development, Southern University of Finance and Economics in China, Chengdu, China,
[2] Department of Educational Psychology, The Chinese University of Hong Kong, Shatin, Hong Kong, SAR China

The present study investigated how young children understand the sophisticated concept of restorative justice in unintentional moral transgressions. A sex-balanced sample of 5-year-old ($M = 5.67$, $SD = 0.34$, 49.3% girls) and 8-year-old ($M = 7.86$, $SD = 0.29$, 46.0% girls) Chinese children ($N = 193$) participated in the study. In designing the materials, we distilled the multidimensional meanings of restorative justice into two stories, one addressing the theme of property violation and the other physical harm; both stories were set in an animal community. We then engaged the children in joint reading and an interview, during which they showed preference for the given treatments for the transgressor (two restorative treatments vs. two retributive treatments) and ranked two further sets of restorative vs. retributive treatments at the community level. The results indicated that most children favored restorative treatments over retributive treatments for a transgressor, and the 8-year-olds viewed psychological restoration more favorably and behavioral punishment less favorably than the 5-year-olds. The children also tended to endorse restorative treatments at the community level, revealing an understanding of the needs, and obligations of all parties concerned. Notably, more 8- than 5-year-olds showed a consistency in restorative orientation at this level. Interpreting our data through the lens of the Representational Redescription model, we attained a more refined account of young children's levels of understanding regarding restorative justice. These results provide insights for the early cultivation of restorative justice among young children, which is a cornerstone for its successful practice in any society.

Keywords: restorative justice, retributive justice, moral transgression, early moral development, Representational Redescription model, choice-based paradigm

OPEN ACCESS

Edited by:
Miriam Beauchamp,
Université de Montréal, Canada

Reviewed by:
Nara Andrade,
Bahiana School of Medicine and Public Health, Brazil
Tali Gal,
University of Haifa, Israel

***Correspondence:**
Zheng Zhou
zhouzheng@swufe.edu.cn

Specialty section:
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

Received: *26 May 2021*
Accepted: *27 August 2021*
Published: *28 September 2021*

Citation:
Zhou Z and Wong W-c (2021) Young Children's Understanding of Restorative Justice. Front. Psychol. 12:715279. doi: 10.3389/fpsyg.2021.715279

INTRODUCTION

The means to achieve justice in responding to moral transgressions has been a matter of debate for thousands of years. Generally, there are two distinct paradigms of response to wrongdoing: retributive/punitive justice and restorative justice. Retributive justice emphasizes the punishment of wrongdoers and has long been the primary practice in the legal system. However, retributive justice is now criticized for destroying people's social personality (Consedine, 1995), fueling conflict, and deepening harm, especially for relatively minor offenses (Daly and Immarigeon, 1998; Zehr, 2002/2015). In recent decades, restorative justice is regarded as a meaningful solution to the excessive reliance on punishment (Barnett, 1977), and its value has been increasingly recognized globally (see Sullivan and Tifft, 2006; van Wormer and Walker, 2013).

Restorative justice concerns healing the harm caused by wrongdoing and meeting the needs of the involved parties, including the victims, offenders, and communities (Van Ness and Strong, 1997/2015; Daly, 2000; Zehr, 2002/2015). Despite the increasing importance of restorative justice in jurisprudence (see Braithwaite, 2002a,b) and its application extending from the legal system to peacemaking circles, school systems, and family group conferencing (see Strang and Braithwaite, 2001; Sullivan and Tifft, 2006; van Wormer and Walker, 2013), surprisingly little research has focused on how children understand the concept of restorative justice compared to the vibrancy of research on children's understanding of moral concepts such as distributive justice (e.g., Fehr et al., 2008; LoBue et al., 2011; Smith and Warneken, 2016) and procedural justice (e.g., Gold et al., 1984; Shaw and Olson, 2014). The current study addressed the gap in the literature by investigating young children's preference for restorative treatment or punitive treatment in response to varied moral transgressions.

Research on early moral development has flourished in recent decades. A special issue in Human Development featured the state of the art of this research field (Smetana, 2018a). In the introductory essay to this special issue, Smetana (2018b) focused the discussion on the advanced moral capacity of infants and young children shown by various research programs. In the commentary that served as the final paper, Turiel (2018) insightfully remarked that the new findings on moral capacity in the early years are not in line with the influential moral formulation in the field, in which developmental sequences culminate in autonomous morality (postulated by Piaget) and stages of principled morality (postulated by Kohlberg) in a much later period. He further raised the issue of universality vs. cultural specificity in this thriving field of study (Turiel, 2018). This special issue presents readers with diverse theoretical propositions and methods used in the research endeavors of early moral development.

Within the field of early moral development, there is a research focus on the emergence of the sense of justice, particularly young children's understanding of distributive justice. In an experimental study on egalitarianism in young children, Fehr et al. (2008) found that most children aged 7–8 preferred resource allocation that removed advantageous or disadvantageous inequality, while the behaviors of those aged 3–4 were characterized by self-interest. Notably, a research team at the Max-Planck Institute for Evolutionary Anthropology found that children as young as 3.5 and 4.5 years of age were capable of showing an aversion to disadvantageous and advantageous inequities (Ulber et al., 2017). A cross-sectional study of resource allocation conducted among 3-, 5-, and 8-year-olds further revealed that young children's understanding of distributive justice developed from an equality preference to an acceptance of legitimate reasons for unequal allocation of resources, which included the consideration of merits, needs, and agreed-upon rules (Schmidt et al., 2016).

An earlier study on children's understanding of procedural justice was conducted in the context of their reaction to authorities' decisions regarding punishment (Gold et al., 1984). The results indicated that both first and fifth graders were sensitive to the manipulations of procedural justice. Shaw and Olson (2014) conducted a series of experiments on young children's preference for partial vs. impartial procedures in the context of resource distribution. In a sample of 5- to 8-year-olds, the older children demonstrated a stronger aversion to the use of partial procedures, suggesting an increasingly positive valuation of procedural justice in middle childhood. Through the research design of a real-life allocation activity in small groups, Xu and Wong (2014) investigated Chinese children's understanding of procedural justice along the implicit-explicit spectrum. The mastery of procedural justice among 5- to 9-year-old children was found to be relatively low, as reflected by both behavioral performance and verbal explanation at the individual level. Based on the assessment indices for group performance developed from the procedural perspective of Habermas (1983/1990), the 7-year-old group showed a significantly enhanced implicit understanding of procedural justice compared with the 5-year-old group.

Young children's reactions to norm violation and their understanding of punitive justice constitute a further research interest in the field. It was found that children as young as 3 years of age exhibited normative responses such as protest, critique, and teaching when encountering a mistake made by a puppet in conventional games (Rakoczy et al., 2008). Three-year-olds also tended to protest when a third party's property rights were violated (Rossano et al., 2011). Past research has also suggested that young children have a sense of what and why proper treatments should be adopted in responding to wrongdoing. Piaget (1932/1997) found that younger children in the egocentric stage selected more severe punishments than older children. Researchers in developmental and evolutionary psychology have further investigated children's use of punishment in different contexts and its rationale (e.g., Helwig et al., 2001; Salali et al., 2015; Smetana and Ball, 2018). A study conducted by Marshall et al. (2021) further investigated the retributive and consequentialist motives of children in using punishment. However, children's endorsement of restorative treatment, which serves to heal the harm done to the victim and the community, has largely been a neglected research area.

A pioneering study on restorative justice in young children was conducted by Riedl et al. (2015). Through the special design of a turnable table, the researchers applied an innovative action-based paradigm to examine the respective punitive and restorative responses of young German children in an experimental setting. The results of the first experiment indicated that both the 3- and 5-year-old children tended to remove the toy or food away from a puppet, who had grabbed the item away from its owner. A further experiment found that children as young as 3 years of age tended to return the toy or food to the original owner, among other options, when the item was grabbed away by a puppet. In both experiments, children showed the tendency to intervene in a violation where they were a third-party witness just as they did in the case where they were personally affected in a second-party condition. Riedl et al. (2015) interpreted such behavioral responses as reflecting a sense of justice, which might be attributable to an understanding of the harm caused to the victim. There have been emerging interests in

children's punishment and restoration among researchers most recently. Notably, McAuliffe and Dunham (2021) found that 6- to 9-year-olds in an American sample favored punishment over restoration. However, Yang et al. (2021) found that 3- to 6-year-old Chinese children preferred restoration to punishment in both the roles of second-party victims and third-party bystanders, with older children showing a stronger preference for restoration than the younger ones while they were victims of transgressions.

It is worthy to note that the above-mentioned pioneering studies on restorative justice have not yet taken the multidimensional meanings of restorative justice into account. Recognizing that the restorative conception of justice has its roots in both Western and non-Western traditions, some of its proponents have regarded the contemporary discourse and practice as a revival of old traditions (Llewellyn and Howse, 1998; Johnstone, 2001/2011). Eglash is credited with coining the term "restorative justice" in his article entitled "Beyond restitution: Creative restitution" (Eglash, 1977). In Eglash's conceptualization, the concern of creative restitution or restorative justice primarily lies in recognizing the harm caused by the offense and considering the victim's needs (Eglash, 1958, 1977). Zehr, generally regarded as the grandfather of the contemporary restorative justice movement, provided important clarification of the multidimensional meanings of restorative justice. With respect to the restorative process, Zehr (2002/2015) highlights the identification of three major stakeholders, namely, the victim, the offender, and the community. With its aim of righting wrongs and harms, Zehr (2002/2015) proposes three central concepts or pillars of restorative justice. The first pillar constitutes the harms and related needs that involve the victim, the offender, and the community; the second pillar concerns itself with the obligations caused by the harms; the third pillar involves the engagement of all concerned parties in the justice-seeking process. The multidimensional meanings of restorative justice clarified by Zehr have been embraced by subsequent discourses in the field (see Van Ness and Strong, 1997/2015; van Wormer and Walker, 2013).

Although research on restorative justice in developmental psychology is only a recent endeavor, studies related to restorative and retributive treatment have been conducted. A typical psychological restoration to alleviate the harms caused by wrongdoing is an apology. There is evidence that children aged 4–9 years could have a basic understanding of the emotional effects of apology on a transgressor and a victim (Smith et al., 2010). A recent study found that children as young as 4 years of age were more forgiving of a transgressor who had apologized than one who had not, and 5-year-olds were more forgiving of a remorseful wrongdoer than an unremorseful wrongdoer even when the wrongdoer did not explicitly apologize to the victim (Oostenbroek and Vaish, 2019).

In contrast to a verbal apology, actual or behavioral restoration provides the victim with actual compensation for the harm or loss. Transgressors performing actual restitution are believed to express a greater commitment to rectifying their wrongdoings than those offering a mere apology (Carlisle et al., 2012). Drell and Jaswal (2016) found that 6- to 7-year-olds' negative feelings decreased when an offender offered behavioral restitution. A study focusing on college students also suggested that restitution enhances forgiveness (Carlisle et al., 2012).

Unlike restorative justice, retributive justice focuses on punishing an offender (Daly, 2000). It is worth noting that previous research mostly asked a general and abstract question about how much punishment the offender deserved (see Cushman, 2008; Jambon and Smetana, 2014; Smetana and Ball, 2018). Research that differentiates and compares children's judgment on psychological and actual punishment is lacking.

When the community's role is taken into consideration, treatments for a wrongdoer can take additional forms, such as exclusion and education. Although children in one study considered it generally morally wrong to exclude others from the group, exclusion was regarded as more acceptable if it was done for the sake of group norms and group functioning (Killen and de Waal, 2000). In another study, exclusion was also endorsed by children in the context of a member's unequal distribution of resources in a group (Hitti et al., 2014). A line of multidisciplinary research has endeavored to examine, in contrast with exclusion as a punitive response to transgression, the creation of inner and outer spaces for making changes to attain restorative justice (see Gavrielides, 2015).

Given the limited ability of young children to express themselves through language, investigating their understanding of a sophisticated justice concept, such as restorative justice, is a challenging task. However, such analyses are possible by interpreting young children's preferences for restorative treatments vs. retributive treatments through the lens of the Representational Redescription model (RR model). The RR model postulates that the acquisition of concepts and knowledge is achieved at different levels along a spectrum of the implicit-explicit dimension (Karmiloff-Smith, 1992/1999). Specifically, four levels of representation are postulated along the spectrum, namely, Implicit (I), Explicit-1 (E1), Explicit-2 (E2), and Explicit-3 (E3). At the implicit level (Level I), children represent knowledge in procedural form, and interdomain representational links are not yet developed. Thus, the behaviors generated by implicit understanding appear inflexible. At Level E1, representation is more cognitively flexible, and such flexibility could be reflected in children's consistency in performance across domains based on their understanding of a certain concept. However, the representations at Level E1 are not yet consciously accessed until Level E2 is reached. Level E3, which is considered the most explicit level of understanding, is characterized by the ability to verbally articulate representations. This RR model has been applied to explain concept and knowledge acquisition in the domains of linguistics, physics, mathematics, notation, and theory of mind (see Karmiloff-Smith, 1992/1999). In recent years, young children's understanding of distributive justice and procedural justice was further examined through the lens of the RR model (Xu and Wong, 2014, 2016).

The RR model could also serve as a valuable lens in examining young children's understanding of restorative justice. In the case that a child can indicate a preference for a restorative option in a moral transgression scenario, he or she might have understood the concept at an implicit level. If he or she consistently prefers restorative choices across different situations in response to

a moral transgression, this behavior could be interpreted as demonstrating an understanding of restorative justice at least at Level E1, which according to the RR model, is the initial stage of explicit understanding. When a clear explanation of the meaning of restorative justice is given by a child as a justification for his or her choices, this indicates that the child has reached Level E3, a high level of explicit understanding at which representations are conscious and can be verbally articulated. Level E2, which is characterized by representations that are consciously accessible but not yet verbally articulable, is difficult to detect in empirical studies. Hence, Karmiloff-Smith (1992/1999) names E2 and E3 collectively E2/3 in the research context.

In adherence to the RR model, the current study applied a choice-based paradigm and integrated it with consistent analysis of restorative orientation. This is a unique characteristic of this study, and the specific details are subsequently discussed.

While the practice of restorative justice has been gaining momentum in the legal system and social institutions of occidental countries in recent decades (Johnstone, 2001/2011; Sullivan and Tifft, 2006; van Wormer and Walker, 2013), how children understand restorative justice is still unclear. Because the successful practice of restorative justice relies much on the endorsement and engagement of the community as a whole, the cultivation of the values of restorative justice becomes important. In this light, an investigation into how young children make sense of restorative justice serves as a fundamental step by which the preconditions for cultivating the restorative orientation could be uncovered. Though some recent studies have been interested in children's restorative behavior, no research has investigated children's preference for restorative justice at both individual level and community level. The current study originated from a desire to capture young children's understanding of the multidimensional meanings of restorative justice in a possibly comprehensive way, which the abovementioned experimental studies might not achieve. Its uniqueness lies in the design of an interactive story-reading activity, through which the researchers could distill the key components of restorative justice in the created scenarios of moral transgressions and make them comprehensible to young children.

Applying a choice-based paradigm embedded in interactive story reading, the overarching goal of the current study is to investigate young children's understanding of restorative justice in unintentional moral transgression scenarios. The intricate roles of intentionality and harmful consequences in the moral judgment exercised by different age groups have been well-documented in the literature (Piaget, 1932/1997; Yuill and Perner, 1988; Zelazo et al., 1996). In full recognition of the complexity of restorative justice and the different possibilities of studying it from the perspective of developmental psychology, in the current study, we first chose to focus on examining young children's understanding of this multidimensional concept in the context of unintentional actions with harmful consequences.

The first research aim of the study lies in examining young children's preference for restorative treatment vs. retributive treatment with regard to two unintentional moral transgressions, one involving a property violation and the other involving physical harm. Children randomly assigned to read one of the stories were asked to rank their preference regarding the four treatments of the transgressor, two of which were restorative, and two of which were retributive. For each treatment type, one psychological treatment and one actual or behavioral treatment were designed. Based on the results of the pioneering study on children's sense of restorative justice (Riedl et al., 2015), we expected that young children would be in general capable of showing a preference for restorative treatments in unintentional moral transgressions.

The second research aim of the current study is to investigate whether the respective hardship backgrounds of the victim and the transgressor affect young children's treatment preference ranking. To investigate this issue, two rounds of treatment ranking were built into the design of the interactive story reading, one before and one after the introduction of the hardship background of either the victim or the transgressor. The consistency in children's restorative orientation despite the manipulation of the hardship background of the involved parties will be assessed in light of the RR model. Considering Hoffman's (1990) thesis that empathic bias could have an impact on the justice-seeking process, we expected that young children's treatment rankings regarding the transgressor to be affected due to the empathy aroused by the background story of either the victim or the transgressor, which might mean a lower consistency in the restorative orientation. Specifically, we expected that children's empathy for the victim would be conducive to a stronger treatment preference in the punitive orientation, whereas their empathy for the transgressor would be conducive to a stronger treatment preference in the restorative orientation.

The third research aim of the current study further addressed how young children endorse restorative treatment at the community level. We presented children with options involving community-wide engagement for treating the victim, the transgressor, and the community as a whole in the aftermath of a moral transgression. Except for the community-level treatments for the victim, which were all restorative treatments, the design of the community-level treatments for the transgressor and the community were differentiated into restorative treatments and retributive treatments. Furthermore, the consistency in children's restorative orientation with regard to community-level treatments for the transgressor and the community was assessed in light of the RR model. Young children's understanding of restorative justice along the dimension of community engagement is a hitherto unexplored area.

Finally, we were interested in exploring the developmental features of young children's understanding of restorative justice by observing the similarities and differences between the 5- and the 8-year-olds involved in the current study with regard to the above three research aims. Past research in early moral development has shown age differences in the understanding of distributive justice and procedural justice, where the age range of five to eight appears to be a critical period of change. In line with the knowledge that older children in early and middle childhood are more advanced in perspective taking (Selman, 1975; see Elfers et al., 2008), we predicted that 8-year-olds would have a higher level

of understanding of restorative justice at the community level than their 5-year-old counterparts, which would be reflected by the higher consistency in the preference for a restorative orientation.

METHODS

Participants

The participants were from a medium-sized city in Southwest China, including ninety-three 5- to 6-year-old children at a senior grade in a local public kindergarten ($M_{age} = 5.67$, $SD_{age} = 0.34$, 49.3% girls) and 100 second-graders in a local public primary school ($M_{age} = 7.86$, $SD_{age} = 0.29$, 46.0% girls). The kindergarten subsample and primary school subsample are referred to as 5- and 8-year-olds, respectively. To determine the sample size, we conducted power analysis using G*Power 3.1 and found that we would need 188 participants in total in our research design, so that the difference of young children's preference between restorative and retributive treatments could be detected with 80% power (two tails, $\alpha = 0.05$, assuming small to medium effect size equals 0.3). Considering the possibility that some children might withdraw from the study, we recruited 193 children. It turned out that all recruited children agreed to participate at the beginning of the study and all of them completed the research process.

Consent forms were distributed and collected from the parents through an online platform before the implementation of the study. The children recruited were from middle-class families in the urban area. The educational level of the parents should be noted, as 6.2% had a graduate school education, and 45.6% had completed university education. A further 25.4% of the parents had finished vocational school, while 12.4% had completed high school. The percentage of parents with an education level of middle school or below was 5.2. Another 5.2% of parents did not provide information on their educational level.

The story type and background condition were the two between-subject factors. Participants were randomly assigned to four cells within each age group, resulting in the following distribution: 24 5- and 27 8-year-olds in the Stealing Story and transgressor background condition, 24 5- and 24 8-year-olds in the Stealing Story and victim background condition, 22 5- and 24 8-year-olds in the Harm-causing Story and transgressor background condition, and 23 5- and 25 8-year-olds in the Harm-causing Story and victim background condition.

Materials

We undertook an interactive story reading and interview process to assess children's understanding of restorative justice in moral transgressions. The multidimensional meanings and abstract moral rules of restorative justice were embedded in the stories along with a series of questions that could be easily understood and answered by young children.

Story stimuli

The study used two colored picture books of A4 size depicting the following two prototypical moral transgressions that occurred in an animal community: property violation and physical harm. Children were randomly assigned to reading one of the storybooks. The Stealing Story was about a property violation (see online **Supplementary Material**), and the Harm-causing Story was about physical harm (see online **Supplementary Material**). The structure of the interview questions was the same for both stories.

Initial moral judgments

The children's responses to the act's acceptability were scored on a five-point scale with the following specifications: 1 (very wrong), 2 (wrong), 3 (neither wrong nor right), 4 (right), 5 (very right). In a similar vein, their responses to the actor's acceptability were scored on a 5-point scale with the following specification: 1 (very bad), 2 (bad), 3 (neither bad nor good), 4 (good), 5 (very good). In both cases, the scales were adapted from the rating scale used in the study of Smetana and Ball (2018).

Ranking of Restorative and Retributive Treatments

The experimenter informed the children that four treatments were proposed at an animal meeting and that they needed to rank the treatments from the most preferable to the least preferable using four red ballots of decreasing size. The children were further asked to explain their ranking. The first treatment was a psychological restorative solution suggesting that the transgressor should apologize to the victim. The second treatment was an actual restorative solution or an action-oriented solution, suggesting that the transgressor should restore the situation. In the case of the Stealing Story, it involved returning the property; in the case of the Harm causing Story, it involved helping the victim clean the farm. The third treatment was an actual retributive solution of imprisoning the transgressor. The fourth treatment was a psychological retributive solution involving criticizing the transgressor. The four treatments were presented to the children on one page of the storybook with a consistent format across participants. Some necessary explanation of the treatments was made to ensure that the children understood their meanings.

The Hardship Background and Second-Round Judgment

Next, the children were randomly assigned to read the hardship background of the transgressor or the victim. The hardship of the transgressor centered on its fate of being an orphan, and the hardship of the victim was also attributed to its fate of being an orphan, both of which led to the lack of socialization. Then, the experimenter repeated the questions regarding the act's and actor's acceptability and the ranking of the four treatments posed in the initial phase.

Community-Involved Judgments

The experimenter informed the children that the transgression had negative consequences on the community even though the transgressor conducted restorative action and apologized to the victim. Then, the experimenter turned to three sets of treatments that involved the participation of the community in

the justice-seeking process in the future, namely, treatments for the victim, the transgressor, and the community as a whole. The experimenter asked the children to rank the options within the three sets of treatments from the most preferable to the least preferable using plastic stars of decreasing sizes. The options were presented in a consistent format across the participants. Whether they had a further explanation for their ranking was probed.

Set 1: Treatments for the Victim

(a) Help the victim recover from the harm or loss; (b) teach the victim how to protect itself or its property in the future; or (c) make friends with the victim to alleviate its sadness.

Set 2: Treatments for the Transgressor

(a) Expel the transgressor from the community; (b) educate the transgressor; (c) exclude the transgressor from group activities; and (d) let the transgressor serve the community. Options (a) and (c) are retributive treatments, while options (b) and (d) are restorative treatments.

Set 3: Treatments for the Community

(a) Expel the animal who misbehaved; (b) educate the community members (in the case of the Stealing Story, teaching each animal to protect its property; in the case of the Harm-causing Story, teaching each animal to keep the environment clean); (c) exclude the wrongdoer from group activities; and (d) ask every member of the community to shoulder the responsibility of helping the needy community members (with respect to the Stealing Story) or protecting the environment (with respect to the Harm-causing Story). While options (a) and (c) are retributive treatments, options (b) and (d) are restorative treatments.

Procedure

Ethics approval for the current study was obtained from the Survey and Behavioral Research Ethics Committee of the Chinese University of Hong Kong. To test the feasibility of the materials, we conducted a pilot study at a kindergarten and a primary school in a medium-sized city in southwestern China. A total of eleven 5-year-olds (six boys and five girls) and ten 8-year-olds (five boys and five girls) participated in the pilot study. Generally, the instructions, stories, and questions were comprehensible to the children. Minor changes to the wording of the storybooks were made according to the children's feedback.

In the main study, participants were randomly assigned to one of the four conditions (i.e., two different types of stories across two different manipulations of hardship backgrounds). As in the case of the pilot study, an experimenter with doctoral training in developmental psychology conducted the main study. Subsequent to a short warm-up that involved playing Legos with the participants, the experimenter read the storybook and interviewed the children individually in a quiet room at their schools. At the end of each interview, the experimenter thanked the child and gave him or her a set of stationary as a small souvenir.

RESULTS

Children's Ranking of Treatments at the Individual Level

Analyses of the data revealed that most children preferred restorative treatments to retributive treatments for the transgressor in both moral transgression situations. The results showed that 45.1 and 38.9% of the children ranked actual restoration as their first choice and second choice, respectively; 43.5 and 36.3% of the children ranked apology as their first choice and second choice, respectively; 64.8% of the children chose criticism as the third choice; and 85.0% of the children chose imprisonment as the least preferable choice. This pattern of results echoes our prediction related to the first research aim of the current study, which specified that young children were capable of showing preference for restorative treatments in unintentional moral transgressions. No gender difference was found in the ranking [apology: $\chi^2_{(df=3)} = 1.04$, $p = 0.791$; restoration: $\chi^2_{(df=3)} = 0.51$, $p = 0.917$; criticism: $\chi^2_{(df=3)} = 5.13$, $p = 0.163$; jail: $\chi^2_{(df=3)} = 6.48$, $p = 0.091$], and gender was not considered in further analyses. The percentage of each ranking (from 1st to 4th) of the four treatments, further differentiated into the percentage in the two age groups and the two types of stories concerning moral transgression, is reported in **Table 1**.

The Effect of Age and Story Type on Individual-Level Treatment Ranking

As the children's rankings of the four treatments were ordinal data by nature, the following analysis used ordinal regression to test the effect of independent variables on the children's preference for the four treatments. We were interested in how age and story type and their interaction impacted the children's ranking of the treatments. Considering that the use of a continuous variable would be conducive to empty cells and lead to a violation of the parallel assumption that is required for ordinal regression (O'Connell, 2006), we treated children's age as a categorical predictor in the following ordinal regression analysis. Thus, age was dummy coded as 0 (5-year-olds) and 1 (8-year-olds). Story type was also dummy coded as 0 (Stealing Story) and 1 (Harm-causing Story). The interaction product of the two factors was created by multiplying dummy-coded age by dummy-coded story type. In four ordinal regression models, age, story type, and their interaction were predictors, whereas the children's ranking of actual restoration, apology, imprisonment, and criticism were the dependent variables.

The parameter estimates and model-fit outcomes of the ordinal regression models are shown in **Table 2**. Regarding actual restoration, the respective main effects of age and story type on its ranking were not significant ($logit_{\text{age}} = -0.67$, *odds ratio* $= 0.51$, $p = 0.078$, 95% CI[−1.41, 0.08]; $logit_{\text{story}} = -0.49$, *odds ratio* $= 0.61$, $p = 0.210$, 95% CI[−1.26, 0.28]), but the interaction effect of the two variables was significant ($logit_{\text{interation}} = 2.41$, *odds ratio* $= 11.11$, $p < 0.001$, 95% CI[1.27, 3.55]), as was the overall model fit (see **Table 2**). Further analysis revealed that for the Stealing Story, the children's age was not

TABLE 1 | Percentage of each ranking (1st-4th) for the four treatments.

		Percentage of rankings for actual restoration (%)			
		1st	**2nd**	**3rd**	**4th**
Age	5	40.9	39.8	15.1	4.3
	8	49.0	38.0	13.0	0.0
Story	Stealing	35.0	47.0	16.0	2.0
	Harm-causing	55.9	30.1	11.8	2.2
		Percentage of rankings for apology (%)			
		1st	**2nd**	**3rd**	**4th**
Age	5	39.8	30.1	15.1	15.1
	8	47.0	42.0	11.0	0.0
Story	Stealing	58.0	26.0	11.0	5.0
	Harm-causing	28.0	47.3	15.1	9.7
		Percentage of rankings for imprisonment (%)			
		1st	**2nd**	**3rd**	**4th**
Age	5	8.6	5.4	12.9	73.1
	8	0.0	0.0	4.0	96.0
Story	Stealing	3.0	1.0	8.0	88.0
	Harm-causing	5.4	4.3	8.6	81.7
		Percentage of rankings for criticism (%)			
		1st	**2nd**	**3rd**	**4th**
Age	5	10.8	24.7	57.0	7.5
	8	4.0	20.0	72.0	4.0
Story	Stealing	4.0	26.0	65.0	5.0
	Harm-causing	10.8	18.3	64.5	6.5

TABLE 2 | Results of ordinal regression analysis with the ranking of actual restoration, apology, imprisonment, and criticism as respective outcome variables.

Outcome variable	**Predictors**	**Logistic coefficient (*p*-value)**	**Odds ratio[a]**	**Model fitting χ^2 (*df* = 3) (*p*-value)**	**Pseudo R^2 (Nagelkerke)**
Actual restoration	Age	−0.67 (0.078)	0.51	26.66 (<0.001)	0.15
	Story	−0.49 (0.210)	0.61		
	Age*Story	2.41 (<0.001)	11.11		
Apology	Age	1.23 (0.002)	3.41	24.16 (<0.001)	0.13
	Story	−0.58 (0.133)	0.56		
	Age*Story	−0.96 (0.082)	0.38		
Imprisonment	Age	−1.90 (0.019)	0.15	24.64 (<0.001)	0.18
	Story	0.70 (0.136)	2.01		
	Age*Story	−0.62 (0.586)	0.54		
Criticism	Age	−0.02 (0.960)	0.98	3.76 (0.288)	0.02
	Story	0.44 (0.288)	1.56		
	Age*Story	−0.82 (0.165)	0.44		

[a] *5-year-olds and the Stealing Story group served as the reference groups.*

significantly associated with their ranking of actual restoration (*logit* = −0.71, *odds ratio* = 0.49, $p = 0.064$, 95% CI[−1.47, 0.04]). However, for the Harm-causing Story, the children's age was significantly associated with their ranking of actual restoration (*logit* = 1.65, *odds ratio* = 5.22, $p < 0.001$, 95% CI[0.79, 2.51]), which indicated that the older children were more likely to give a higher ranking to actual restoration than younger children in the context of the Harm-causing Story. The children's ranking of apology was significantly predicted by age ($logit_{age} = 1.23$, *odds ratio* = 3.41, $p = 0.002$, 95% CI[0.44, 2.02]), which indicated that older children gave more credit to psychological restoration; but the ranking of apology was not

predicted by story type and the interaction ($logit_{story} = -0.58$, *odds ratio* = 0.56, $p = 0.133$, 95% CI[−0.13, 0.18]; $logit_{interaction} = -0.96$, *odds ratio* = 0.38, $p = 0.082$, 95% CI[−2.05, 0.12]), Regarding the ranking of imprisonment, the effect of age was significant ($logit_{age} = -1.90$, *odds ratio* = 0.15, $p = 0.019$, 95% CI[−3.48, −0.32]), while the effect of story type and the interaction was not significant ($logit_{story} = 0.70$, *odds ratio* = 2.01, $p = 0.136$, 95% CI[−0.22, 1.61]; $logit_{interaction} = -0.62$, *odds ratio* = 0.54, $p = 0.586$, 95% CI[−2.83, 1.60]). These results indicated that older children gave less priority to imprisonment as a way to address moral transgression. Regarding the ranking of criticism, the respective effects of age ($logit_{age} = -0.02$, *odds ratio* = 0.98, $p = 0.960$, 95% CI[−0.83, 0.79]), story type ($logit_{story} = 0.44$, $p = 0.288$, *odds ratio* = 1.56, 95% CI[−0.37, 1.26]), and their interaction ($logit_{interaction} = -0.82$, *odds ratio* = 0.44, $p = 0.165$, 95% CI[−1.99, 0.34]) were all found to be non-significant.

Correlation Analysis of Act and Actor Acceptability and Treatment Ranking

To examine whether the children's ranking of the treatments was correlated with their judgment of act and actor acceptability, we computed the correlation between the children's ratings of acceptability and their treatment ranking using Spearman's rho test (see **Table 3**). The results showed that the worse the act was judged by the children, the stronger their preference was for actual restoration ($\rho = -0.17$, $p = 0.017$) and the lower their preference for apology ($\rho = 0.18$, $p = 0.011$). Regarding actor acceptability, the correlation analysis indicated that the worse the actor was judged by the children, the higher they ranked imprisonment ($\rho = -0.20$, $p = 0.005$) and criticism ($\rho = -0.20$, $p = 0.005$) and the lower they ranked apology ($\rho = 0.23$, $p = 0.001$). These results revealed that the children's preference for retributive treatment was correlated with their judgment of actor acceptability rather than act acceptability.

The Effects of Transgressor and Victim Hardship on Treatment Ranking

To examine whether knowing the hardship background of the transgressor or the victim would lead the children to change their ranking of the four treatments, we compared children's pre- and post-ranking of the treatments using the Wilcoxon signed-rank test. The results of the between-subject comparisons are listed in **Table 4**.

The results indicated that the ranking of apology as a solution was higher after than it was before the children learned about the hardship of the transgressor ($Z = -2.36$, $p = 0.018$). However, after hearing the victim's hardship story, the children ranked apology as a less preferable solution ($Z = -3.07$, $p = 0.002$) and believed that a harsher punishment of putting the transgressor into prison was a more appropriate solution ($Z = -3.18$, $p = 0.001$). Other comparisons between the pre- and post-rankings were not significant (see **Table 4**). These findings suggested that the hardship of the involved parties had an impact on the priority given to psychological restoration and actual punishment by the children: In the case when they were told that the transgressor had experienced hardship, they were more likely to support apology as a solution; if they were told that the victim had experienced hardship, they were more likely to support putting the transgressor into prison and less likely to support accepting an apology from the transgressor. These results echo our prediction related to the second research aim of the current study, which specified that children's consistency in their restorative vs. punitive orientation would be affected by their empathy for the victim or the transgressor.

Children's Preference for Community-Involved Treatments

In the last part of the storybook, we presented children with a series of solutions at the community level focusing separately on the victim, the transgressor, and the community. We were interested in examining the ways in which the children's rankings at the community level were associated with age, story type, and the background of the involved parties. Along with listing the percentages associated with the children's rankings of the solutions that addressed the victim, transgressor, and community, **Table 5** shows the results of ordinal regression. Each solution's ranking appears as the outcome variable, and age (0 = 5-year-olds, 1 = 8-year-olds), story type (0 = Stealing Story, 1 = Harm-causing Story) and hardship background (0 = transgressor background, 1 = victim background) are the predictors.

Table 5 indicates that compared to the children who knew the hardship story of the transgressor, those who knew the hardship background of the victim were more likely to prioritize helping the victim recover from the harm or restore the loss (*logit* = 0.76, *odds ratio* = 2.13, $p = 0.006$, 95% CI[0.21, 1.30]) but less likely to teach the victim to protect itself or its property (*logit* = −0.82, *odds ratio* = 0.44, $p = 0.003$, 95% CI[−1.36, −0.28]).

With regard to the treatment of the transgressor, the older children were more likely to favor educating the transgressor (*logit* = 1.04, *odds ratio* = 2.83, $p = 0.002$, 95% CI[0.39, 1.69]) when they were informed of its hardship

TABLE 3 | Correlations of acceptability and preference for the four treatments.

	Actual restoration	Apology	Imprisonment	Criticism
Act acceptability	−0.17 ($p = 0.017$)	0.18 ($p = 0.011$)	−0.07 ($p = 0.324$)	0.07 ($p = 0.370$)
Actor acceptability	0.08 ($p = 0.290$)	0.23 ($p = 0.001$)	−0.20 ($p = 0.005$)	−0.20 ($p = 0.005$)

Correlations were computed by Spearman's rho value. Act acceptability and actor acceptability were scored on 5-point scales with 5 = very right and very good, respectively; the four treatments were rated on a scale ranging from 1 = the least preferable choice to 4 = the most preferable choice.

TABLE 4 | Comparisons of children's pre- and post-preference for moral treatments.

	Transgressor hardship				Victim hardship			
	Negative rank[a]	Positive rank[b]	Ties	*Z* (*p*-value)	Negative rank	Positive rank	Ties	*Z* (*p*-value)
Res_{post}-Res_{pre}[c]	35	22	39	−1.71 (0.088)	23	27	47	−0.26 (0.797)
Apo_{post}-Apo_{pre}[c]	18	38	40	−2.36 (0.018)	37	17	43	−3.07 (0.002)
Imp_{post}-Imp_{pre}[c]	10	7	79	−0.03 (0.980)	2	15	80	−3.18 (0.001)
Cri_{post}-Cri_{pre}[c]	24	18	54	−1.23 (0.218)	15	21	61	−1.03 (0.304)

[a] *The number of pairs for which the post-ranking was lower than the pre-ranking.*
[b] *The number of pairs for which the post-ranking was higher than the pre-ranking.*
[c] *Res_{pre} and Res_{post} refer to the children's pre- and post-ranking of actual restoration, respectively; Apo_{pre} and Apo_{post} refer to the children's pre- and post-ranking of apology, respectively; Imp_{pre} and Imp_{post} refer to the children's pre- and post-ranking of imprisonment, respectively; and Cri_{pre} and Cri_{post} refer to the children's pre- and post-ranking of criticism, respectively.*

TABLE 5 | Percentage of ranking for community-level solutions and the results of ordinal regressions predicting rankings from age, story type, and hardship background.

		Percentage of ranking (%)			Logits[a] (*p* value)		
		1st	2nd	3rd	Age	Story	Background
Treatments for the victim	Helping	49.2	24.4	25.9	−0.45 (0.108)	−0.28 (0.308)	0.76 (0.006)
	Teaching	25.4	44.0	30.1	0.07 (0.800)	−0.07 (0.808)	−0.82 (0.003)
	Comforting	25.4	31.1	43.0	0.27 (0.315)	−0.20 (0.448)	−0.05 (0.842)
Treatments for the transgressor	Expelling	3.6	4.1	6.2	– –	0.47 (0.295)	0.37 (0.412)
	Excluding	1.6	6.2	69.9	0.37 (0.237)	0.50 (0.116)	0.80 (0.013)
	Educating	70.5	24.9	2.1	1.04 (0.002)	−0.28 (0.392)	−0.73 (0.026)
	Serving	24.4	62.7	8.3	– –	−0.12 (0.677)	0.23 (0.430)
Treatments for the community	Expelling	8.8	8.8	6.2	−2.60 (< 0.001)	−0.34 (0.358)	0.83 (0.026)
	Excluding	3.6	13.5	62.2	−0.53 (0.069)	0.13 (0.645)	0.31 (0.282)
	Educating	39.9	44.6	6.7	0.42 (0.123)	0.68 (0.014)	0.24 (0.383)
	Responsibility-taking	48.7	31.6	14.0	1.48 (<0.001)	−0.67 (0.018)	−0.48 (0.086)

For the logits of age, story type, and background, the reference groups are 5-year-olds, the Stealing Story group, and the transgressor background group, respectively.

than the group that received the information about the hardship of the victim. In addition, compared to the children who knew the hardship of the transgressor, the children who knew the hardship of the victim were more inclined to exclude the transgressor from group activities (*logit* = 0.80, *odds ratio* = 2.22, $p = 0.013$, 95% CI[0.17, 1.43]) and were less likely to give support to educating the transgressor (*logit* = −0.73, *odds ratio* = 0.48, $p = 0.026$, 95% CI[−1.38, −0.09]).

Concerning the solution for the whole community, the results showed that the older children were more likely to give priority to the solution requiring the community members to share the responsibility (*logit* = 1.48, *odds ratio* = 4.39, $p < 0.001$, 95% CI[0.91, 2.05]) and were less willing to support the solution that suggested expelling the transgressor from the community (*logit* = −2.60, *odds ratio* = 0.07, $p < 0.001$, 95% CI[−3.54, −1.66]). Compared to the Stealing Story group, the Harm-causing Story group was more likely to favor the solution that suggested educating community members to prevent future transgression (*logit* = 0.68, *odds ratio* = 1.97, $p = 0.014$, 95% CI[0.14, 1.22]) but less likely to ask each community member to shoulder the responsibility (*logit* = −0.67, *odds ratio* = 0.51, $p = 0.018$, 95% CI[−1.23, −0.12]). In addition, compared with the children who knew the hardship background of the transgressor, the children

who knew the hardship background of the victim were more likely to support the expulsion of future transgressors from the community (*logit* = 0.83, *odds ratio* = 2.30, $p = 0.026$, 95% CI[0.10, 1.57]). The pattern of results reported in this subsection addressed the third research aim of the current study in uncovering young children's understanding of restorative justice at the community level, which fills a knowledge gap in the field.

The Consistency of the Children's Restorative Orientation

In our current study, the consistency of the children's restorative orientation refers to their consistent preference for restorative treatments across different situations within each story of moral transgression. First, we were interested in the consistency of the children's preference for restorative justice before and after learning about the hardship background of the victim or the transgressor at the individual level. The treatments were ranked from one to four, and there were two restorative treatments and two retributive treatments. In the case that a child's pre- and post-rankings for the two restorative treatments were always at or above the second-ranking (i.e., first or second), he or she was considered to be consistently supportive of restorative treatments.

Regarding the community-level treatments, the treatments for the transgressor and the community each included two restorative treatments and two retributive treatments. The children who top-ranked the two restorative treatments for the transgressor and top-ranked the two restorative treatments for the community were regarded as having a consistent restorative orientation. **Table 6** shows the number of participants with a consistent or inconsistent restorative orientation at both the individual and community levels. Those participants who showed a consistent retributive orientation were excluded from the analysis.

As shown in **Table 6,** there was no age difference in the number of children with consistent and inconsistent restorative orientations at the individual level [$\chi^2_{(df=1)} = 0.39$, $p = 0.531$]. However, at the community level, there were significantly more 8-year-olds who chose the restorative treatments consistently than 5-year-olds [$\chi^2_{(df=1)} = 28.04$, $p < 0.001$]. When combining the children's choices at the individual level and the community level, significantly more 8-year-olds had a consistent restorative orientation than 5-year-olds [$\chi^2_{(df=1)} = 6.19$, $p = 0.013$]. In other words, when considering treatments that involved the transgressor and the victim only at the individual level, there was a similar proportion of 5- and 8-year-olds who held a consistent restorative orientation before and after learning about the hardship story of one of the involved parties. In contrast, when considering treatments involving the engagement of the whole community, more 8- than 5-year-olds held a consistent restorative orientation. This finding echoes our prediction made on the age differences in the understanding of restorative justice, based on the development of perspective-taking during early and middle childhood.

DISCUSSION

The current study systematically investigated children's understanding of the multidimensional meanings of restorative justice. We embedded a choice-based paradigm into newly developed story materials to examine children's preference for restorative or retributive treatments. The results revealed that most 5- and 8-year-olds demonstrated a sense of restorative justice in response to unintentional moral transgressions. Specifically, children preferred restorative treatments to retributive treatments in both property violation and harm-causing scenarios, and their endorsement of apology, a type of psychological restoration, developed with age. We also observed that the children's preference for restorative treatments was associated with their judgment of the acceptability of the moral transgression and could be influenced by knowledge of the hardship background of the victim or transgressor. In addition, the children also extended their preference for restorative treatment to community-level engagement.

Knowledge Advancement Centered on the Three Critical Differentiations of the Treatments for Moral Transgressions

A central aim of the current study was to examine the endorsement of restorative treatments in young children. The study addressed this aim by applying three critical treatment differentiations in our materials to gain deeper insights into children's understanding of restorative justice. The first differentiation involves children's preference for restorative options vs. retributive options. Restorative treatments, including both apology and actual restitution, are important for maintaining social harmony and repairing relationships and trust (Drell and Jaswal, 2016; Ma et al., 2018). Our findings indicated that most children in our sample, including those as young as 5 years of age, showed a preference for both forms of restoration. Notably, these findings, elicited by a choice-based paradigm, concur with the restoration orientation found in other pioneering studies using an action-based paradigm (Riedl et al., 2015; Yang et al., 2021). This implied that a sense of restorative justice could be fostered in young children's minds.

The second differentiation that we made was between the psychological treatments and actual or behavioral treatments within the restorative and retributive orientations. Our findings showed that most children preferred actual restoration as the first choice and psychological restoration as the second choice. The more severe the transgression was judged to be, the more endorsement was given to actual restoration, which required the transgressor to provide actual compensation to the victim. Previous studies found no age differences in children's judgment regarding whether an offender was deserving of punishment for certain moral transgressions (Smetana, 1981). However, when we differentiated punishment into psychological and actual punishment, we found that the endorsement of actual punishment decreased with age, although the endorsement of psychological punishment did not change. Given that the moral transgressions described in both stories were unintentional, the

TABLE 6 | Number of children with a consistent or an inconsistent restorative orientation.

	Individual level		Community level		Both levels	
Age	**Incons.**	**Cons.**	**Incons.**	**Cons.**	**Incons.**	**Cons.**
5	45	44	47	43	66	24
8	46	54	16	84	56	44
χ^2 ($df = 1$)	0.39 ($p = 0.315$)		28.04 ($p < 0.001$)		6.19 ($p = 0.013$)	

Incons., inconsistent; cons., consistent. Children who had a consistent retributive orientation were excluded from the analyses (four 5-year-olds at the individual level, three 5-year-olds at the community level, and three 5-year-olds at both levels).

decreased endorsement of punishment in older children might be attributable to their developing theory of mind, which enables them to better understand other's mental states, intentions, emotions, and thoughts (see Flavell, 1999). The results also echo Piaget's observation that older children have a reduced preference for punishment as they develop from the egocentric stage to the cooperation stage (Piaget, 1932/1997).

The third differentiation was between restorative treatments at the individual level and the community level. The individual-level treatments focused on only the two parties—the transgressor and the victim. Previous studies on punishment and apology have seldom taken the well-being of the community into consideration (e.g., Cushman, 2008; Oostenbroek and Vaish, 2019). Community-level treatments consider the community as a whole in addressing wrongdoing. Piaget (1932/1997) posited that exclusion causes the offender to feel isolated and thus motivates him or her to return to normal social relations. Previous studies have found that children accepted the social exclusion of in-group members out of consideration for group norms and group functioning (Killen et al., 2013; Hitti et al., 2014). The current study suggested that even when the community is harmed by a transgression, many young children are reluctant to exclude an unintentional transgressor and overwhelmingly endorse restorative options, including educating or responsibility taking. The findings also implied that young children opt for social support in addressing wrongdoings instead of using harsh measures of societal deterrence.

Consistency of the Restorative Orientation

Through the lens of the RR model, the consistency analysis in the preference for restorative treatments revealed that a substantial portion of 5- and 8-year-olds had reached the initial level of explicit understanding of restorative justice. Children who indicated certain restorative preferences but did not demonstrate consistency might be perceived as having an understanding of restorative justice at the implicit level. Although the children were asked to justify their treatment preference for moral transgressions in the study, articulated explanations of restorative justice were not witnessed. The most typical answer for a justification of their treatment preference was "I don't know." In the remaining cases, the children somewhat restated their preferences in their own words or talked about unrelated issues. Thus, regarding the participants' understanding of restorative justice, no evidence was found to support their general attainment of Level E2/3, the high level of explicit understanding postulated in the RR model, at which representations are conscious and verbally articulable. It was evident that the older children showed higher consistency in their restorative orientation at the community level, reflecting a more comprehensive mastery of the multidimensional concept of restorative justice. No significant difference was found between the 5- and the 8-year-olds in their restorative orientation consistency at the individual level. Consistency in choosing restorative treatments for the transgressor in this case required a mature and consolidated mastery of the concept of restorative justice despite hearing the hardship story of the victim. This might explain why even the 8-year-olds did not show higher consistency than their 5-year-old counterparts with respect to the restorative orientation. As explicated by Karmiloff-Smith (1992/1999), the RR model postulates the development of different levels of conceptual understanding along the implicit-explicit continuum and is not an aged-based theoretical framework. Nonetheless, with respect to young children's understanding of a certain concept, such as restorative justice, along this continuum, the RR model can also serve as a lens to discern the developmental characteristics of different age groups.

Limitations and Future Directions

Despite the interesting findings of the current study, it has several limitations. First, the participants in the current study were all from families of middle socioeconomic status with relatively good educational backgrounds living in a medium-sized city in Southwest China. Thus, it might not be possible to generalize the results to explain the restorative orientation of children from other cultural and socioeconomic backgrounds. In the future, examining the preference for restorative treatments of children from different backgrounds in cross-cultural settings would be worthy of our endeavors. Cultural differences in moral development during early and middle childhood have been found in empirical studies (see, for instance, Lau et al., 2012; Chiu Loke et al., 2014; Paulus, 2015). As the current study on young children's understanding of restorative justice is pioneering and involves only a Chinese sample, we are not yet in a position to contribute to the discourse on the issue of universality vs. cultural specificity in early moral development. Considering that the understanding and practice of restorative justice is a global concern, cross-cultural research in this topic area constitutes a valuable future direction. Second, we acknowledge that both of the moral transgression scenarios designed in the current study were unintentional by nature. Future research can explore how children endorse restorative treatments in the context of intentional transgression and how children's restorative

preferences vary across different levels of transgression severity. Preference in the practice of restorative justice vs. retributive justice is a complex jurisprudence issue that is value-laden and depends on the nature of the transgression and its contexts. In the current study, we found that children's preference for retributive treatment was negatively and significantly correlated with actor acceptability. Investigation into children's restorative vs. punitive orientation in the context of intentional moral transgressions would be an interesting focus for future studies. Finally, in employing a cross-sectional design in the current study, we were not able to rule out the possibility that the differences between the two age groups were due to the birth cohort effect induced by shared temporal experiences. In the future, the application of a rigorous longitudinal design is recommended to verify the developmental characteristics of the understanding of restorative justice.

CONCLUSIONS

Notwithstanding the limitations of the current study, it has important theoretical, methodological, and practical implications. Interpreting our data through the lens of the RR model, we attained a more refined perception of young children's levels of understanding regarding restorative justice. The study also serves to extend the RR model's scope of application from cognitive development (e.g., mathematical cognition, language learning) to social cognition. On the methodological plane, the interactive storybooks on moral transgressions, in their current or further modified form, can serve as a new research tool for eliciting children's understanding of restorative justice vs. retributive justice. On the practical plane, the findings of the current study can provide insights for parents and teachers with regard to the cultivation of restorative justice in young children. Such an early cultivation is the most desirable way to facilitate a good understanding of the moral concept among citizens, representing the cornerstone of the successful practice of restorative justice vs. retributive justice in any society. It is encouraging to find that children as young as 5 years of age in this sample showed an understanding of restorative justice, admittedly a sophisticated concept. The finding that young children have different levels of understanding of this moral concept calls for further deliberation of how different educational means can be applied to children to enhance their acquisition and mastery of this concept. It is recommended to promote interactive reading activities related to restorative justice in family and kindergarten settings. Small group discussions regarding a restorative orientation vs. a retributive orientation can be arranged for primary school students using moral transgression scenarios. To facilitate young children's understanding of restorative justice at the implicit and explicit levels, intervention programs can be further designed to examine the role of exposure to stories of restorative justice and the role of various types of dialogues in the restorative orientation.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because authors have not asked for parents' permission to make public the interview record of the young children, we are not in the position to make such raw data available. Requests to access the datasets should be directed to Zheng Zhou, zhouzheng@swufe.edu.cn.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Survey and Behavioral Research Ethics Committee of the Chinese University of Hong Kong. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

Both ZZ and WW contributed to the conception and design of the study. ZZ conducted the experiment and analyzed the data. Both authors engaged in the interpretation and discussion of the findings. WW reviewed and revised the first draft of the manuscript written by ZZ. Both authors contributed to the final revision and approved the submitted version subsequent to a close reading.

FUNDING

The data collection of the present study was supported the Department of Educational Psychology at the Chinese University of Hong Kong. The publication fee of this manuscript is supported by The Research Funds for the Young Scholars, Southwestern University of Finance and Economics (Grant Number: 230600001002020019).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.715279/full#supplementary-material

REFERENCES

Barnett, R. E. (1977). Restitution: a new paradigm of criminal justice. *Ethics* 87, 279–301.

Braithwaite, J. (2002a). "In search of restorative jurisprudence," in *Restorative Justice and the Law,* ed L. Walgrave (Devon: Willan Publishing), 150–167.

Braithwaite, J. (2002b). Restorative justice and therapeutic jurisprudence. *Crim. Law Bull.* 38, 244–262.

Carlisle, R. D., Tsang, J. A., Ahmad, N. Y., Worthington, E. L., van Witvliet, C. O., and Wade, N. (2012). Do actions speak louder than words? Differential effects of apology and restitution on behavioral and self-report measures of forgiveness. *J. Posit. Psychol.* 7, 294–305. doi: 10.1080/17439760.2012.690444

Chiu Loke, I., Heyman, G. D., Itakura, S., and Toriyama, R. (2014). Japanese and American children's moral evaluations of reporting on transgressions. *Dev. Psychol.* 50, 1520–1531. doi: 10.1037/a0035993

Consedine, J. (1995). Restorative *Justice: Healing the Effects of Crime*. Boston, MA: Ploughshares Publications.

Cushman, F. (2008). Crime and punishment: distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition* 108, 353–380. doi: 10.1016/j.cognition.2008.03.006

Daly, K. (2000). "Revisiting the relationship between retributive and restorative justice," in *Restorative Justice: Philosophy to Practice,* eds H. Strang and J. Braithwaite (Hanover, NH: Dartmouth), 33–54.

Daly, K., and Immarigeon, R. (1998). The past, present, and future of restorative justice: some critical reflections. *Contemp. Justice Rev.*1, 21–45.

Drell, M. B., and Jaswal, V. K. (2016). Making amends: children's expectations about and responses to apologies. *Soc. Dev.* 25, 742–758. doi: 10.1111/sode.12168

Eglash, A. (1958). Creative restitution. A broader meaning for an old term. *J Crim. Law Criminol. Police Sci.* 48, 619–622.

Eglash, A. (1977). "Beyond restitution: creative restitution," in *Restitution in Critical Justice,* eds J. Hudson and B. Galaway (Lexington, KY: D. C. Health and Company), 91–99.

Elfers, T., Martin, J., and Sokol, B. (2008). "Perspective taking: a review of research and theory extending Selman's developmental model of perspective taking," in *Advances in Psychology Research, Vol. 54, Advances in Psychology Research*, ed A. M. Columbus (Hauppauge, NY: Nova Science Publishers), 229–262.

Fehr, E., Bernhard, H., and Rockenbach, B. (2008). Egalitarianism in young children. *Nature* 454, 1079–1083. doi: 10.1038/nature07155

Flavell, J. H. (1999). Cognitive development: children's knowledge about the mind. *Annu. Rev. Psychol.* 50, 21–45. doi: 10.1146/annurev.psych.50.1.21

Gavrielides, T. (2015). *The Psychology of Restorative Justice.* London: Routledge.

Gold, L. J., Darley, J. M., Hilton, J. L., and Zanna, M. P. (1984). Children's perceptions of procedural justice. *Child Dev.* 55, 1752–1759. doi: 10.2307/1129922

Habermas, J. (1983/1990). *Moral Consciousness and Communicative Action.* Transl by C. Lenhardt and S.W. Nicholsen. Cambridge: MIT Press. Original Work published 1983.

Helwig, C. C., Zelazo, P. D., and Wilson, M. (2001). Children's judgments of psychological harm in normal and noncanonical situations. *Child Dev.* 72, 66–81. doi: 10.2307/1132472

Hitti, A., Mulvey, K. L., Rutland, A., Abrams, D., and Killen, M. (2014). When is it okay to exclude a member of the ingroup? Children's and adolescents' social reasoning. *Soc. Dev.* 23, 451–469. doi: 10.1111/sode.12047

Hoffman, M. L. (1990). Empathy and justice motivation. *Motiv. Emot.* 14, 151–172.

Jambon, M., and Smetana, J. G. (2014). Moral complexity in middle childhood: children's evaluations of necessary harm. *Dev. Psychol.* 50, 22–33. doi: 10.1037/a0032992

Johnstone, G. (2001/2011). *Restorative Justice. Ideas, Values, Debates*. Denvon: Routledge. Original work published 2001).

Karmiloff-Smith, A. (1992/1999). Beyond modularity: a developmental perspective on cognitive science. Cambridge: MIT Press (Original work published 1992)

Killen, M., and de Waal, F. B. M. (2000). "The evolution and development of morality," in *Natural Conflict Resolution*, eds. F. Aureli and F. B. M. de Waal (Oakland, CA: University of California Press), 352–372.

Killen, M., Mulvey, K. L., and Hitti, A. (2013). Social exclusion in childhood: a developmental intergroup perspective. *Child Dev.* 84, 772–790. doi: 10.1111/cdev.12012

Lau, Y. L., Cameron, C. A., Chieh, K. M., O'Leary, F.u, G., and Lee, K (2012). Cultural differences in moral justifications enhance understanding of Chinese and Canadian children's moral decisions. *J. Cross Cult. Psychol.* 44, 461–477. doi: 10.1177/0022022112453315

Llewellyn, J. J., and Howse, R. (1998). *Restorative Justice – A Conceptual Framework.* Law Commission of Canada.

LoBue, V., Nishida, T., Chiong, C., DeLoache, J., and Haidt, J. (2011). When getting something good is bad: even three-year-olds react to inequality. *Soc. Dev.* 20, 154–170. doi: 10.1111/j.1467-9507.2009.00560.x

Ma, F., Wylie, B. E., Luo, X., He, Z., Xu, F., and Evans, A. D. (2018). Apologies repair children's trust: the mediating role of emotions. *J. Exp. Child Psychol.* 176, 1–12. doi: 10.1016/j.jecp.2018.05.008

Marshall, J., Yudkin, D. A., and Crockett, M. J. (2021). Children punish third parties to satisfy both consequentialist and retributive motives. *Nat. Hum. Behav.* 5, 361–368. doi: 10.1038/s41562-020-00975-9

McAuliffe, K., and Dunham, Y. (2021). Children favor punishment over restoration. *Dev. Sci.* 2021, e13093. doi: 10.1111/desc.13093

O'Connell, A. A. (2006). *Logistic Regression Models for Ordinal Response Variables*. Thousand Oaks, CA: Sage Publications.

Oostenbroek, J., and Vaish, A. (2019). The emergence of forgiveness in young children. *Child Dev.* 90, 1969–1986. doi: 10.1111/cdev.13069

Paulus, M. (2015). Children's inequity aversion depends on culture: a cross-cultural comparison. *J. Exp. Child Psychol.* 132, 240–246. doi: 10.1016/j.jecp.2014.12.007

Piaget, J. (1932/1997). *The Moral Judgment of the Child*. New York, NY: Free Press. Original work published 1932.

Rakoczy, H., Warneken, F., and Tomasello, M. (2008). The sources of normativity: young children's awareness of the normative structure of games. *Dev. Psychol.* 44, 875–881. doi: 10.1037/0012-1649.44.3.875

Riedl, K., Jensen, K., Call, J., and Tomasello, M. (2015). Restorative justice in children. *Curr. Biol.* 25, 1731–1735. doi: 10.1016/j.cub.2015.05.014

Rossano, F., Rakoczy, H., and Tomasello, M. (2011). Young children's understanding of violations of property rights. *Cognition* 121, 219–227. doi: 10.1016/j.cognition.2011.06.007

Salali, G. D., Juda, M., and Henrich, J. (2015). Transmission and development of costly punishment in children. *Evol. Hum. Behav.* 36, 86–94. doi: 10.1016/j.evolhumbehav.2014.09.004

Schmidt, M. F. H., Svetlova, M., Johe, J., and Tomasello, M. (2016). Children's developing understanding of legitimate reasons for allocating resources unequally. *Cogn. Dev.* 37, 42–52. doi: 10.1016/j.cogdev.2015.11.001

Selman, R. L. (1975). Level of social perspective taking and the development of empathy in children: Speculations from a social-cognitive viewpoint. *J. Moral Educ.* 5, 35–43. doi: 10.1080/0305724750050105

Shaw, A., and Olson, K. (2014). Fairness as partiality aversion: the development of procedural justice. *J. Exp. Child Psychol.* 119, 40–53. doi: 10.1016/j.jecp.2013.10.007

Smetana, J. G. (1981). Preschool children's conceptions of moral and social rules. *Child Dev.* 52, 1333–1336. doi: 10.1111/j.1467-8624.1981.tb03187.x

Smetana, J. G. (ed.). (2018a). Early moral development [Special issue]. *Hum. Dev.* 61, 205–310.

Smetana, J. G. (2018b). Becoming moral: introduction to the special issue on early moral development. *Hum. Dev.* 61, 209–213. doi: 10.1159/000492799

Smetana, J. G., and Ball, C. L. (2018). Young children's moral judgments, justifications, and emotion attributions in peer relationship contexts. *Child Dev.* 89, 2245–2263. doi: 10.1111/cdev.12846

Smith, C. E., Chen, D., and Harris, P. L. (2010). When the happy victimizer says sorry: children's understanding of apology and emotion. *Br. J. Dev. Psychol.* 28, 727–746. doi: 10.1348/026151009X475343

Smith, C. E., and Warneken, F. (2016). Children's reasoning about distributive and retributive justice across development. *Dev. Psychol.* 52, 613–628. doi: 10.1037/a0040069

Strang, H., and Braithwaite, J. (eds.). (2001). *Restorative Justice and Civil Society*. Cambridge: Cambridge University Press.

Sullivan, D., and Tifft, L. (eds.). (2006). Handbook of *Restorative Justice*: *A Global Perspective*. New York, NY: Routledge.

Turiel,. E. (2018). Moral development in the early years: when and How. *Hum. Dev.* 61, 297–308. doi: 10.1159/000492805

Ulber, J., Hamann, K., and Tomasello, M. (2017). Young children, but not chimpanzees, are averse to disadvantageous and advantageous inequities. *J. Exp. Child Psychol.* 155, 48–66. doi: 10.1016/j.jecp.2016.10.013

Van Ness, D. W., and Strong, K. H. (1997/2015). *Restoring Justice: An Introduction to Restorative Justice*. New York, NY: Routledge. Original work published 1997.

van Wormer, K. S., and Walker, L. (eds.). (2013). *Restorative Justice Today: Practical applications*. Sage Thousand Oaks, CA: Publications.

Xu, H., and Wong, W. C. (2014). Children's mastery of the concepts of procedural and distributive justice reflected in a distribution activity (in Chinese). *Adv. Psychol.* 4, 239–251. doi: 10.12677/AP.2014.42036

Xu, H., and Wong, W. C. (2016). Understanding children's conceptual development of distributive justice along the implicit-explicit dimension (in Chinese). *Bull. Educ. Psychol.* 48, 37–51. doi: 10.6251/BEP.20150708

Yang, X., Wu, Z., and Dunham, Y. (2021). Children's restorative justice in an intergroup context. *Soc. Dev.* 30, 663–83. doi: 10.1111/sode.12508

Yuill, N., and Perner, J. (1988). Intentionality and knowledge in children's judgments of actor's responsibility and recipient's emotional reaction. *Dev. Psychol.* 24, 358–365. doi: 10.1037/0012-1649.24.3.358

Zehr, H. (2002/2015). *The Little Book of Restorative Justice*. New York, NY: Skyhorse Publishing. Original work published 2002.

Zelazo, P. D., Helwig, C. C., and Lau, A. (1996). Intention, act, and outcome in behavioral prediction and moral judgement. *Child Dev.* 67, 2478–2492. doi: 10.1111/j.1467-8624.1996.tb01869.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 08 November 2021
doi: 10.3389/fpsyg.2021.742199

Sociomoral Temperament: A Mediator Between Wellbeing and Social Outcomes in Young Children

Darcia Narvaez[1*†], *Tracy Gleason*[2†], *Mary Tarsha*[1†], *Ryan Woodbury*[1†], *Ying Cheng*[1†] *and Lijuan Wang*[1†]

[1] Department of Psychology, University of Notre Dame, Notre Dame, IN, United States, [2] Department of Psychology, Wellesley College, Wellesley, MA, United States

Social outcomes, such as empathy, conscience, and behavioral self-regulation, might require a baseline of psychological wellbeing. According to Triune Ethics Metatheory (TEM), early experience influences the neuropsychology underlying a child's orientation toward the social and moral world. Theoretically, a child's wellbeing, fostered through early caregiving, promotes sociomoral temperaments that correspond to the child's experience, such as social approach or withdrawal in face-to-face situations. These temperaments may represent an individual's default sociomoral perspective on the world. We hypothesized that sociomoral temperament emerges as a function of wellbeing and would be related to social outcomes measured by moral socialization and self-regulation. Further, we hypothesized that sociomoral temperament would mediate the relationship between wellbeing and social outcomes. To investigate, we collected items reflective of sociomoral temperament, asking mothers from two countries (USA: $n = 525$; China: $n = 379$) to report on their 3- to 5-year-old children. They also reported on their child's wellbeing (anxiety, depression, happiness) and social outcomes, including moral socialization (concern after wrong doing, internalized conduct and empathy) and behavioral self-regulation (inhibitory control and misbehavior). As expected, correlations identified connections between wellbeing, sociomoral temperament, and social outcomes. Mediation analyses demonstrated that sociomoral temperament mediated relations between wellbeing and social outcomes in both samples, though in slightly different patterns. Fostering early wellbeing may influence social outcomes through a child's developing sociomoral temperament.

Keywords: sociomoral, wellbeing, temperament, self-regulation, social behavior, child development, socialization, triune ethics

OPEN ACCESS

Edited by:
Simona Carla Silvia Caravita, University of Stavanger, Norway

Reviewed by:
Sha Xie, Shenzhen University, China
Concetta Esposito, University of Naples Federico II, Italy

***Correspondence:**
Darcia Narvaez
dnarvaez@nd.edu

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

Received: *15 July 2021*
Accepted: *13 October 2021*
Published: *08 November 2021*

Citation:
Narvaez D, Gleason T, Tarsha M, Woodbury R, Cheng Y and Wang L (2021) Sociomoral Temperament: A Mediator Between Wellbeing and Social Outcomes in Young Children. Front. Psychol. 12:742199. doi: 10.3389/fpsyg.2021.742199

INTRODUCTION

Historically, research on morality has focused on cognition, emphasizing the development of moral reasoning and judgment (Kohlberg, 1984; Rest, 1986; Turiel, 2006). More recently, psychologists have started integrating the socio-emotional aspects of moral functioning, such as empathy and social cooperation (Killen and Smetana, 2015), with the cognitive domain (Padilla-Walker and Carlo, 2014), but little work to date has examined the role of psychological wellbeing as a precursor to moral outcomes. This oversight is puzzling given that attending to the needs of others is compromised if individuals are distracted by their own distress (Batson and Oleson, 1991).

Wellbeing, including not just the absence of psychopathology but also happiness and thriving, might well be essential for sociomoral capacities, such as empathy and conscience. As a foundation for interacting with the social world, a child's wellbeing theoretically fosters a particular "sociomoral temperament"—an orientation toward others—that either hinders or enables moral social behavior. Given the connections between poor wellbeing and chronic distress (e.g., Lanius et al., 2010), we hypothesized that wellbeing established early in development might be linked to sociomoral temperament, and that individual differences in wellbeing and consequent sociomoral temperament might contribute to variations in moral behavior. If so, intervention efforts designed to promote wellbeing and positive sociomoral temperaments in early childhood might have significant implications for children's social outcomes.

Empirical work supports the notion of a tie between wellbeing and moral outcomes. For example, the significantly compromised wellbeing caused by early toxic stress undermines self-regulatory capacities fundamental for sociality, such as the physiological stress response (Lupien et al., 2009) and vagal tone (Porges, 2011). Similarly, Kochanska and colleagues (e.g., Kochanska, 2002; Kochanska et al., 2005) have written extensively on the Mutually Responsive Orientation (MRO)—the emotional climate of the early parent-child relationship that promotes the development of conscience and the internalization of social mores. While the MRO and wellbeing are not synonymous, a positive emotional environment is a likely component of wellbeing in early childhood. How exactly positive wellbeing connects to outcomes is unclear, so in this work we focused on the details of the implicit systems that result in the "ethical knowhow" (Varela, 1999) necessary for children to exhibit moral behavior. These implicit systems mediating between psychological wellbeing and children's social outcomes is what we are assessing through our measure of sociomoral temperament.

WELLBEING AND SOCIOMORAL TEMPERAMENT

The idea that wellbeing might have implications for sociomoral temperament has its roots in the relational developmental systems view (Overton, 2013). This perspective suggests that human functioning is not just psychological but deeply embodied, such that a child's wellbeing is a function of neurobiological architecture that shapes implicit assumptions and expectations about social interactions and relationships (Schore, 2019). Animal studies demonstrate that poor wellbeing is associated with cacostatic responses to social situations, such as aggression or withdrawal (Harlow, 1958). We hypothesized that in humans, this link between wellbeing and behavior might be mediated by implicit expectations regarding social interactions, or sociomoral temperament.

Although the relation between wellbeing and sociomoral temperament has not received significant attention in children, for adults, states of wellbeing are related to moral functioning (e.g., Frederickson and Branigan, 2005), and dispositional wellbeing (e.g., agreeableness) predicts prosociality (Meier et al., 2006). Likewise, in adult retrospective reports, Narvaez et al. (2016b) found that poor wellbeing, identified as subclinical psychopathology (depression and anxiety), predicted self-protective sociomoral temperaments emphasizing social opposition or withdrawal, whereas higher wellbeing (a lack of subclinical psychopathology) was associated with a sociomoral temperament characterized by engagement and social approach. We tested whether similar relations could be established as early as the preschool years, given our theory that sociomoral temperament emerges early in development when personality temperament is likewise being formed (Rothbart and Bates, 1998).

SOCIOMORAL TEMPERAMENT AND SOCIAL OUTCOMES

Our hypothesis that sociomoral temperament would be linked to social outcomes is based in Triune Ethics Meta-theory (TEM; Narvaez, 2008, 2014, 2016). TEM joins the trend toward studying the effects of implicit embodied functioning (rather than cognitive structures only) on psychosocial functioning (Varela et al., 1991), particularly in moral psychology (Narvaez and Lapsley, 2005; Narvaez, 2010). It integrates the interaction of developmentally-relevant experience, neurobiological development, and personality formation to clinical and sub-clinical moral outcomes with the idea that these may help explain the gap between moral judgment and moral action (Lapsley and Narvaez, 2004b). For example, despite having learned and internalized social rules, a person might feel justified to act from fear or rage in the heat of the moment if a social stimulus triggers an intense stress response for protection (e.g., Gilligan, 1997; van der Kolk, 2014).

TEM also holds that neurobiological dispositions are influenced by early wellbeing. For example, the functioning of the hypothalamic–pituitary–adrenal (HPA) system is affected by wellbeing (Lupien et al., 2009), such that with highly stressful experiences, an individual might develop a disposition toward a hyper- or hypo-reactive stress response, possibly undermining capacities for the social attunement required for socioemotional intelligence (Goleman, 1995a,b). Further, early wellbeing is associated with vagus nerve function (vagal tone), which appears fundamental to capacities related to compassionate moral behavior such as social approach and social closeness (Porges, 2011). Experiences that promote physiological and psychological wellbeing may thus influence sociomoral temperament, promoting social *approach* rather than *withdrawal*.

The TEM framework for understanding sociomoral functioning focuses on three basic orientations rooted in global brain states (MacLean, 1990): protectionism, engagement, and imagination. When a particular global brain state or mindset guides behavior, it becomes an ethic. A protectionist orientation emerges from the activation of survival systems (e.g., stress response: fight-flight-freeze-faint; Sapolsky, 2004) and focuses on self-preservation through dominance or withdrawal. The stress response directs perception, thought, and

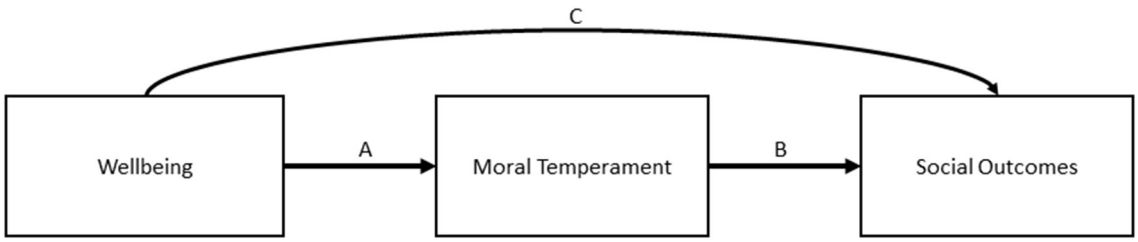


FIGURE 1 | Overall model predicting sociomoral temperament and social outcomes from wellbeing.

action in self-protective ways. For example, individuals whose neurobiological systems have a low threshold for activation of the stress response are likely to default to protectionist orientations (e.g., perceiving intentional harm when accidentally bumped, Crick and Dodge, 1994). In contrast, an engagement orientation draws on capacities for emotional presence, relational attunement, and unconditional positive regard (Rogers, 1961), which rely on developmental neurobiological capacities like vagal regulation (Porges, 2011) and social oxytocin release (Feldman, 2007). Well-functioning self-regulation, in combination with these neurobiological capacities, enables an engagement orientation. The ethic of imagination is undergirded by abilities to abstract and imagine possibilities outside the present moment. It adds creativity, intentionality, and abstraction into social relations and can be fueled by engagement or protectionist mechanisms.

In adults, triune ethics orientations have been examined with self-report measures that address how much a respondent thinks a list of characteristics represents an orientation they have and how much they think their friends and family would consider the list part of the respondent's characteristics (Narvaez, 2014; Narvaez and Hardy, 2016). Each list of characteristics relates to a particular type of mindset: protectionism (social opposition or withdrawal), engagement, or imagination [one that is generally reflective, detached (withdrawn), vicious (oppositional), or communal]. Although a person can shift among mindsets based on the situation, an adult sociomoral temperament defaulting to self-protectionism generally has been related to a personality pattern of distrust, aggression, and social dominance, as well as less prosocial behavior. In contrast, an engagement temperament has been related to greater agreeableness, conscientiousness, and prosocial behavior variably measured (Narvaez, 2014; Narvaez and Hardy, 2016; Narvaez et al., 2016a).

We tested whether individual differences in sociomoral temperaments would lead to variations in social behavior outcomes in the preschool years. Specifically, we examined variations in aspects of moral socialization, such as empathy and conscience, and in socially-relevant self-regulation, such as inhibitory control and misbehavior. Each of these outcomes is dependent upon a child's ability to regulate internal states sufficiently so as to attend to the requirements of a social situation and the needs of others. A child who has developed a sociomoral temperament of self-protectionism has likely encountered consistent social stress (i.e., associated with compromised wellbeing), which has resulted in behavior that prioritizes safety of the self over the wellbeing of others. Poor wellbeing, associated with repeated stress, could create a default focus on self-preservation (Shanker, 2016) that undermines social self-regulation. As a consequence, regulation in social settings might be compromised, resulting in impulsivity or externalizing behaviors toward others. In contrast, a child whose wellbeing is high will likely have developed the foundations for engagement, allowing for attention to the needs and concerns of others and successful regulation of behavior. Our hypothesis was that early positive wellbeing would be linked to a sociomoral temperament of approach and openness to social experience rather than avoidance or withdrawal. We also expected that sociomoral temperament would connect to social outcomes with implications for moral behavior, such as moral socialization and social self-regulation. Lastly, we expected the relation between wellbeing and outcomes would be mediated by sociomoral temperament.

CURRENT STUDIES

Our first goal was to develop a measure of sociomoral temperament, which we named the measure of Child SocioMoral Orientation (CSMO), and to confirm its factor structure. Our second goal was to examine each component of our hypothesized model of the relations between wellbeing, sociomoral temperament, and children's social outcomes (see **Figure 1**). Given that our overarching hypothesis was that children's wellbeing predicts individual differences in moral socialization and social self-regulation mediated by sociomoral temperament, we tested the relations between (a) our measures of wellbeing (happiness, thriving, anxiety, depression) and the CSMO scores (**Figure 1**, path A); (b) the CSMO scores and social outcomes (**Figure 1**, path B), including measures of moral socialization (empathy, concern after wrong doing, internalized conduct) and social self-regulation (inhibitory control, misbehavior); and (c) the extent to which CSMO mediated between wellbeing and social outcomes.

We expected strong positive relationships between wellbeing measures (e.g., thriving) and engagement scores on CSMO, and similarly positive relationships between illbeing (e.g., anxiety), and self-protectionism scores; likewise, we expected illbeing to be negatively correlated with engagement scores, and wellbeing

to be negatively correlated with self-protectionism. We also hypothesized that CSMO engagement scores would positively predict social outcomes (socialization and self-regulation), and that self-protectionism would negatively predict these same outcomes. Lastly, we hypothesized that CSMO scores would mediate the relations between wellbeing and child moral outcomes.

We collected large samples in two countries. We selected the USA and China for their historically distinctive cultures, one more individualistic and one more collectivistic (Triandis et al., 1990). We also assessed measurement invariance across the two cultures. Although USA and Chinese cultures differ in several ways, we anticipated that similar outcomes would be found for each sample, including that CSMO would have similar factor structures across cultures because of its focus on fundamental social approach-avoidance, which characterizes human interactions generally**.** In all our analyses, we controlled for gender and tested for gender differences since they have emerged in past research using some of the measures (e.g., Kochanska, 1994; Clark et al., 2016).

STUDY 1

The purposes of this study were to confirm the factor structure of the CSMO measure and assess whether it would mediate the relation between wellbeing and moral outcomes in young children.

Method

Participants

We collected data from US mothers ($n = 525$; *Mage* = 32.97, SD = 5.06 years, range: 20–49; median household income $50,000–$75,000; 84.4% married or in civil union) who reported on their 2- to 4-year-old children (*Mage* = 3.35, SD = 0.46, range 2–5; 295 sons, 229 daughters, 1 "Other").

Procedure

Participants were recruited through fliers at preschools and through electronic notices delivered by parenting-focused organizations, listservs, and blogs. Consent was gathered prior to the start of the survey. Data were collected *via* Qualtrics Survey Software, taking ∼45 mins to complete; participants were compensated with a gift card. Study design and data collection procedures were approved by the university's institutional review board.

Measures

All measures were parent report. In addition to measuring children's sociomoral temperament through CSMO, we included measures of children's wellbeing as predictors of CSMO, the mediator, and measures of self-regulation and moral socialization as child outcomes.

Predictors

Child Wellbeing. Child wellbeing was measured *via* four constructs. Five items measured the frequency with which the child demonstrated *happiness* (Gleason et al., 2016; e.g., "Dances spontaneously;" $\alpha = 0.72$) on a 6-point Likert scale (1 = never to 6 = more than once a day). *Thriving* was assessed using an adapted version (Gleason et al., 2016) of the Warwick-Edinburgh Wellbeing Scale ($n = 14$; e.g., "My child deals well with problems;" $\alpha = 0.91$) using a 6-point Likert scale (1 = never to 6 = always). We used a 17-item *depression* measure (Gleason et al., 2016; $\alpha = 0.92$) assessing frequency of depression-related behaviors (e.g., "How often does your child lack confidence?") using a 6-point Likert scale (1 = never to 6 = several times a day). *Anxiety* was measured using the Preschool Anxiety Scale (Spence et al., 2001; $n = 29$; e.g., "Is afraid of meeting or talking to unfamiliar people," $\alpha = 0.94$) rated on a 5-point Likert scale (0 = not true at all to 4 = very often true or not applicable).

Mediator

In order to measure sociomoral orientation in children, we ran a pilot study using items from the adult measure of triune ethics orientation (Narvaez and Hardy, 2016; Narvaez et al., 2016a), which measures various forms of self-protectionism (e.g., social withdrawal, social opposition), social engagement, and types of imagination (e.g., reflective, vicious, detached, communal), and adapted them for maternal report of the child in social situations. We added terms to capture children's visible behavior (e.g., freezes, excited). Items ($n = 71$) were randomly presented. Whether the items within the orientations of self-protectionism, engagement, and imagination would fall into factors similar to those in the adult measure was unclear (Narvaez, 2013), so we conducted exploratory factor analyses (EFA) within each of the three orientations (Kline, 2013). Participants included US mothers of preschool children ($N = 166$; 58.9% boys) recruited in the United States through parenting blogs, flyers, a parenting organization, and parenting listservs in the Midwest and Northeast to fill out an online survey in exchange for a gift card. Participants ranged in age from 18 to 48 years (M = 33.56, SD = 5.54). Most mothers (92.0%) were married and 96.9% had at least some college education. Yearly household income varied substantially. And the sample was 82.2% Caucasian Euro-American, with family size ranging from 2-9 people ($M = 4.43$, SD = 1.13), including an average of 2.12 adults (SD = 0.55) and 2.31 children (SD = 0.93).

The extraction method was principal component analysis (PCA) using oblimin rotation because this method is recommended (Kim and Mueller, 1978; Brown, 2009) when factors are expected to correlate. Retention of factors was based on eigenvalues larger than 1 (Kaiser, 1960), inspecting the scree plot (i.e., identifying an elbow), and the variance explained by retained factors.

Eighteen items did not fit into any factor; the remaining 53 items fell into seven factors. Two factors emerged with respect to Social Engagement: (a) Social enjoyment ($n = 9$; $\alpha = 0.93$) and (b) Social attunement ($n = 8$; $\alpha = 0.88$). Two factors were associated with Imagination: (c) Social consideration ($n = 6$; $\alpha = 0.84$) and (d) Social imagination ($n = 6$; $\alpha = 0.84$). Three factors emerged from the items associated with Self-Protectionism: (e) Social opposition ($n = 10$; $\alpha = 0.92$), (f) Social distrust ($n = 4$; $\alpha = 0.61$), and (g) Social withdrawal ($n = 10$; $\alpha = 0.91$). The seven subscales were then used in the current

studies (see **Supplemental Materials** for factor loadings and notes; the final set of items is listed by factor in the **Appendix**). We called the final version of the measure Child SocioMoral Orientation (CSMO).

Outcomes

Child Moral Socialization. Three different aspects of child morality were assessed *via* maternal report using subscales of the My Child survey (Kochanska et al., 1994) which have been demonstrated to correlate with child behavior (Kochanska, 1995): *Empathy* (n =13, $\alpha = 0.86$, "Will try to comfort or reassure another in distress"), *concern after wrongdoing* ($n = 8$, $\alpha = 0.88$, "When she or he does something wrong, seems to feel relieved when forgiven") and *internalized conduct* ($n = 9$, $\alpha = 0.86$, "Clearly hesitates before doing something forbidden, even when alone"). For all scales, mothers rated their children's behavior using a 7-point Likert scale (1 = extremely untrue of your child to 7 = extremely true of your child).

Child Self-Regulation. We measured self-regulation using the *inhibitory control* subscale from the Child Behavior Questionnaire (CBQ, Putnam and Rothbart, 2006; $n = 6$; e.g., "Can lower his/her voice when asked to do so"; $\alpha = 0.83$) rated on a 7-point Likert scale (1 = extremely untrue of your child to 7 = extremely true of your child). We measured frequency of *misbehavior* (Gleason et al., 2016; $n = 6$; e.g., "How often does your child misbehave?" $\alpha = 0.75$) using a 4-point Likert scale (1 = once a week or less, 2 = several times a week, 3 = every day, 4 = several times a day), except for one question that assessed recent misbehavior (i.e., "How often did your child misbehave in the last week?") rated on a 5-pt. scale (1 = "not at all," 5 = "over a dozen times").

Analytic Plan

To investigate factor structure, we used parallel analysis (Horn, 1965) on CSMO subscale scores. After assessing reliability, we ran correlation analyses between CSMO subscale and composite scores with established measures of wellbeing, self-regulation, and moral socialization. Because there were gender differences in prior research with self-regulation and moral socialization measures (i.e., girls scoring higher on inhibitory control, guilt, moral conduct; Kochanska et al., 1997, 2005), we investigated gender differences as well.

EFA analyses were conducted using R Statistical Language (R Core Team, 2016) and add-on packages, such as lavaan (Rosseel, 2012), psych (Revelle, 2016), car (Fox and Weisberg, 2011), Hmisc (Harrell, 2016), and QuantPsyc (Fletcher, 2012). Mediation models were conducted using Mplus (Muthén and Muthén, 2019).

For the EFAs, goodness of fit was assessed using root mean square error of approximation (RMSEA) values below 0.08 and comparative fit index (CFI) values greater than or equal to 0.95 (Hu and Bentler, 1999; Hooper et al., 2008). Chi-square is largely influenced by sample size and considering the number of participants in our studies, CFI and RMSEA were the indices used to evaluate model fit (Shi et al., 2019). However, chi-square was still assessed for all models. In addition, all mediation models are saturated and they have perfect model fit with chi-squares = 0, RMSEA = 0, and CFI =1.

The proportion of missing data differed by variable. In all models, we handled missing data with full-information-maximum-likelihood (FIML), which is suitable for handling data missing completely at random and missing at random (Enders and Bandalos, 2001).

TABLE 1 | Study 1 (USA) and Study 2 (China) EFA factor loadings for the child sociomoral orientation measure (CSMO).

	USA		China	
	Factor 1	**Factor 2**	**Factor 1**	**Factor 2**
Opposition	**0.46**	0.03	**0.57**	0.05
Distrust	**0.77**	0.03	**0.82**	−0.01
Withdrawal	**0.77**	−0.02	**0.62**	−0.03
Attunement	0.03	**0.93**	0.00	**0.84**
Consideration	0.02	**0.87**	0.06	**0.84**
Imagination	0.00	**0.77**	−0.01	**0.81**
Enjoyment	−0.07	**0.79**	−0.07	**0.79**

Bold values show significant loadings, $p < 0.05$.

Results

Factor Structure

Parallel analysis identified two factors. The first factor, labeled Self-Protectionism (SP), was measured by Social Opposition, Distrust, and Withdrawal. The second factor, termed Imaginative Relational Attunement (IRA), was measured by Social Enjoyment, Attunement, Imagination, and Consideration. This factor was named IRA because it was a combination of both Engagement and Imagination orientations. Utilizing PROMAX rotation, the two-factor model had good fit and a clean factor structure [robust $\chi^2_{(9)} = 32.22$, $p < 0.001$, CFI = 0.986, RMSEA = 0.07] and all standardized factor loadings were higher than 0.40 (see **Table 1**). Further, the two factors were slightly negatively correlated ($r = -0.12$, $p = 0.04$) and corresponded with our notions of sociomoral temperaments connected to social withdrawal and approach. Given that two factors accounted for CSMO items, SP and IRA, the composite scores of these subscales were used in proceeding analyses.

Descriptives and Correlations

Table 2 presents descriptive statistics. In this study, the proportion of missing data ranged from 2.7 to 9.6%, and sample size ranges are listed in under USA for boys, girls, and the total sample. After correcting for multiple tests, three gender differences emerged: girls had significantly higher means than boys on happiness $t_{(457.18)} = -4.42$, $p < 0.001$, empathy, $t_{(503.39)} = -3.74$, $p < 0.001$, and inhibitory control, $t_{(482.62)} = -5.14$, $p < 0.001$. Because of these differences, we used gender as a control variable in our models.

Table 3 contains the correlation coefficients (above diagonal). As predicted, positive correlations emerged between wellbeing and IRA and between illbeing and SP. The reverse was also mostly supported, although SP did not correlate with happiness.

TABLE 2 | Descriptive statistics for Study 1 (USA) and Study 2 (China) by country and gender with between-group t-tests.

	USA				China				Between-country comparisons		
	Boys (*n* = 253-295)	**Girls (*n* = 213-229)**	**Total (*n* = 475-511)**	**Between-gender *difference* for USA**	**Boys (*n* = 179-188)**	**Girls (*n* = 185-191)**	**Total (*n* = 365-380)**	**Between-gender *difference* for China**	**Boys**	**Girls**	**Total**
	***M* (SD)**	***M* (SD)**	***M* (SD)**	***t* (Cohen's d)**	***M* (SD)**	***M* (SD)**	***M* (SD)**	***t* (Cohen's *d*)**	***t* (Cohen's *d*)**	***t* (Cohen's *d*)**	***t* (Cohen's *d*)**
Wellbeing											
Happiness	5.24 (0.69)	5.49(0.52)	5.35 (0.63)	−4.42* (−0.40)	4.39 (0.95)	4.54(0.96)	4.47 (0.96)	−1.52 (−0.16)	10.38* (1.06)	12.01* (1.26)	15.42* (1.12)
Thriving	5.37 (0.55)	5.50(0.45)	5.43 (0.51)	−2.82 (−0.25)	4.73 (0.74)	4.73(0.78)	4.73 (0.76)	−0.03 (0.00)	10.01* (1.01)	11.85* (1.23)	15.30* (1.11)
Depression	2.08 (0.70)	1.90(0.64)	1.99 (0.68)	2.90 (0.26)	1.77 (0.66)	1.63(0.64)	1.70 (0.65)	2.03 (0.22)	4.69* (0.45)	4.02* (0.42)	6.38* (0.43)
Anxiety	1.49 (0.53)	1.42(0.48)	1.46 (0.51)	1.48 (0.14)	2.38 (0.88)	2.27(0.94)	2.32 (0.91)	1.19 (0.12)	−12.00* (−1.29)	−10.84* (−1.17)	−16.01* (−1.21)
Sociomoral temperament											
CSMO subscales											
Opposition	2.72 (0.99)	2.51 (0.94)	2.63 (0.98)	2.37 (0.21)	2.07 (0.83)	1.89 (0.69)	1.98 (0.77)	2.32 (0.23)	7.56* (0.70)	7.61* (0.74)	10.90* (0.73)
Distrust	2.12 (0.92)	2.02 (0.87)	2.07 (0.90)	1.17 (0.11)	1.73 (0.88)	1.59 (0.77)	1.66 (0.83)	1.52 (0.17)	4.56* (0.43)	5.21* (0.52)	6.96* (0.47)
Withdrawal	2.44 (0.80)	2.42 (0.80)	2.43 (0.80)	0.26 (0.025)	2.16 (0.75)	2.15 (0.72)	2.15 (0.73)	0.09 (0.01)	3.83* (0.36)	3.57* (0.35)	5.26* (0.36)
Attunement	5.09 (0.72)	5.25 (0.71)	5.16 (0.72)	−2.52 (0.22)	3.90 (1.06)	3.86 (1.21)	3.88 (1.14)	0.41 (0.04)	13.23* (1.36)	13.85* (1.43)	19.07* (1.38)
Consideration	4.92 (0.80)	5.06 (0.73)	5.00 (0.77)	−1.73 (−0.18)	3.71 (1.18)	3.55 (1.07)	3.64 (1.13)	1.37 (0.14)	12.22* (1.25)	16.06* (1.67)	19.75* (1.44)
Imagination	4.87 (0.83)	5.09 (0.83)	4.97 (0.84)	−2.89 (−0.26)	3.90 (1.18)	3.79 (1.17)	3.87 (1.17)	1.11 (0.09)	9.32* (0.98)	12.57* (1.30)	15.37* (1.10)
Enjoyment	5.58 (0.60)	5.65 (0.73)	5.61 (0.58)	−1.35 (−0.12)	4.72 (1.14)	4.62 (1.17)	4.67 (1.15)	0.78 (0.09)	9.45* (1.00)	10.98* (1.08)	14.43* (1.08)
CSMO factor scores											
Self-protectionism	2.43 (0.73)	2.32 (0.41)	2.38 (0.71)	1.69 (0.15)	1.99 (0.65)	1.88 (0.59)	1.93 (0.62)	1.65 (0.18)	14.53* (0.63)	15.81* (0.88)	9.75* (0.67)
IRA	5.12 (0.65)	5.26 (0.63)	5.18 (0.64)	−2.47 (−0.22)	4.07 (0.98)	3.96 (1.00)	4.02 (0.99)	0.87 (0.11)	7.09* (1.31)	11.03* (1.58)	12.55* (1.43)
Social outcomes											
Moral socialization											
Empathy	5.22 (0.61)	5.42 (0.58)	5.31 (0.60)	−3.74* (−0.33)	5.00 (0.75)	5.09 (0.60)	5.05 (0.68)	−1.35 (−0.13)	3.43* (0.33)	5.57* (0.56)	6.02* (0.41)
Concern after wrongdoing	5.06 (1.17)	5.28 (1.18)	5.16 (1.18)	−1.99 (−0.19)	5.20 (0.98)	5.22 (0.91)	5.21 (0.94)	−0.22 (−0.02)	−1.37 (−0.13)	0.50 (0.06)	−0.67 (−0.05)
Internalized conduct	4.24 (1.08)	4.42 (1.11)	4.32 (1.09)	−1.87 (−0.17)	4.10 (0.55)	4.15 (0.49)	4.13 (0.52)	−1.03 (−0.10)	1.85 (0.15)	3.25* (0.31)	3.42* (0.21)
Self-regulation											
Inhibitory control	4.79 (0.85)	5.18 (0.81)	4.97 (0.85)	−5.14* (−0.46)	4.64 (0.80)	4.86 (0.78)	4.75 (0.79)	−2.62 (−0.28)	1.95 (0.18)	4.12* (0.40)	3.96* (0.27)
Misbehavior	2.75 (0.50)	2.63 (0.41)	2.69 (0.47)	3.00 (0.26)	2.28 (0.50)	2.20 (0.34)	2.24 (0.45)	1.67 (0.18)	9.32* (0.94)	10.40* (1.13)	13.80* (0.98)

CSMO, child sociomoral orientation; IRA, imaginative relational attunement.
T-tests were performed with a Bonferroni correction (c = 18).
*$p < 0.002778$.

TABLE 3 | Within country pearson's correlations of all variables for Study 1 (USA) and Study 2 (China).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Wellbeing																			
1. Happiness	–	0.56**	−0.29**	−0.26**	0.55**	0.52**	0.46**	0.48**	0.48**	−0.01	−0.05	0.15**	−0.12**	0.37**	0.21**	0.11*	0.23**	−0.04	0.20**
2. Thriving	0.46**	–	−0.46**	−0.40**	0.58**	0.54**	0.51**	0.50**	0.52**	−0.18**	−0.19**	0.09	−0.29**	0.50**	0.22**	0.13**	0.31**	−0.19**	0.13**
3. Depression	−0.10*	−0.25**	–	0.40**	−0.26**	−0.24**	−0.23**	−0.20**	−0.25**	0.54**	0.52**	0.21**	0.47**	−0.28**	0.00	−0.21**	−0.30**	0.43**	−0.14
4. Anxiety	0.01	−0.14**	0.50**	–	−0.27**	−0.24**	−0.20**	−0.23**	−0.31**	0.37**	0.17**	0.20**	0.48**	−0.17**	0.01	−0.03	−0.12*	0.17**	−0.07
CSMO																			
5. IRA (comp.)	0.37**	0.43**	−0.12*	0.08	–	0.92**	0.89**	0.87**	0.86**	0.03	−0.02	0.24**	−0.14**	0.55**	0.22**	0.22**	0.42**	−0.16**	0.12*
6. Attunement	0.36**	0.38**	−0.10	−0.08	0.56**	–	0.82**	0.69**	0.74**	0.03	−0.05	0.23**	−0.10*	0.60**	0.26**	0.25**	0.41**	−0.18**	0.12**
7. Consideration	0.31**	0.34**	−0.05	0.05	0.72**	0.66**	–	0.66**	0.67**	−0.00	−0.16**	0.25**	−0.07	0.53**	0.27**	0.36**	0.52**	−0.27**	0.08
8. Imagination	0.32**	0.44**	−0.08	−0.06	0.55**	0.64**	0.69**	–	0.69**	0.07	0.06	0.23**	−0.14**	0.41**	0.11*	0.09	0.27**	−0.06	0.14**
9. Enjoyment	0.35**	0.40**	−0.16**	−0.13*	0.53**	0.70**	0.60**	0.63**	–	0.01	0.09	0.13**	−0.21**	0.40**	0.12	0.07	0.26**	−0.05	0.06
10. SP (comp.)	−0.02	−0.15**	0.52**	0.28**	−0.03	0.03	0.08	0.03	−0.02	–	0.74**	0.74**	0.79**	−0.11*	0.06	−0.14**	−0.17**	0.40**	−0.05
11. Opposition	−0.06	−0.07	0.45**	0.18**	−0.07	−0.01	0.04	0.11*	0.05	0.77**	–	0.25**	0.36**	−0.22**	−0.10*	−0.37**	−0.40**	0.60**	−0.11*
12. Distrust	0.01	−0.07	0.37**	0.22**	0.05	0.07	0.13*	0.05	−0.01	0.84**	0.48**	–	0.49**	0.13**	0.18**	0.08	0.13**	0.06	0.02
13. Withdrawal	−0.00	−0.21**	0.44**	0.28**	−0.04	0.02	0.03	−0.09	−0.09	0.78**	0.36**	0.53**	–	−0.15**	0.08	0.02	−0.06	0.18**	−0.02
Child outcomes																			
14. Empathy	0.29**	0.30**	−0.16**	0.03	0.72**	0.26**	0.22**	0.18**	0.27**	−0.06	−0.08	−0.01	−0.07	–	0.29**	0.27**	0.44**	−0.28**	0.16**
15. Concern	0.22**	0.31**	−0.09	0.07	0.76**	0.21**	0.18**	0.22**	0.21**	−0.12*	−0.15**	−0.04	−0.10	0.55**	–	0.19**	0.21**	−0.07	0.10*
16. Int. conduct	0.08	0.14**	−0.04	0.10	0.23**	0.06	0.12*	0.03	0.03	−0.02	−0.09	0.00	0.06	0.32**	0.15**	–	0.62**	−0.40**	0.08
17. Inh. control	0.15**	0.28**	0.17**	−0.00	0.54**	0.31**	0.33**	0.22**	0.25**	−0.14**	−0.28**	−0.01	−0.06	0.43**	0.45**	0.31**	–	−0.41**	0.23**
18. Misbehavior	−0.05	−0.01	0.20**	0.03	−0.09	−0.07	−0.13*	0.02	0.01	0.32**	0.45**	0.19**	0.07	−0.02	−0.00	−0.16**	−0.22**	–	−0.15**
19. Gender	0.04	−0.01	−0.11**	−0.06	−0.02	−0.02	−0.08	−0.06	−0.04	−0.10	−0.12*	−0.12*	−0.01	0.07	0.01	0.05	0.14**	−0.07	–

IRA (comp.), Imaginative Relational Attunement (composite), SP (comp.), self-protectionism (composite).

USA and China correlation coefficients are above and below the diagonal respectively.

$^{*}p < 0.0$; $^{**}p < 0.01$.

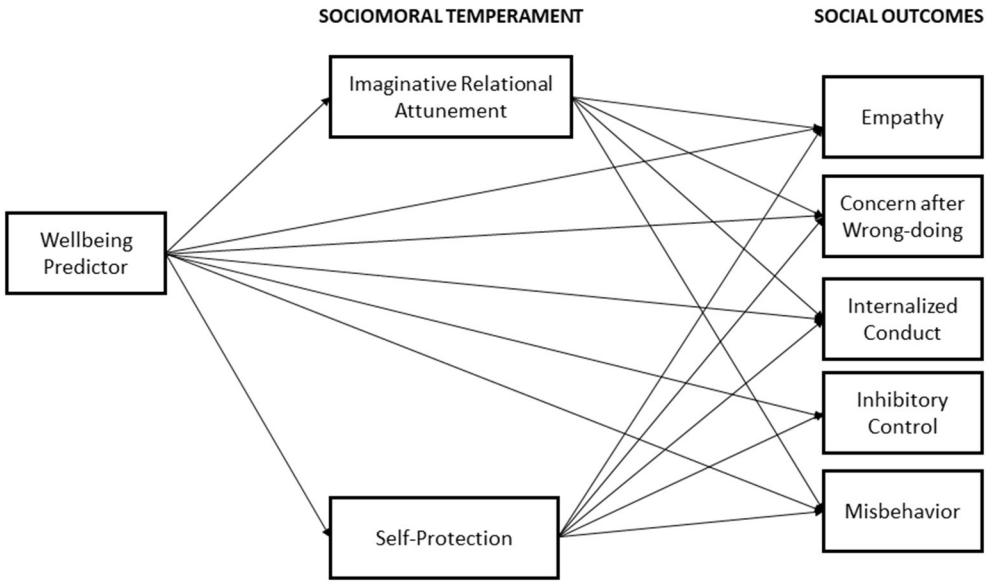


FIGURE 2 | Mediation model with moral temperament mediating relation between wellbeing and social outcomes.

The fact that SP did not correlate with happiness, whereas IRA did, suggests the importance of treating IRA and SP separately. Generally, these findings are consistent with the idea that psychological systems are associated with differentiated sociomoral orientations.

Most correlations between CSMO subscales and the self-regulation and moral socialization measures were in the expected directions. The results for the self-protectionism subscales were mixed. Most correlations were negative, as hypothesized, except that distrust correlated positively with empathy, concern after wrongdoing, and inhibitory control. Correlations were not significant between social distrust and internalized conduct and misbehavior, nor did social withdrawal correlate with concern, internalized conduct, or inhibitory control.

Contributions of Sociomoral Temperament When Mediating the Relationship Between Wellbeing and Child Outcomes

We next tested the hypothesis that CSMO composite scores would mediate the relations between wellbeing and child social outcomes (**Figure 1**, path C). We constructed four models, each of which used a wellbeing measure (happiness, thriving, depression, and anxiety) as a predictor, IRA and SP as mediators, and the moral socialization (empathy, concern after wrong doing, internalized conduct) and self-regulation (inhibitory control, misbehavior) measures as outcomes (see **Figure 2** for the mediation model and **Table 4** for model results). All child outcomes were allowed to correlate, yielding four saturated, perfect fitting models, $\chi^2_{(0)} = 1.00$, $p = 0.00$, CFI = 1.00, RMSEA = 0.00. Gender was included as a control variable.

Mediation analyses tested both total (i.e., both mediators together) and specific (i.e., each CSMO mediator) indirect mediation effects of wellbeing on child outcomes. For specific indirect effects, IRA significantly mediated the relationships from all four wellbeing measures to all five outcomes at $ps \leq 0.002$ with two caveats: the strengths of the specific indirect paths from thriving to concern and depression to misbehavior were lower but still significant ($p \leq 0.045$). For SP, the results were more mixed. None of the specific indirect effects of SP were significant when mediating happiness and the five outcomes. However, SP mediated the relationship from thriving and depression to three outcomes: internalized conduct and inhibitory control ($ps \leq 0.04$), and misbehavior ($ps \leq 0.006$). SP also mediated paths between anxiety and all outcomes except concern ($ps \leq 0.008$).

Regarding direct effects, happiness predicted concern after wrong-doing ($p = 0.03$) and, contrary to our hypotheses, positively predicted misbehavior ($p = 0.02$). Thriving directly predicted empathy and concern ($p \leq 0.009$), whereas depression predicted inhibitory control ($p = 0.02$) and misbehavior ($p < 0.001$). Both depression and anxiety predicted internalized conduct ($p \leq 0.045$).

Discussion

These findings suggest that fostering wellbeing may be a significant contribution to future moral socialization and self-regulation *via* sociomoral temperament. Specifically, the connections between our measures of wellbeing and moral socialization were successfully mediated in whole or in part by our measure of sociomoral temperament. The relations were stronger for the self-regulatory outcomes, in that those predicted by thriving and anxiety were not accompanied by direct effects. These stronger relations to self-regulatory than moral socialization outcomes might be a function of development, in that the former might emerge earlier than the latter (Berger et al., 2007). Also, as the relation between physiological wellbeing and self-regulation is well-established (Lupien et al., 2009), these

TABLE 4 | Study 1 (USA) summary of mediation effects.

Predictor outcome	â: path to IRA (*p*)	b: path from IRA (*p*)	â: path to SP (*p*)	b: path from SP (*p*)	â * b IRA [95% CI]	â * b SP [95% CI]
Happiness	2.18 (<0.001)		0.00 (1.00)			
Empathy		0.12 (<0.001)		−0.04 (0.001)	0.26 (0.20, 0.32)	0.00 (−0.02, 0.01)
Concern		0.60 (0.009)		0.03 (0.234)	0.14 (0.03, 0.26)	0.00 (−0.01, 0.02)
Int. conduct		0.10 (<0.001)		−0.08 (0.001)	0.22 (0.12, 0.34)	0.00 (−0.03, 0.02)
Inh. control		0.14 (<0.001)		−0.08 (<0.001)	0.30 (0.22, 0.39)	0.00 (−0.03, 0.02)
Misbehavior		−0.05 (<0.001)		0.11 (<0.001)	−0.16 (−0.13, −0.05)	0.00 (−0.01, 0.05)
Thriving	2.85 (<0.001)		−0.69 (<0.001)			
Empathy		0.10 (<0.001)		−0.03 (0.025)	0.27 (0.20, 0.36)	0.02 (0.00, 0.04)
Concern		0.05 (0.035)		0.05 (0.087)	0.16 (0.00, 0.32)	−0.03 (−0.09, 0.01)
Int. conduct		0.10 (<0.001)		−0.09 (0.001)	0.29 (0.17, 0.42)	0.06 (0.02, 0.11)
Inh. control		0.12 (<0.001)		−0.07 (<0.001)	0.35 (0.26, 0.45)	0.05 (0.02, 0.09)
Misbehavior		−0.03 (0.002)		0.11 (<0.001)	−0.09 (−0.15, −0.04)	−0.07 (−0.12, −0.03)
Depression	−0.99 (<0.001)		0.05 (0.740)			
Empathy		0.12 (<0.001)		−0.03 (0.063)	−0.11 (−0.16, −0.06)	−0.04 (−0.09, 0.01)
Concern		0.10 (<0.001)		0.01 (0.744)	−0.09 (−0.15, −0.04)	0.02 (−0.10, 0.13)
Int. conduct		0.08 (<0.001)		−0.05 (0.119)	−0.07 (−0.12, −0.04)	−0.07 (−0.16, 0.01)
Inh. control		0.12 (<0.001)		−0.05 (0.018)	−0.11 (−0.16, −0.06)	−0.08 (−0.14, −0.01)
Misbehavior		−0.02 (0.029)		0.07 (<0.001)	0.02 (0.00, 0.04)	0.20 (0.12, 0.28)
Anxiety	−1.22 (<0.001)		1.39 (<0.001)			
Empathy		0.16 (<0.001)		−0.04 (0.005)	−0.15 (−0.24, −0.08)	−0.05 (−0.09, −0.01)
Concern		0.10 (<0.001)		0.03 (0.373)	−0.12 (−0.24, −0.04)	0.04 (−0.06, 0.13)
Int. conduct		0.12 (<0.001)		−0.10 (<0.001)	−0.12 (−0.22, −0.07)	−0.14 (−0.25, −0.06)
Inh. control		0.14 (<0.001)		−0.09 (<0.001)	−0.12 (−0.27, −0.08)	−0.12 (−0.18, −0.07)
Misbehavior		−0.03 (<0.001)		0.11 (<0.001)	0.04 (0.02, 0.07)	0.15 (0.10, 0.22)

IRA, imaginative relational attunement; SP, self-protectionism; CI, confidence interval.
The confidence intervals were bootstrapped. The formula is: point estimate $\pm z^* s/\sqrt{n}$.

results suggest that one mechanism of that connection in early childhood might be sociomoral temperament.

STUDY 2

Our goal in Study 2 was to test whether these models applied equally well in a different culture. Despite the significant cultural variation between the USA and China, we expected CSMO to have a similar factor structure because of its focus on fundamental social approach-avoidance, which characterizes human interactions generally. This second sample afforded examination of measurement invariance across cultures as well as group comparisons with the USA data. We also conducted the same examination of mediation effects of sociomoral temperament in relation to wellbeing and sociomoral outcomes as in Study 1.

Method

Participants

Data were collected from Chinese mothers ($n = 382$; *Mage* = 33.83 years; range: 21–45; median income $120,000–$160,000 US; 95% of mothers married or in civil unions). They were recruited through six Beijing preschools (Child *Mage* = 4.48, *SD* =0.90; 188 girls, 191 boys, 3 missing). Institutional review board approval and consent for participation were gathered before the survey was started. Participants received a parenting book to compensate them for their time.

Procedure

Chinese mothers responded *via* paper and pencil to all the same measures as USA mothers in Study 1.

Measures

Most measures had been translated and validated in Chinese in a previous study (Narvaez et al., 2013). The remainder was translated into Mandarin (including standard procedures for back translation).

Child Wellbeing. The same measures were used as in Study 1. Happiness had an alpha of 0.62; thriving $\alpha = 0.81$; depression $\alpha = 0.86$; anxiety $\alpha = 0.90$.

Child Moral Socialization. The same measures were used as in Study 1. Empathy had an alpha of 0.80; concern after wrongdoing $\alpha = 0.78$; internalized conduct $\alpha = 0.60$.

Child Self-Regulation. The same measures were used as in Study 1. Inhibitory control had an alpha of 0.81; misbehavior $\alpha = 0.67$.

Analytical Plan

We conducted a confirmatory factor analysis. In order to be able to compare country means, we tested for measurement invariance.

TABLE 5 | Goodness-of-fit indices and Chi-square difference tests for models testing measurement invariance with respect to country.

	Model compared	Chi-square	df	Chi. Dif.	df Dif.	*p*-value
1. Baseline (configural)		210.15	26	–	–	–
2a. Invariant loadings	1	235.20	31	19.27	5	0.002
2b. Invariant loadings (free: $\Lambda_{enjoyment}$)	1	214.16	30	3.54	4	0.472
3a. Invariant loadings and intercepts (free: $\Lambda_{enjoyment}$)	2b	269.59	35	44.68	5	<0.001
3b. Invariant loadings and intercepts (free: $\Lambda_{enjoyment}$; $\upsilon_{opposition}$)	2b	226.08	34	8.93	4	0.063
4. Invariant loadings, intercepts, and residuals (free: $\Lambda_{enjoyment}$; $\upsilon_{opposition}$)	3b	693.35	41	374.85	7	<0.001

Λ = *factor loading;* υ = *manifest intercept.*

Results

Factor Structure and Measurement Invariance

We detected a similar factor structure as in the USA data. We found that a two-factor CFA model acceptably fit the data [robust $\chi^2_{(11)} = 31.66$, $p < 0.001$, CFI = 0.98, RMSEA = 0.07] and all standardized factor loadings were higher than 0.50 (see **Table 1**). Self-Protectionism and Imaginative Relational Attunement were not significantly correlated.

After confirming the two-factor structure for the CSMO within China, we tested for measurement invariance between the USA and China at the scale level. We obtained partial invariance by freeing two parameters (factor loadings of "social enjoyment" and "social opposition" intercepts) and constraining all other factor loadings and manifest variable intercepts to be equal across countries (model 3b). Model 3b was the final model and suggested partial strong invariance (see **Table 5**). With such a model, we were able to compare the means of the two factors between countries.

Table 2 includes comparisons on each scale between and within country and gender. In this study, missing data ranged from <1 to 4% and sample ranges are listed under China for boys, girls, the total sample. We also investigated whether the factor score differences by country were linked to specific CSMO subscales. For all CSMO subscales, USA children scored significantly higher than Chinese children, even after a Bonferroni correction. No gender differences in CSMO scales emerged within country. However, we wondered whether CSMO subscale differences between countries were related to gender; the lack of correlations between countries suggested that this difference may not be related to gender. USA boys and girls scored significantly higher than their Chinese counterparts on all CSMO scales and subscales.[1]

Correlations are presented in **Table 3** (below diagonal). The hypothesis that wellbeing would be positively correlated with engagement scores and ill-being with self-protectionism was supported (**Figure 1**, path A), but the reverse was not well-supported. Depression was only negatively correlated with IRA and enjoyment, and anxiety only with the latter; thriving was only negatively correlated with SP and withdrawal. However, these few negative correlations underscore the idea that engagement and self-protection are not exclusive of one another but that an individual can display one or the other in different social situations (Narvaez, 2014), which a social-cognitive theory would predict (Lapsley and Narvaez, 2004a).

Similar to the USA, we found partial support for the hypothesis that CSMO engagement scores and social outcomes of self-regulation and moral socialization would be positively correlated (**Figure 1**, path B). Few negative correlations emerged between CSMO Self-Protectionism scores and child outcomes, although they were positively correlated (as expected) with misbehavior. In general, IRA predicted some aspects of moral socialization, but only misbehavior was linked to SP. Withdrawal in particular was unrelated to any child outcomes, suggesting that Chinese mothers did not associate fearfulness with moral socialization or self-regulation.

Contributions of Sociomoral Temperament When Mediating the Relationship Between Wellbeing and Child Outcomes

We tested for mediation effects using the same model as the USA sample (**Figure 1**, path C), which also yielded four saturated, perfect fitting models, $\chi^2_{(0)} = 1.00$, $p = 0.00$, CFI = 1.00, RMSEA = 0.00 (see **Table 6**). For specific indirect effects, IRA significantly mediated the same relationships as the total indirect effects for happiness and thriving ($ps \leq 0.002$). For depression, IRA significantly mediated the relationship with empathy and inhibitory control ($ps = 0.03$). IRA did not yield any significant specific indirect effects for anxiety. SP did not mediate relationships between happiness, thriving, and the five outcome variables. However, SP significantly mediated the relationship between depression, anxiety and concern after wrong-doing ($ps \leq 0.005$) and between anxiety and inhibitory control ($p = 0.004$).

Direct effects emerged for happiness in relation to empathy and concern after wrong-doing ($ps \leq 0.001$). Thriving likewise directly predicted all but concern ($ps \leq 0.02$). Depression directly predicted empathy ($p = 0.02$) and anxiety directly predicted internalized conduct ($p = 0.02$).

DISCUSSION

In this sample, a two-factor structure for CSMO again emerged, and sociomoral temperament mediated some of the relations

[1]When the same item is given to different groups of respondents, they may read it differently owing to cultural differences, which is why we examined measurement invariance across countries. We did not examine gender invariance as the items were given to mothers of small children, instead of children themselves, and we did not expect child gender to influence a mother's reading or interpretation of our items.

TABLE 6 | Study 2 (China) summary of mediation effects.

Predictor outcome	â: path to IRA (*p*)	b: path from IRA (*p*)	â: path to SP (*p*)	b: path from SP (*p*)	â * b IRA [95% CI]	â * b SP (95% CI)
Happiness	1.29 (<0.001)		−0.03 (0.737)			
Empathy		0.04 (<0.001)		−0.02 (0.374)	0.05 (0.02, 0.09)	0.00 (−0.01, 0.01)
Concern		−0.01 (0.797)		−0.13 (0.001)	−0.01 (−0.10, 0.08)	0.003 (−0.03, 0.02)
Int. conduct		0.01 (0.451)		−0.01 (0.727)	0.01 (−0.01, 0.02)	0.00 (−0.03, 0.02)
Inh. control		0.07 (<0.001)		0.002 (0.738)	0.07 (0.05, 0.14)	−0.06 (−0.02, 0.01)
Misbehavior		−0.02 (<0.001)		0.00 (0.965)	−0.02 (−0.06, −0.001)	0.00 (−0.003, 0.003)
Thriving	2.33 (<0.001)		−0.23 (0.050)			
Empathy		0.03 (0.002)		−0.01 (0.650)	0.03 (0.02, 0.14)	−0.01 (−0.01, 0.01)
Concern		0.03 (0.159)		−0.14 (<0.001)	0.07 (−0.01, 0.17)	0.03 (−0.001, −0.08)
Int. conduct		0.00 (0.997)		−0.00 (0.959)	0.00 (−0.04, 0.04)	−0.00 (−0.01, 0.01)
Inh. control		0.05 (<0.001)		−0.05 (0.008)	0.13 (0.07, 0.21)	0.01 (−0.003, 0.03)
Misbehavior		−0.01 (0.343)		−0.01 (0.532)	−0.01 (−0.03, 0.004)	0.001 (−0.01, 0.009)
Depression	−0.73 (0.020)		1.44 (<0.001)			
Empathy		0.05 (<0.001)		0.01 (0.818)	−0.03 (−0.07,−0.008)	0.01 (−0.04, 0.07)
Concern		0.04 (0.012)		−0.12 (0.002)	−0.03 (−0.08, 0.003)	−0.17 (−0.34, −0.04)
Int. conduct		0.01 (0.221)		−0.00 (0.807)	0.01 (−0.02, 0.01)	−0.00 (−0.06, 0.05)
Inh. control		0.07 (<0.001)		−0.04 (0.066)	−0.05 (−0.01, −0.12)	−0.06 (−0.14, 0.00)
Misbehavior		−0.01 (0.010)		0.00 (0.899)	0.01 (.00, 0.02)	0.002 (−0.05, 0.04)
Anxiety	−0.18 (0.430)		0.60 (<0.001)			
Empathy		0.05 (<0.001)		−0.02 (0.225)	−0.01 (−0.03, 0.01)	−0.01 (−0.04, 0.01)
Concern		0.02 (0.324)		−0.13 (0.001)	−0.08 (−0.03, 0.01)	−0.08 (−0.14, −0.35)
Int. conduct		0.01 (0.150)		−0.02 (0.251)	−0.01 (−0.01, 0.004)	0.07 (−0.03, 0.01)
Inh. control		0.07 (<0.001)		−0.07 (0.001)	−0.01 (−0.05, 0.02)	−0.04 (−0.08, −0.02)
Misbehavior		−0.02 (<0.001)		0.01 (0.628)	0.00 (−0.01, 0.02)	0.04 (−0.09, 0.01)

IRA, Imaginative Relational Attunement; SP, self-protectionism; CI, confidence interval.
The confidence intervals were bootstrapped. The formula is: point estimate $\pm z^* s/\sqrt{n}$.

between wellbeing and social outcomes. The mediation analyses suggest the greatest role for sociomoral temperament was between the illbeing predictors (depression and anxiety) and concern and inhibitory control, at least with respect to total indirect effects. Similar patterns emerged for total indirect and IRA-specific mediation between happiness and inhibitory control and misbehavior. In comparison to the USA sample, these findings suggest possible cultural differences with respect to the role of sociomoral temperament in connecting wellbeing and social outcomes.

Most of the CSMO differences between subscales that emerged were country rather than gender differences, with the USA having significantly higher scores even after adjustments for multiple comparisons. For the subscales associated with IRA, these findings are consistent with other research indicating that American mothers tend to emphasize their children's successes and deemphasize their failures, whereas Chinese mothers do the opposite (Ng et al., 2007). Although CSMO factors are not measuring successes and failures per se, the generally positive valence of the items might have elicited higher endorsement from American mothers than Chinese mothers. Additionally, if expectations about the appropriateness of displaying fear, anger, or timidity differ between cultures, what children show and how parents rate those behaviors might also differ (Louie et al., 2013). At the same time, the Chinese mothers scored their children lower on every CSMO factor subscale except anxiety. Chinese culture is a shame-socialized culture, emphasizing maintenance of others' approval, especially elders (filial piety), and avoidance of disappointing others (Schoenhals, 1993); consequently, parenting even of young children tends to emphasize right behavior, compliance after wrongdoing and making amends (Fung, 1999). Perhaps low scores in China occurred because Chinese citizens tend to minimize emotions in their lives (Ryder et al., 2008), and they have specific longstanding etiquette rules for behavior such as self-control and obedience to elders, which they may consequently judge more harshly (Conrad, 2019).

GENERAL DISCUSSION

In these studies, we hypothesized that sociomoral temperament would mediate the relationship between wellbeing and sociomoral outcomes in two countries, the USA and China. As hypothesized, we found a two-factor solution for our measure of sociomoral temperament in both samples for Imaginative Relational Attunement (IRA) and Self-Protectionism (SP), and in both samples, mediation analyses demonstrated that these factors influenced the relationship between wellbeing and social outcomes. These findings support the theory that psychological wellbeing influences moral development

through children's orientation toward the social world in early childhood, though results need to be understood in light of their cross-sectional nature.

Similarities and differences emerged in the patterns when examining models across countries. For example, IRA mediated everything in the USA but not in China. With respect to child outcomes, at least one factor of sociomoral temperament effectively mediated the relationship with the wellbeing measures in the USA, and connections emerged for all outcomes except internalized conduct for China. For internalized conduct, the Chinese scores were less varied than in the US sample, which might account for the lack of mediation, and further exploration of parents' conceptualization of internalized conduct across cultures would illuminate these relationships at this point in development.

Across both countries IRA demonstrated a higher number of specific and direct effects than SP. One interpretation of this result is that an engaged sociomoral temperament plays a greater and more varied role in mediating between wellbeing and social outcomes than does self-protectionism. However, the samples were drawn from typically-developing, middle-class populations that scored higher on thriving and happiness and lower on anxiety and depression. They exhibited more variation on the IRA than the SP scores. A wider range of scores on SP, such as that obtainable in clinical populations, might result in greater variation. In typically-developing samples like these, who have higher happiness and thriving scores, IRA might be a more typical outcome. Thus, as a positive mediator, IRA might have captured more variance than SP. This idea does not negate SP as a mediator but does suggest a need for exploration in samples that have less attenuated scores.

The Importance of Wellbeing for Sociomoral Temperament, Moral Socialization, and Self-Regulation

Our analyses demonstrated relations between wellbeing, self-regulation, and moral socialization with many relations mediated by CSMO scales in the USA and fewer in China. These results have several implications for understanding the development of morality in early life and possible directions for future research, including the embodied nature of moral functioning, the developmental progression of capacities relevant to moral behavior, and the usefulness of sociomoral temperament as a construct and component of moral development.

Theoretically, both wellbeing and CSMO orientations are based in biosocial functioning. Understanding the psychobiological mechanisms that lead to individual variations in wellbeing, and concomitantly, sociomoral temperament, is a promising avenue for future research. For example, parental responsivity, especially given its relation to moral socialization measures (Kochanska, 2002), is likely predictive of the enhanced wellbeing that would foster development of an IRA orientation. However, other experiences, such as those with direct links to biosocial processes such as self-regulation, might have significant influence on the neurological underpinnings of a self-protectionist versus a relational attunement social orientation (Narvaez, 2014).

The fact that sociomoral temperament mediated the relations between wellbeing and self-regulation more so than moral socialization—and that these findings emerged in both samples (albeit in different patterns)—might be explained by the point in development at which these processes were measured (Berger et al., 2007). Children's development of empathy and concern after wrongdoing, for example, requires a significant cognitive component that includes attention to another's feelings—a capacity that is not fully formed in early childhood. Inhibitory control and the prevention of misbehavior, in contrast, are self-focused behaviors influenced by social contexts. These behaviors are components of moral development because of their implications for interpersonal relationships, but their mastery, depending as it does on curbing impulsivity, might be experienced intrapersonally at least some of the time, meaning that children might confront, practice, and master these developmental tasks prior to those that involve the perspectives of others. If so, the implication of these findings is that at this point in early childhood, wellbeing might be particularly important for promoting self-regulation both directly and through the construction of the implicit systems measured by sociomoral temperament. Whether these relations emerge similarly at other points in development is a topic for future research.

Mediating effects of sociomoral temperament across both countries also point to the possible explanatory power of TEM. Regardless of cultural influences, sociomoral temperament, as understood through social approach (IRA) and social withdrawal (SP) temperaments, helps explain the relationship between wellbeing and child social outcomes in early childhood. Similar to the TEM adult studies (Narvaez, 2014; Narvaez and Hardy, 2016; Narvaez et al., 2016a), evidence from this study suggests that TEM global temperaments are important variables in understanding the links between wellbeing and social outcomes. Our results contribute to this growing literature but also extend it by providing evidence that sociomoral temperament is important for understanding *child* wellbeing and social outcomes. The mediation model suggests that the exclusion of sociomoral temperament may lead to misrepresentation of the relationship between child wellbeing and social outcomes. For both research and practice, including sociomoral temperament may be necessary to both adequately investigate and understand the way in which child wellbeing relates to social outcomes.

Limitations and Future Directions

This paper has several limitations. First, it was a cross-sectional set of studies which allow only a glimpse into potential developmental trajectories. Second, the data were collected in different ways, online across the USA in one sample and on paper in several preschools in China. These strategies may have had an effect on the nature of who was recruited and how they responded. Third, the correlations found between our measures of sociomoral temperament and other child outcomes might have been inflated by the use of maternal report for both. Observational measures should be used in future studies. Fourth,

while the comparison of the USA and China is useful, truly establishing cultural invariance of CSMO requires more samples. Fifth, we did not include tests of the childhood environment so we cannot relate CSMO scores to childhood experiences (although see Tarsha et al., 2020). Future work might use CSMO to assess relationships between aspects of wellbeing and changes or stability in sociomoral temperament over time. Such an approach might illustrate whether the timing and intensity of experiences related to wellbeing are critical to the development of a more flexible and open sociomoral temperament.

Despite these limitations, our results do provide some support for the role of sociomoral temperaments in children's social outcomes. According to TEM, these different sociomoral orientations are based upon global brain states or neurobiological dispositions influenced by wellbeing. Future research should consider taking into consideration the components of wellbeing that promote or deter social approach or withdrawal.

CONCLUSION

The findings presented here support the idea that early wellbeing influences social outcomes and that sociomoral temperament helps explain this relationship. Our results suggest that sociomoral temperament is a mechanism worth investigating in explaining the connection between psychological wellbeing and young children's moral development. The pattern of findings between the USA and China samples suggests that some relation between wellbeing, sociomoral temperament, and social outcomes might be universal, but that differences in cultural expectations, perhaps with respect to children's obedience to adult authority (Fung, 1999), might govern the exact connections that emerge. As children's wellbeing helps formulate children's sociomoral temperament, the inclination to approach or withdraw from social interaction, the ways in which they learn to function in the world appear to have significance for later social capacities. Viewing morality as a function of holistic wellbeing might have significant ramifications for understanding how psychological functioning influences children's nascent morality and consequent social functioning.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Notre Dame Institutional Review Board. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

We thank the Institute for Scholarship in the Liberal Arts at the University of Notre Dame for its financial support.

ACKNOWLEDGMENTS

We thank Lifang Deng for her assistance in collecting data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.742199/full#supplementary-material

REFERENCES

Batson, C. D., and Oleson, K. C. (1991). "Current status of the empathy-altruism hypothesis," in *Prosocial Behavior*, ed M. S. Clark (Sage Publications, Inc.), 62–85.

Berger, A., Kofman, O., Livneh, U., and Henik, A. (2007). Multidisciplinary perspectives on attention and the development of self-regulation. *Prog. Neurobiol.* 82, 256–286. doi: 10.1016/j.pneurobio.2007.06.004

Brown, J.D. (2009). *Choosing the Right Type of Rotation in PCA and EFA*. Shiken: JALT Testing and Evaluation SIG Newsletter.

Clark, D. A., Listro, C. J., Lo, S. L., Durbin, C. E., Donnellan, M. B., and Neppl, T. K. (2016). Measurement invariance and child temperament: an evaluation of sex and informant differences on the Child Behavior Questionnaire. *Psychol. Assess.* 28:1646. doi: 10.1037/pas0000299

Conrad, R. (2019). *Culture Hacks: Deciphering Differences in American, Chinese, and Japanese Thinking*. New York, NY: Lioncrest.

Crick, N. R., and Dodge, K. A. (1994). A review and reformulation of social information processing mechanisms in children's social adjustment. *Psychol. Bull.* 115, 74–101. doi: 10.1037/0033-2909.115.1.74

Enders, C. K., and Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Struct. Equ. Model.* 8, 430–457. doi: 10.1207/S15328007SEM0803_5

Feldman, R. (2007). Parent-infant synchrony: biological foundations and developmental outcomes. *Curr. Direct. Psychol. Sci.* 16, 340–345. doi: 10.1111/j.1467-8721.2007.00532.x

Fletcher, T. D. (2012). *QuantPsyc: Quantitative Psychology Tools (Version 1.5)* [R package]. Available online at: https://CRAN.R-project.org/package=QuantPsyc (accessed June 15, 2015).

Fox, J., and Weisberg, S. (2011). *An R Companion to Applied Regression (2nd Edn.)*. Thousand Oaks, CA: Sage.

Frederickson, B. L., and Branigan, C. (2005). Positive emotions broaden the scope of attention and thought-action repertoires. *Cogn. Emot.* 19, 313–332. doi: 10.1080/02699930441000238

Fung, H. (1999). Becoming a moral child: the socialization of shame among young Chinese children. *Ethos* 27, 180–209. doi: 10.1525/eth.1999.27.2.180

Gilligan, J. (1997). *Violence: Reflections on a National Epidemic*. New York, NY: Vintage.

Gleason, T., Narvaez, D., Cheng, Y., Wang, L., and Brooks, J. (2016). "Well-being and sociomoral development in preschoolers: the role of maternal parenting attitudes consistent with the evolved developmental niche," in *Contexts for Young Child Flourishing: Evolution, Family and Society*, eds. D. Narvaez,

J. Braungart-Rieker, L. Miller, L. Gettler, and P. Hastings (New York, NY: Oxford University Press), 166–184. doi: 10.1093/acprof:oso/9780190237790.003.0008

Goleman, D. (1995a). *Social Intelligence: The New Science of Human Relationships*. New York, NY: Arrow.

Goleman, D. (1995b). *Emotional Intelligence: Why it Can Matter More Than IQ*. New York, NY: Bantam Books.

Harlow, H. (1958). The nature of love. *Am. Psychol.* 13, 673–685. doi: 10.1037/h0047884

Harrell, F. E. Jr. (2016). *Hmisc: Harrell Miscellaneous (Version 4.0-2)* [R package]. Available online at: https://CRAN.R-project.org/package$=$Hmisc (accessed June 15, 2015).

Hooper, D., Coughlan, J., and Mullen, M. R. (2008). Structural equation modelling: guidelines for determining model fit. *Electron. J. Bus. Res. Methods* 6, 53–60. doi: 10.21427/D7CF7R

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika* 30, 179–185. doi: 10.1007/BF02289447

Hu, L. T., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model.* 6, 1–55. doi: 10.1080/10705519909540118

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educ. Psychol. Meas.* 20, 141–151. doi: 10.1177/001316446002000116

Killen, M., and Smetana, J. G. (2015). "Origins and development of morality," in *Handbook of Child Psychology and Developmental Science: Socioemotional Processes*, eds M. E. Lamb and R. M. Lerner (Hoboken, NJ: Wiley), 701–749.

Kim, J. O., and Mueller, C. W. (1978). *Introduction to Factor Analysis*. Beverly Hills, CA: Sage. doi: 10.4135/9781412984652

Kline, R. B. (2013). "Exploratory and confirmatory factor analysis," in *Applied Quantitative Analysis in the Social Sciences*, eds Y. Petscher and C. Schatschneider (New York, NY: Routledge), 171–207.

Kochanska, G. (1994). Beyond cognition: expanding the search for the early roots of internalization and conscience. *Dev. Psychol.* 30, 20–22.

Kochanska, G. (1995). Children's temperament, mothers' discipline, and security of attachment: multiple pathways to emerging internalization. *Child Dev.* 66, 597–615. doi: 10.2307/1131937

Kochanska, G. (2002). Mutually responsive orientation between mothers and their young children: a context for the early development of conscience. *Curr. Dir. Psychol. Sci.* 11, 191–195. doi: 10.1111/1467-8721.00198

Kochanska, G., DeVet, K., Goldman, M., Murray, K., and Putnam, S. (1994). Maternal reports of conscience development and temperament in young children. *Child Dev.* 65, 852–869. doi: 10.2307/1131423

Kochanska, G., Forman, D. R., Aksan, N., and Dunbar, S. B. (2005). Pathways to conscience: early mother–child mutually responsive orientation and children's moral emotion, conduct, and cognition. *J. Child Psychol. Psychiatry* 46, 19–34. doi: 10.1111/j.1469-7610.2004.00348.x

Kochanska, G., Murray, K., and Coy, K. C. (1997). Inhibitory control as a contributor to conscience in childhood: from toddler to early school age. *Child Dev.* 68, 263–277. doi: 10.2307/1131849

Kohlberg, L. (1984). *Essays on Moral Development, Volume 2: The Psychology of Moral Development*. San Francisco: Harper and Row.

Lanius, R. A., Vermetten, E., and Pain, C. (2010). *The Impact of Early Life Trauma on Health and Disease: The Hidden Epidemic*. New York, NY: Cambridge University Press.

Lapsley, D.K., and Narvaez, D. (eds.) (2004b). *Moral Development, Self and Identity: Essays in Honor of Augusto Blasi*. Mahwah, NJ: Erlbaum.

Lapsley, D. K., and Narvaez, D. (2004a). "A social-cognitive view of moral character," in *Moral Development: Self and Identity*, eds D. K. Lapsley and D. Narvaez (Mahwah, NJ: Erlbaum), 189–212.

Louie, J. Y., Oh, B. J., and Lau, A. S. (2013). Cultural differences in the links between parental control and children's emotional expressivity. *Cult. Diver. Ethnic Minor. Psychol.* 19, 424–434. doi: 10.1037/a0032820

Lupien, S. J., McEwen, B. S., Gunnar, M. R., and Heim, C. (2009). Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nat. Rev. Neurosci.* 10, 434–445. doi: 10.1038/nrn2639

MacLean, P. D. (1990). *The Triune Brain in Evolution: Role in Paleocerebral Functions*. New York, NY: Plenum.

Meier, B. P., Robinson, M. D., and Wilkowski, B. M. (2006). Turning the other cheek: agreeableness and the regulation of aggression-related primes. *Psychol. Sci.* 17, 136–142. doi: 10.1111/j.1467-9280.2006.01676.x

Muthén, L. K., and Muthén, B. (2019). *Mplus. The Comprehensive Modelling Program for Applied Researchers: User's Guide* 5. Los Angeles, CA: Muthén & Muthén

Narvaez, D. (2008). Triune ethics: the neurobiological roots of our multiple moralities. *New Ideas Psychol.* 26, 95–119. doi: 10.1016/j.newideapsych.2007.07.008

Narvaez, D. (2010). "The emotional foundations of high moral intelligence," in *Children's Moral Emotions and Moral Cognition: Developmental and Educational Perspectives. New Directions for Child and Adolescent Development, vol 129*, eds B. Latzko and T. Malti (San Francisco: Jossey-Bass), 77–94.

Narvaez, D. (2013). "Neurobiology and moral mindsets," in *Moral and Immoral Behavior: Theoretical and Empirical Perspectives on Moral Motivation*, eds K. Heinrichs and F. Oser (Rotterdam: Sense Publishers), 289–307. doi: 10.1007/978-94-6209-275-4_19

Narvaez, D. (2014). *Neurobiology and the Development of Human Morality: Evolution, Culture and Wisdom*. New York, NY: W.W. Norton.

Narvaez, D. (2016). *Embodied Morality: Protectionism, Engagement and Imagination*. New York, NY: Palgrave-Macmillan.

Narvaez, D., and Hardy, S. (2016). "Measuring triune ethics orientations," in *Embodied Morality: Protectionism, Engagement and Imagination*, ed D. Narvaez (New York: Palgrave-Macmillan), 47–72.

Narvaez, D., and Lapsley, D. (2005). "The psychological foundations of everyday morality and moral expertise," in *Character Psychology and Character Education*, eds D. Lapsley and C. Power (Notre Dame, IN: University of Notre Dame Press), 140–165.

Narvaez, D., Thiel, A., Kurth, A., and Renfus, K. (2016a). "Past moral action and ethical orientation," in *Embodied Morality: Protectionism, Engagement and Imagination,* ed D. Narvaez (New York, NY: Palgrave-Macmillan), 99–118.

Narvaez, D., Wang, L., and Cheng, A. (2016b). Evolved developmental niche history: relation to adult psychopathology and morality. *Appl. Dev. Sci.* 4, 294–309. doi: 10.1080/10888691.2015.1128835

Narvaez, D., Wang, L., Gleason, T., Cheng, Y., Lefever, J., and Deng, L. (2013). The evolved developmental niche and sociomoral outcomes in Chinese 3-year-olds. *Eur. J. Dev. Psychol.* 10, 106–127.

Ng, F. F.-Y., Pomerantz, E., and Lam, S. (2007). European American and Chinese parents' responses to children's success and failure: implications for children's responses. *Dev. Psychol.* 43, 1239–1255. doi: 10.1037/0012-1649.43.5.1239

Overton, W. F. (2013). A new paradigm for developmental science: relationism and relational-developmental-systems. *Appl. Dev. Sci.* 17, 94–107. doi: 10.1080/10888691.2013.778717

Padilla-Walker, L. M., and Carlo, G. (eds.) (2014). *Prosocial Development: A Multidimensional Approach*. New York, NY: Oxford University Press.

Porges, S. (2011). *Polyvagal Theory*. New York, NY: Norton.

Putnam, S. P., and Rothbart, M. K. (2006). Development of short and very short forms of the children's behavior questionnaire. *J. Personal. Assess.* 87, 103–113. doi: 10.1207/s15327752jpa,8701_09

R Core Team (2016). *R: A LANGUAGE and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: http://www.R-project.org (accessed June 15, 2015).

Rest, J. R. (1986). *Moral Development: Advances in Research and Theory*. New York, NY: Praeger Press.

Revelle, W. (2016). p*sych: Procedures for Personality and Psychological Research (Version 1.6.12)* [R package]. Available online at: https://CRAN.R-project.org/package=psych (accessed June 15, 2015).

Rogers, C. (1961). *On Becoming a Person: A Therapist's View of Psychotherapy*. London: Constable.

Rosseel, Y. (2012). lavaan: an R package for structural equation modeling. *J. Stat. Softw.* 48, 1–36. doi: 10.18637/jss.v048.i02

Rothbart, M. K., and Bates, J. E. (1998). "Temperament," in *Handbook of Child Psychology: Vol. 3. Social, Emotional, and Personality Development,* ed N. Eisenberg (New York, NY: Wiley), 105–176.

Ryder, A. G., Yang, J., Zhu, X., Yao, S., Yi, J., Heine, S. J., and Bagby, R. M. (2008). The cultural shaping of depression: somatic symptoms in China, psychological symptoms in North America? *J. Abnorm. Psychol.* 117, 300–313. doi: 10.1037/0021-843X.117.2.300

Sapolsky, R. (2004). *Why Zebras Don't Get Ulcers, 3rd Edn*. New York, NY: Holt.

Schoenhals, M. (1993). *The Paradox of Power in a People's Republic of China Middle School*. New York, NY: Routledge.

Schore, A. N. (2019). *The Development of the Unconscious Mind*. New York, NY: W.W. Norton.

Shanker, S. (2016). *Self-Reg: How to Help Your Child (and You) Break the Stress Cycle and Successfully Engage With Life*. Toronto: Penguin Canada.

Shi, D., Lee, T., and Maydeu-Olivares, A. (2019). Understanding the model size effect on SEM fit indices. *Educ. Psychol. Meas.* 79, 310–334. doi: 10.1177/0013164418783530

Spence, S. H., Rapee, R., McDonald, C., and Ingram, M. (2001). The structure of anxiety symptoms among preschoolers. *Behav. Res. Therapy* 39, 1293–1316. doi: 10.1016/S0005-7967(00)00098-X

Tarsha, M. S., Narvaez, D., and Gleason, T. (2020). *Children's Play Is Associated with Wellbeing, Autonomic Regulation and Sociomoral Development*. Chicago, IL: Association for Psychological Science.

Triandis, H., McCusker, C., and Hui, C. (1990). Multimethod probes of individualism and collectivism. *J. Personal. Soc. Psychol.* 59, 1006–1020. doi: 10.1037/0022-3514.59.5.1006

Turiel, E. (2006). "The development of morality (revised edition)," in *Handbook of Child Psychology (Sixth Edition)*, *Volume 3,* eds W. Damon and R. M. Lerner. New York, NY: Wiley.

van der Kolk, B. (2014). *The Body Keeps the Score*. New York: Penguin.

Varela, F. (1999). *Ethical Know-How: Action, Wisdom, and Cognition*. Stanford, CA: Stanford University Press.

Varela, F. J., Thompson, E., and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

APPENDIX

Child Sociomoral Orientation Measure (CSMO)

Think of your child in SOCIAL SITUATIONS. Indicate how much your child shows the following behaviors. Response scale: I have never seen my child be this way or heard that my child acts this way.
I have seen my child be this way or heard that my child acts this way a couple of times ever.
I have seen my child be this way or heard that my child acts this way multiple times in the past.
I have seen my child be this way or heard that my child acts this way every week.
I have seen my child be this way or heard that my child acts this way every day.
I have seen my child be this way or heard that my child acts this way several times a day.

Ethic of Self-Protection.

Social opposition	Social distrust	Social withdrawal
Combative	Watchful	Timid
Easily upset	Suspicious	Withdrawing
Hostile	Untrusting	Anxious
Argumentative	Vigilant	Cowardly
Uncooperative		Fearful
Aggressive		Nervous
Fights easily		Scared
Angry		Hesitant
Threatening		Wallflower
Hot-tempered		Freezes

Ethic of Engagement.

Social enjoyment	Social attunement	Social consideration
Excited	Forgiving	Thoughtful
Laughs	Gentle	Attentive
Happy	Kind hearted	Considerate of others
Pleasant	Cuddly	Moral
Cheerful	Sympathetic	Honorable
Loving	Empathic	Respectful
Affectionate	Supportive	
Playful	Comforting	
Cheerfully interactive		

Ethic of Imagination

Social Imagination
Creative
Original
Enterprising
Thinks of new ideas
Artistic
Innovative

frontiers in Psychology

ORIGINAL RESEARCH
published: 24 November 2021
doi: 10.3389/fpsyg.2021.770891

Individual and Environmental Correlates of Adolescents' Moral Decision-Making in Moral Dilemmas

Dario Bacchini[1], Grazia De Angelis[1], Mirella Dragone[1], Concetta Esposito[1] and Gaetana Affuso[2]*

[1]Department of Humanities, University of Study of Naples Federico II, Naples, Italy, [2]Department of Psychology, University of Campania "Luigi Vanvitelli", Caserta, Italy

While extensive research has been conducted on adults' judgments in moral sacrificial dilemmas, there is little research on adolescents. The present study aimed at: (1) adding further empirical evidence about adolescents' moral decisions (deontological vs. utilitarian) in sacrificial moral dilemmas and (2) investigating how these moral decisions relate with gender, school grade, emotional traits (callous-unemotional traits), context-related experiences (perceived parental rejection and community violence exposure), and moral-related factors (moral disengagement and universalism value). A sample of 755 Italian adolescents (54.7% females; Mean age = 16.45, *SD* = 1.61) attending the second and the fifth year of secondary school took part in the study. Two sacrificial trolley-type dilemmas (where harmful actions promote the greater good) were presented. In the "switch" scenario (impersonal sacrificial dilemma), the choice is whether to hit a switch to save five people killing only one person. In the "footbridge" scenario (personal sacrificial dilemma), the choice is whether to push a large man off a footbridge saving five persons. For each scenario, participants had to indicate whether the proposed action was "morally acceptable" or not. Data were analyzed performing generalized linear mixed models. Our results showed that: (1) Adolescents were more likely to indicate as admissible to hit the switch rather than to push the large man; (2) male adolescents, compared to females, were more likely to say it was morally acceptable to intervene in the footbridge dilemma, whereas younger adolescents said it was morally acceptable both in the switch and the footbridge situations; and (3) higher levels of callous-unemotional traits, perceived parental rejection, and moral disengagement, on the one hand, and lower levels of universalism, on the other hand, were associated to higher admissibility to intervene in the footbridge scenario. Higher community violence exposure was associated with a lower propensity to intervene in the switch scenario. Overall, the present study expands the research on sacrificial dilemmas involving a sample of adolescents. The findings support previous studies concerning the role of emotions in making moral decisions but, at the same, open new perspectives regarding the role of contextual experiences and moral-related factors.

Keywords: moral judgment, trolley problem, utilitarian vs. deontological judgment, moral disengagement, community violence, values, callous-unemotional traits, parental rejection

OPEN ACCESS

Edited by:
Robert Thornberg, Linköping University, Sweden

Reviewed by:
Zijing He, Sun Yat-sen University, China
Valerio Capraro, Middlesex University, United Kingdom
Sergio Da Silva, Federal University of Santa Catarina, Brazil

****Correspondence:***
Dario Bacchini
dario.bacchini@unina.it

Specialty section:
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

Received: *05 September 2021*
Accepted: *31 October 2021*
Published: *24 November 2021*

Citation:
Bacchini D, De Angelis G, Dragone M, Esposito C and Affuso G (2021) Individual and Environmental Correlates of Adolescents' Moral Decision-Making in Moral Dilemmas. Front. Psychol. 12:770891. doi: 10.3389/fpsyg.2021.770891

INTRODUCTION

In the last two decades, psychological research on morality assigned a prominent role to the intuitionist models, according to which moral judgments are driven by automatic, unconscious, and affective reactions (Haidt, 2001; Cosmides and Tooby, 2004; Cushman et al., 2006; Hauser et al., 2007; Haidt, 2012) rather than by conscious cognitive-based processes (Kohlberg, 1968; Turiel, 1989; Turiel, 2006).

From the sixties to the 2000s, psychological research on moral development has yielded a wealth of evidence in support of the assumption that reasoning is the milestone of moral judgment. In Kohlberg's (1969) theory, morality is believed to evolve toward increasingly advanced stages accordingly to cognitive development. Social domain theorists (Turiel, 1989; Smetana, 2006, 2013) emphasized the capability of children to form judgments entailing the distinction between a moral domain, focused on impersonal and compulsory rules pertaining to welfare, justice, and rights, and a conventional domain, pertaining to not generalizable, arbitrary, shared social rules. From a different perspective, within the framework of the social cognitive theory, Bandura (1986) introduced the concept of moral agency according to which the conscious acceptance of moral standards guides individuals' beliefs and behaviors. Similarly, the theory of values (Schwartz, 1992) claimed that values are enduring goals that refer to "what people consider important" (Roccas et al., 2002) and identified ten universal values, among which "universalism," in particular, represents intrinsically moral goals guiding individuals' moral evaluations (Schwartz, 2007).

Since 2000, the assumption that moral judgment is founded on rational adherence to moral principles has been strongly challenged. An increasing number of theorists proposed an alternative approach – embedded within the evolutionary paradigm – according to which the human mind is pre-programmed to automatically react to social cues implying moral decisions. Intuitionist approaches minimized the role of cognitive processes in making moral judgments, emphasizing, on the contrary, the role of automatic and innate mechanisms. In this perspective, the role of reasoning is reduced to producing justifications, post hoc rationalizations following a pre-existing judgment (Haidt, 2001). This change of paradigm represents a challenge when shifting toward a developmental perspective, as the emphasis on innate mechanisms dramatically overshadows the role of the development. As moral decisions are triggered by automatic responses, differences between children and adults in making moral decisions seem to disappear.

In early 2000, the dual-process theory (Greene et al., 2001; Greene, 2007) proposed a synthesis between the recent intuitionist models and the more traditional approaches to morality, suggesting that both intuitive emotional responses and more controlled cognitive responses play a crucial role in moral judgment. More specifically, while emotional processes were identified as the basis for deontological judgments, cognitive processes were considered the basis of utilitarian judgments. Consistently, the so-called trolley problem, introduced in late `60s just to investigate the processes underlying utilitarian vs. deontological moral judgments, became one of the most widely used tools in the research in this theoretical framework. Utilitarian judgments can be defined as judgments endorsing actions (even harmful) that promote the greater good (Greene, 2007) and as judgments that privilege aggregate welfare over that of a small number of individuals. Deontological judgments, on the other hand, are based on an immediate feeling that a specific action could be intrinsically "wrong," irrespective of its consequences. The trolley problem and its variants (see later for a detailed description) are a prototypical dilemma where individuals have to choose whether it is permissible to sacrifice one life to save five others. Individuals answering "no" think that killing a person is intrinsically wrong, irrespective of its consequences and thus not acceptable, even if it would save the lives of several people. Individuals answering "yes" think that the consequences of any action are the focus and, thus, that sacrificing the welfare of one person can be considered right if it leads to saving the lives of several people. The first one is considered a "deontological response," in which the emphasis is on moral rules, most often articulated in terms of rights and duties; the second one is considered a "utilitarian response," in which the emphasis is on the consequences, more specifically, on maximizing benefits for the largest number of people. From a rationalist perspective, utilitarian and deontological choices would express an individual's personal philosophical perspective. Nonetheless, extensive research has provided evidence that individuals make their judgments based on specific triggers which are present in the proposed dilemmas. More specifically, the dual-process theory (Greene et al., 2001; Greene, 2007) posits that when an impersonal action is required to save five human lives sacrificing the life of one, most individuals tend to make a utilitarian choice, whereas when a personal action is required most individuals tend to make a deontological choice, beyond their personal beliefs and reasoning.

In the present research, we aimed to increase the knowledge in this field of study in two ways. Firstly, investigating adolescents' responses to trolley dilemma (and its variants) since, with few exceptions (e.g., Dahl et al., 2018), only a few studies involving individuals in the developmental age have been conducted so far. Secondly, analyzing the concurrent contribution of emotional (i.e., callous-unemotional traits), contextual (i.e., family and neighborhood), and moral-related (i.e., moral disengagement and moral values) variables to deontological vs. utilitarian judgments. Many studies investigated how emotions affected the responses to the trolley dilemma and its variants, but the samples they used included only adults. Furthermore, only a few studies have taken into account the role of family and we are not aware of any study investigating the role of neighborhood and moral cognitions with respect to the tendency to make utilitarian vs. deontological choices.

Utilitarian Vs. Deontological Approach: Individuals' Responses to Sacrificial Moral Dilemmas and Their Correlates

In Foot (1967) proposed a dilemma in which: "A runaway trolley is headed for five people who will be killed if it proceeds on its present course. The only way to save these people is to hit a switch that will turn the trolley onto a sidetrack, where it will run over and kill one person instead of five" (Greene, 2007,

pp. 41–42). This is the text of the well-known trolley dilemma, of which many variants have been proposed in the following years. One of the most famous variants was the footbridge case (Thomson, 1985), in which: "A runaway trolley threatens to kill five people, but this time you are standing next to a large stranger on a footbridge spanning the tracks, in between the oncoming trolley and the five people. The only way to save the five people is to push this stranger off the bridge and onto the tracks below. He will die as a result, but his body will stop the trolley from reaching the others" (Greene, 2007, p. 42). Respondents had to choose if it was okay to turn the trolley, in the trolley case, or push the large man onto the tracks, in the footbridge case, "in order to save five people at the expense of one" (Greene, 2007, pp. 41–42). The interesting evidence emerging from a large literature was that, while people primarily consider acceptable turning the trolley, on the contrary, they primarily consider it unacceptable pushing the large man onto the tracks (see Awad et al., 2020). In other words, differently from what it would be expected on the basis of the rationalist perspective, while people primarily exhibit a utilitarian response to the standard trolley case, they primarily exhibit a deontological response to the footbridge case. Greene and colleagues' dual-process model (Greene et al., 2001; Greene, 2007) provides an explanation of this difference positing that, depending on the characteristics of the situation, cognitive-driven or emotion-driven processes are primarily activated. More specifically, the footbridge case describes an "up-close and personal" (Greene, 2007, p. 43) situation, since the large man has to be personally pushed onto the tracks to stop the trolley, while the trolley situation, although bringing to the same consequence (the death of a person), requires just an impersonal action like hitting a switch. Therefore, the footbridge case turns out to be more morally salient and tends to evoke a dominant negative emotional response, leading to a primarily deontological response. On the contrary, the trolley situation, which is not associated with this dominant emotional response, allows more pragmatic cost-benefit analysis (Greene et al., 2001, 2004; Greene, 2007), leading to a primarily utilitarian response. Many studies investigating the link between emotional reactions to dilemmas and moral judgment supported the basic assumptions of the dual-process model (see Greene, 2014). Individuals with brain injuries altering affective reactions or individuals with a low level of affective empathy made more utilitarian judgments in sacrificial moral dilemmas (Koleva et al., 2014; Patil and Silani, 2014; Takamatsu and Takai, 2019; Dinić et al., 2021). Conversely, the examination of the role of cognitive empathy so far has provided controversial results: While some studies evaluating responses to moral everyday dilemmas evidenced a reduced tendency to make utilitarian choices in individuals with lower levels of cognitive empathy (Takamatsu, 2019), other researchers found that in sacrificial moral dilemmas the tendency to make utilitarian judgments is associated with even a selective impairment of cognitive empathy (Gleichgerrcht et al., 2013; Bacchini et al., 2018).

Other studies have focused on other possible correlates of the utilitarian choice, starting from the dimensions expected to be related to an empathic deficit, as some personality traits, like psychopathic traits and, more generally, Dark Triad traits. Research about psychopathy has evidenced that both incarcerated, clinical psychopaths (Koenigs et al., 2012; Rosas and Koenigs, 2014) and non-incarcerated, subclinical individuals with psychopathic tendencies show a preference for utilitarian solutions on emotionally aversive moral dilemmas (Glenn et al., 2010; Bartels and Pizarro, 2011; Langdon and Delmas, 2012; Gao and Tang, 2013; Djeriouat and Trémolière, 2014; Kahane et al., 2015; Patil, 2015; Balash and Falkenbach, 2018) confirming that emotionally callous personalities are more prone to endorse utilitarian judgment. However, other studies failed to find significant associations (Glenn et al., 2009; Cima et al., 2010; Pujol et al., 2012) or showed less consistent results (Gao and Tang, 2013). However, the association between dark triad traits and utilitarian judgment seems to be reduced when other personality traits, such as Honesty/Humility (Dinić et al., 2021) or moral foundations, such as Care/Harm (Djeriouat and Trémolière, 2014) were controlled. The studies evidencing the negative relation between utilitarian responses to sacrificial moral dilemmas and the endorsement of moral foundations (Koleva et al., 2014) suggested the importance to examine the role of moral values and belief in sacrificial dilemmas although we have just a little evidence about the role of ideological beliefs on utilitarian tendencies. Individuals higher on social dominance orientation and more likely to dehumanize others were more prone to utilitarian responses (Takamatsu, 2019). This last result is particularly interesting, as dehumanizing beliefs are conceptually close to one of Bandura's moral disengagement mechanisms and, although no study has investigated so far the association between these mechanisms and utilitarian tendencies, there is empirical evidence that moral disengagement is positively associated with increased unethical decision-making (Detert et al., 2008; Johnson and Connelly, 2016).

Surprisingly, there is a lack of research investigating the role of daylife contexts on deontological vs. utilitarian judgments and most of the studies have been carried out with the adults' population. Just a few studies, to our knowledge, took into consideration family variables evaluating the effects of adult attachment style (Koleva et al., 2014) and childhood adversity on adult moral decision-making (Larsen et al., 2019). Findings of these studies evidenced that avoidant attachment (Koleva et al., 2014), as well as higher levels of physical neglect during childhood (Larsen et al., 2019), were associated with greater acceptability of causing harm to utilitarian ends.

Finally, as regards the role of community context, to date, we are not aware of any study having investigated its association with moral judgment in sacrificial moral dilemmas. However, there is a great amount of evidence highlighting that growing up in violent contexts and being repeatedly exposed to the observation of violent models within the community makes youth desensitized to the effects of violence (Huesmann and Kirwil, 2007), thus disrupting their ability to empathize with other's pain and suffering and making them more prone to condone harmful actions toward others.

Utilitarian Vs. Deontological Choice in Developmental Age

Despite the wide literature on the mechanisms and psychological correlates of sacrificial dilemmas, very few studies have been

carried out with non-adult participants. Research has therefore largely disregarded the developmental perspective, ignoring that adolescence is a critical period in the consolidation of personal values, in the formation of the moral self and identity, and, in general, for the development of cognitive capabilities which could reasonably influence moral choices. A possible reason is that in a nativist perspective (e.g., Mikhail, 2000), children and adults should not differ in making moral judgments that are undergone/subjected to the same mechanisms.

A study by Pellizzoni et al. (2010) seems to confirm this assumption. The authors proposed to 3-year-old children the switch and the footbridge dilemma using Lego constructions to adapt scenarios. They found similar patterns in utilitarian vs. deontological responses of children compared to adults. Both groups preferred utilitarian choice (benefit for the greater number of people) only if the required action did not imply a personal contact: 87% of children advocated action to save five lives in the trolley dilemma, against the 27% in the footbridge dilemma. These percentages were similar to those found in adults: 91% of subjects judged the action of hitting the switch as appropriate, whereas only 31% judged it as appropriate to push the man onto the tracks. The authors concluded that in some situations, fast and automatic intuitions, based on emotional arousal, are the primary source of many moral judgments and that deliberation is used mostly to construct post hoc justifications for judgments that have already occurred (Haidt, 2001).

In line with the literature evidencing an association between utilitarian responses to sacrificial dilemmas and psychopathic/antisocial tendencies (Kahane et al., 2015), two studies investigated this association in adolescents (Bacchini et al., 2018; Dickinson and Masclet, 2019) finding results consistent with those from studies conducted with adults. More specifically, Bacchini et al. (2018) found a higher tendency to make utilitarian choices in incarcerated adolescents compared to a community control group, and the mediating role of utilitarian choice in the relationship between perspective-taking and delinquent behavior.

A study with two groups of adolescents (9th- and 12th-grade students) was realized by Stey et al. (2013) to investigate whether judgments in sacrificial dilemmas were influenced by affective considerations (Greene et al., 2001) and whether judgments of permissible harm were the product of implicit principles (Cushman et al., 2006). The authors did not find age differences between younger and older adolescents in the frequency of utilitarian vs. deontological choices, even though 12th-grade students provided significantly more sufficient justifications than 9th-grade students when asked to justify their judgments. Their conclusions agreed with Greene's point of view, since participants were more likely to use emotion words rather than refer to implicit principles, like contact and action, in their justifications, and this tendency to use emotion words in justifications was related to more deontological responses.

On the other hand, harsh criticism of Greene's model inspired the research involving adolescents carried out by Dahl et al. (2018). In their study, the authors proposed a divergent interpretation of the trolley problem and its variants, contrasting the dichotomy (emotion vs. cognition) postulated by Greene. They argued that the moral reasoning about sacrificing and saving lives involve multiple moral considerations about the value of life, that cannot simply be considered post hoc rationalizations (Haidt, 2013). In line with the literature, also Dahl et al. (2018) found that 71% of adolescents judged it permissible to activate the switch, whereas only 19% judged it permissible to push the large man onto the tracks. Nevertheless, investigating qualitative differences in participants' reasoning about the standard switch and footbridge situation, they found that in the switch dilemma people reasoned in a utilitarian way, whereas in the footbridge dilemma people's reasoning was multifaceted, implying other morally relevant issues concerning the value of life (based on the number of saved lives or the right to life of the potential victim), the natural course of events, the responsibility for consequences of actions, and the evaluation of the consequences for self. In addition, consistently with Bleske-Rechek et al. (2010), they found that choices and justifications of adolescents and adults change accordingly to other variants added to the dilemmas (e.g., the man on the bridge will be only scratched but not dead, or the victim on the sidetrack was a relative of the observer). In conclusion, Dahl et al. (2018, p. 14) considered the trolley problem (in both variants) as a dilemma evoking multifaceted conflicts which individuals try to solve according to their moral beliefs. In this regard, they affirm that "changes in evaluations about multifaceted situations reflect developmental changes in how children coordinate competing moral and non-moral considerations."

The Present Study

An extensive amount of research on sacrificial dilemmas has been produced in the last two decades. However, only a few studies have been conducted with adolescents, although adolescence is a crucial developmental period for the consolidation of moral beliefs. The present study aims to fill this gap in the literature, investigating adolescents' judgments in sacrificial dilemmas and the role of emotional-, contextual-, and moral-related variables.

The first research question was whether adolescents made more utilitarian vs. deontological choices in the switch dilemma compared to the footbridge dilemma. Consistently with the large number of studies carried out with adults and the few existing findings concerning adolescents, we expected more utilitarian responses to the switch dilemma and more deontological responses to the footbridge dilemma.

The second research question was whether adolescents' responses varied by gender and school grade, used as a proxy of adolescent age. Previous studies evidenced a higher prevalence of deontological responses among females (Fumagalli et al., 2010; Bartels and Pizarro, 2011; Friesdorf et al., 2015; Capraro and Sippel, 2017; Armstrong et al., 2019). On the other hand, no study has systematically investigated differences between younger and older adolescents; therefore, we were not able to advance specific hypotheses concerning this issue.

The third research question was whether emotional- (i.e., callous-unemotional traits), contextual- (i.e., community violence

exposure and parental rejection), and moral-related variables (i.e., moral disengagement and universalism value) could differently affect the tendency to give utilitarian vs. deontological responses across the two moral dilemmas. In this regard, according to the assumption of the dual-process theory, the effect of the aforementioned variables should be especially notable in the footbridge dilemma, given the higher emotional activation that this dilemma is expected to produce. More specifically, with respect to the callous-unemotional traits, we hypothesized that individuals with low levels of emotional activation as those with higher callous-unemotional traits were more likely to give utilitarian responses to the footbridge dilemma compared to the switch dilemma.

At the contextual level, we investigated the role of some negative experiences that adolescents could have encountered within their family context and in the neighborhood. Only two studies carried out with adults investigated the role of family-related dimensions on moral judgment in sacrificial moral dilemmas, finding that avoidant attachment (Koleva et al., 2014) and childhood adversity, such as physical neglect (Larsen et al., 2019), were associated with a higher frequency of utilitarian choices. Based on the findings of these studies, we hypothesized that higher levels of perceived parental rejection were associated with a higher tendency to give utilitarian responses in the footbridge dilemma, as parental rejection could affect emotional responsiveness in children and obstacle the internalization of moral values (Grusec et al., 2000). We also investigated the role of adolescents' exposure to community violence. Although, to our knowledge, no study has so far investigated whether experiencing violence within the neighborhood/community could affect the moral decision-making in sacrificial moral dilemmas, previous research (e.g., Dodge et al., 2006) highlighted that growing up in a violent neighborhood/community might undermine the normative process of moral development. Therefore, we speculated that, due to a process of desensitization to violence resulting from repeated experiences of exposure to community violence (Huesmann and Kirwil, 2007), youth could reduce their emotional aversion to performing even "up-close and personal" harm to others, as in the case of footbridge dilemma. Lastly, we investigated the role of two moral-related variables by considering the contribution of moral disengagement and the value of universalism. As a utilitarian solution in sacrificial moral dilemmas requires that individuals come to consider acceptable harming others for the sake of a greater good, and moral disengagement mechanisms are defined as leading individuals to disengage moral self-sanctions from their harmful practices, it is plausible to hypothesize that higher levels of moral disengagement could be associated with a higher prevalence of utilitarian responses. Moreover, since utilitarian responses to the footbridge dilemma require more sophisticated reasoning to justify the choice of sacrificing one life through direct action, we hypothesized that the more youth make use of moral disengagement mechanisms to justify their actions, the more they tend to make utilitarian choices in the footbridge dilemma.

Regarding the role of values, despite the lack of studies investigating their link with moral decision-making in sacrificial dilemmas, it seems reasonable to assume that, among others, the value of universalism which focused on the importance of preserving human life, could encourage the adoption of a deontological rather than utilitarian perspective in sacrificial moral dilemmas.

MATERIALS AND METHODS

Participants and Procedure

The sample consisted of 755 Italian adolescents (54.7% females) enrolled in grade 10 ($n = 459$, 60.8%; $M_{age} = 15.25$, $SD = 0.63$) and grade 13 ($n = 296$, 39.2%; $M_{age} = 18.27$, $SD = 0.63$) of several public schools located in the metropolitan area of Naples. The mean age of the total sample was 16.45 ($SD = 1.61$), ranging from 14 to 20 years. Although the mean age of students enrolled in the grade 13 is very close to the age of undergraduates participating in other studies, in Italy these subjects still involved in their high school careers, are usually considered representative of the adolescent age group. The socioeconomic distribution of participants' families reflected the Italian national statistics [Istituto Nazionale di Statistica (ISTAT), 2018], with most of the fathers and mothers having obtained at least a high school degree (30.9% of fathers and 34.3% of mothers) or a junior high school license (55.8% of fathers and 51.1% of mothers).

The study was approved by the Ethical Committee of the Department of Humanistic Studies, University of Naples Federico II (project identification code: 2/2020). Data were collected by trained research assistants in 2017, during regular school hours. Parents', or child guardians' written informed consent and adolescents' assents were obtained before the administration of the questionnaires. Privacy was guaranteed to participants in accordance with Italian laws 196/2003 and 101/2018. Participation in the study was voluntary, and participants could withdraw at any time without any adverse consequence.

Measures

Moral Dilemmas

Participants were presented with two scenarios involving hypothetical moral dilemmas extracted from Greene et al. (2009) and Paxton et al. (2012). The problem was presented as follows: A runaway trolley is about to run over and kill five people. In the "switch" scenario (impersonal sacrificial dilemma), one can save them by hitting a switch that will divert the trolley onto a sidetrack, where it will kill only one person. In the "footbridge" scenario (personal sacrificial dilemma), one can save them by pushing a large man off a footbridge and onto the trolley's path, killing him, but stopping the trolley. Following Greene et al. (2001) and Paxton et al. (2012) for each scenario, participants had to indicate whether the proposed action was "morally acceptable" or not. Choosing "no" (i.e., it is not morally acceptable switching tracks, or pushing the person off the bridge) can be classified as a deontological moral judgment. Choosing "yes" (i.e., it is morally acceptable switching tracks, or pushing the person off the bridge) can be classified as a utilitarian moral judgment. No/Yes answers were used in the analyses as a dichotomous dependent variable.

Moral Disengagement

Moral disengagement was measured through the moral disengagement scale developed by Caprara et al. (2006). The questionnaire specifically assesses the proneness to morally disengage with reference to different forms of detrimental conduct, in different contexts and interpersonal relationships. It consisted of 24 items that participants rated on a 5-point Likert-type scale (from 1 = "agree not at all" to 5 = "completely agree"). Sample items were as follows: "If people leave their belongings around, it is their fault if someone steals them" and "People cannot be held responsible for crimes committed at the instigation of others." Cronbach's alpha and McDonald's omega were 0.95.

Universalism

Value of universalism was self-reported by participants using the Portrait Values Questionnaire – short version (PVQ; Schwartz et al., 2001; Italian validation by Capanna et al., 2005). The PVQ – short version includes 21 verbal portraits of different people that describe a person's goals, aspirations or wishes, and point implicitly to the importance that the person attaches to a specific value. For each portrait, respondents answered the question "How much like you is this person"? using a 6-point Likert scale (from 1 = "not at all like me" to 6 = "very much like me"). For the purposes of this study, only the items measuring universalism were considered (3 items). A sample item was "He/she thinks it is important that every person in the world should be treated equally." Cronbach's alpha and McDonald's omega were 0.62.

Callous-Unemotional Traits

Callous-unemotional traits were measured using the 24-item Inventory of Callous-Unemotional Traits (ICU; Kimonis et al., 2008; Italian validation by Ciucci et al., 2014). Items of the questionnaire were scored along a 4-point Likert-type scale (from 0 = "not at all true" to 3 = "definitely true"). The factor structure of the ICU, as it has been demonstrated in several previous studies (e.g., Kimonis et al., 2008; Roose et al., 2010; Ciucci et al., 2014), consists of a general callous-unemotional factor and three subfactors: callousness (e.g., "The feelings of others are unimportant to me"), unemotional (e.g., "I hide my feelings from others"), and uncaring (e.g., "I try not to hurt others' feelings" – reversed scored item). For this study's purposes, items were averaged and used as a general callous-unemotional factor. Cronbach's alpha and McDonald's omega for the global scale were 0.81 and 0.80, respectively.

Parental Rejection

The Parental Acceptance-Rejection Questionnaire (PARQ; Rohner et al., 2005) was used to measure adolescents' perceptions of maternal and paternal rejection. Participants completed the mother version of the PARQ and then the father version. The PARQ is a 24-item self-report instrument that assesses respondents' perceptions of parental warmth, affection, hostility, aggression, indifference, neglect, and undifferentiated rejection. Items were rated on a 4-point Likert scale (from 1 = "almost never true" to 4 = "almost always true"). Sample items were as follows: "My [mother/father] makes me feel wanted and needed"; "My [mother/father] goes out of [her/his] way to hurt my feelings." Scores for each subscale were averaged to compute global scores of maternal and paternal rejection, with high values indicating high rejection. The two scores, one referring to the mother and the other one referring to the father, were then averaged to create a composite score of parental rejection, which demonstrated high reliability (Cronbach's alpha = 0.95; McDonald's omega = 0.94).

Exposure to Community Violence

Exposure to community violence was assessed using the witnessing subscale of the Exposure to Community Violence Questionnaire (Esposito et al., 2017), consisting of 6 items. Adolescents were asked to report the frequency with which they have been witnessed violent incidents that had occurred during the last year in their neighborhood using a scale ranging from 1 ("never") to 5 ("more than five times"). A sample item was "How many times have you seen somebody get robbed"? Scores for each item were averaged to create the score for community violence exposure. Reliability statistics were adequate, Cronbach's alpha and McDonald's omega = 0.88.

Statistical Analysis

Before testing our hypotheses, the univariate normality of data distribution was tested, finding that no study's variables approached skewness > |3| or kurtosis > |10|.

Then, we firstly identified within the sample participants who made deontological vs. utilitarian judgments to impersonal (i.e., switch dilemma) and personal (i.e., footbridge dilemma) moral scenarios (Research question 1) and compared them by gender (males vs. females) and school grade (10th vs. 13th graders; Research question 2) using a set of chi-square statistics performed in IBM SPSS 21 (IBM, Armonk, NY, United States).

The third research question (Research question 3, i.e., correlates of moral choices) was then examined using generalized linear mixed models with binomial family and logit link function, performed with JASP statistical software (JASP Team, 2020). The dependent variable was a dichotomous variable indicating whether the action required in the dilemmas was considered morally acceptable or not (no vs. yes). Models included random intercepts for participants and fixed effects of the variables considered in the study as potential predictors. More specifically, three separate generalized linear mixed models were performed. The first one tested the effects of emotional traits (namely, callous-unemotional traits), controlling for gender, school grade, and type of moral scenario (switch vs. footbridge). Then, the two-way interaction between emotional traits and type of scenario was included as a second step. The second and third models examined the effects of contextual factors (exposure to community violence and parental rejection) and moral-related variables (moral disengagement and universalism), respectively. Also in these cases, the effects of gender and school grade were controlled, and interactions with the moral scenario were tested. Continuous variables were mean-centered

before running the analyses. Model terms were tested with the likelihood ratio tests method.

RESULTS

Deontological Vs. Utilitarian Responses: Differences by Type of Scenario, Gender, and School Grade

In order to examine whether adolescents' judgments differed according to the type (personal vs. impersonal) of moral scenarios (Research question 1) and whether there were gender- and school grade-related differences (Research question 2), a set of chi-square statistics was performed. Results revealed that there were significant differences [χ^2 (1) = 84.31; $p < 0.001$] in the frequency of deontological (vs. utilitarian) responses to the switch dilemma (268 subjects; 35.5%; $M = 109$, $M_{age} = 16.60$ years, $SD = 1.56$) compared to the footbridge dilemma (543 subjects; 71.9%; $M = 221$, $M_{age} = 16.49$ years, $SD = 1.60$). Moreover, significant differences by gender [χ^2 (1) = 16.50; $p < 0.001$] emerged only when youth were faced with the footbridge dilemma, with 121 males out of 342 (35.4%) making utilitarian judgments compared to 91 females out of 413 (22%). Conversely, no significant gender difference emerged when subjects were presented with the switch dilemma, with 233 males out of 342 (68.1%) making utilitarian judgments compared to 254 females out of 413 (61.5%). Finally, significant school grade differences emerged when youth were faced with both the switch [χ^2 (1) = 4.06; $p < 0.05$] and the footbridge dilemma [χ^2 (1) = 4.73; $p < 0.05$], with 10th-grade participants more likely to make utilitarian judgments than 13th-grade participants (67.3% vs. 60.1 and 30.9% vs. 23.6%, for switch and footbridge dilemma, respectively).

Generalized Linear Mixed Models

The results of the generalized linear mixed models (Research question 3) are displayed in **Tables 1**–**3** and described in the following sections.

The Effect of Emotional Traits

The results showed a significant interaction effect between callous-unemotional traits and the type of moral dilemma (**Table 1**). The analysis of simple slopes (**Figure 1A**) indicated that those who reported higher levels of callous-unemotional traits were more likely to rate the intervention in the footbridge scenario as permissible, $B = 0.75$, $SE = 0.29$, 95% CI [0.18, 1.32], whereas no significant effect was found in the switch situation, $B = -0.14$, $SE = 0.27$, 95% CI [−0.67, 0.39].

The Effects of Contextual Factors

The examination of the effects of parental rejection and community violence witnessing revealed significant interaction effects with the type of moral dilemma (**Table 2**). More specifically, parental rejection was found to have a significant positive effect only in the footbridge scenario, $B = 0.66$, SE = 0.24, 95% CI [0.20, 1.12], whereas no significant effect emerged in the switch scenario, $B = -0.03$, SE = 0.23, 95% CI [−0.48, 0.41] (**Figure 1B**). Conversely, high levels of exposure to community violence negatively predicted the ratings of the intervention as permissible in the switch situation, $B = -0.31$, SE = 0.14, 95% CI [−0.58, −0.04], whereas no significant effect was found in the footbridge scenario, $B = 0.12$, SE = 0.14, 95% CI [−0.16, 0.39] (**Figure 1C**).

TABLE 1 | Generalized linear mixed model (1) – Emotional traits predicting No/Yes answers in moral dilemmas.

Terms	*B*	*SE*	*t*	*p*
Intercept	−0.27	0.09	−2.84	0.005
Type of moral scenario (Switch)	1.18	0.10	11.98	<0.001
Gender (Male)	0.30	0.09	3.22	0.001
School Grade (10th)	0.20	0.09	2.11	0.035
Callous-unemotional traits	0.28	0.22	1.27	0.203
Model with interactions				
Intercept	−0.28	0.10	−2.91	0.004
Type of moral scenario (Switch)	1.21	0.10	11.85	<0.001
Gender (Male)	0.31	0.10	3.22	0.001
School Grade (10th)	0.20	0.10	2.14	0.033
Callous-unemotional traits	0.31	0.22	1.37	0.170
Type of moral scenario * Callous-unemotional traits	−0.44	0.17	−2.57	0.010

The Effects of Moral-Related Variables

Moral disengagement and universalism were found to be significant predictors of participants' ratings of the intervention as permissible, conditional on the type of moral dilemma (interaction effects: $ps < 0.001$ and 0.05, for moral disengagement and universalism, respectively; see **Table 3**). The simple slopes analysis revealed that both had a significant effect only in the footbridge situation. Moral disengagement was associated with an increased likelihood to rate the intervention as permissible, $B = 0.57$, SE = 0.17, 95% CI [0.23, 0.90] (**Figure 1D**), whereas high levels of universalism negatively predicted the permissibility of the intervention, $B = -0.26$, SE = 0.12, 95% CI [−0.50, −0.01] (**Figure 1E**).

DISCUSSION

In the last two decades, a tremendous amount of research has been carried out using the "trolley dilemma" in which participants are faced with the choice of whether or not it is permissible to sacrifice one human life to save five others. Such sacrificial dilemma (and its variants, e.g., the footbridge dilemma) has become a prototypical tool of investigation because it would allow to reveal two competing mechanisms implicated in making moral judgments: a cognitive-driven process (addressing utilitarian choices, i.e., based on an evaluation of cost-benefit ratio, it is morally acceptable to kill one person in order to save five others) as opposed to an emotion-driven process (addressing deontological choices, i.e., based on an immediate, automatic and unconscious feeling that it is not morally acceptable to kill one person in order to save five others). According to the dual-process theory of moral judgment

TABLE 2 | Generalized linear mixed model (2) – Contextual-related factors predicting No/Yes answers in moral dilemmas.

Terms	*B*	*SE*	*t*	*p*
Intercept	−0.27	0.09	−2.89	0.004
Type of moral scenario (Switch)	1.18	0.10	11.98	<0.001
Gender (Male)	0.31	0.09	3.32	<0.001
School Grade (10th)	0.22	0.09	2.37	0.018
Parental rejection	0.38	0.18	2.07	0.039
Community violence	−0.12	0.11	−1.10	0.273
Model with interactions				
Intercept	−0.32	0.10	−3.25	0.001
Type of moral scenario (Switch)	1.20	0.10	11.70	<0.001
Gender (Male)	0.31	0.10	3.26	0.001
School Grade (10th)	0.22	0.10	2.34	0.019
Parental rejection	0.32	0.18	1.73	0.085
Community violence	−0.10	0.11	−0.86	0.393
Type of moral scenario * Parental rejection	−0.35	0.14	−2.46	0.014
Type of moral scenario * Community violence	−0.21	0.08	−2.56	0.010

TABLE 3 | Generalized linear mixed model (3) – Moral-related variables predicting No/Yes answers in moral dilemmas.

Terms	*B*	*SE*	*t*	*p*
Intercept	−0.27	0.09	−2.83	0.005
Type of moral scenario (Switch)	1.18	0.10	11.98	<0.001
Gender (Male)	0.28	0.10	2.93	0.003
School Grade (10th)	0.19	0.10	1.99	0.046
Moral disengagement	0.16	0.13	1.28	0.200
Universalism value	−0.08	0.09	−0.82	0.412
Model with interactions				
Intercept	−0.30	0.10	−3.01	0.003
Type of moral scenario (Switch)	1.26	0.11	11.52	<0.001
Gender (Male)	0.30	0.10	2.93	0.003
School Grade (10th)	0.21	0.10	2.08	0.037
Moral disengagement	0.18	0.13	1.37	0.172
Universalism value	−0.09	0.10	−0.92	0.356
Type of moral scenario * Moral disengagement	−0.38	0.10	−3.81	<0.001
Type of moral scenario * Universalism value	0.17	0.07	2.30	0.021

(Greene et al., 2001; Greene, 2007), while utilitarian choices are activated when individuals are faced with a low-impact dilemma (i.e., "impersonal" trolley dilemma) that elicits lower levels of emotional distress, deontological choices are activated when the proposed high-impact dilemma (i.e., "personal" footbridge dilemma) elicits higher levels of emotional distress.

The present study aimed to increase the knowledge about the processes underlying moral decision-making into sacrificial dilemmas in two ways: (1) by exploring the adolescents' responses to trolley dilemma across different, "personal" (i.e., the switch scenario) and "impersonal" (i.e., the footbridge scenario) variants since, to date, only a few studies (e.g., Dahl et al., 2018) on this topic have involved non-adults populations; (2) by investigating the concurrent contribution of gender, age, emotional-, contextual- (i.e., family and neighborhood), and moral-related variables in making moral judgments (deontological vs. utilitarian) in both switch and footbridge dilemmas.

Consistent with the literature, in our study, we found that adolescents' choices in sacrificial moral dilemmas significantly varied according to the type (impersonal vs. personal) of scenario, with the majority of youths more prone to the utilitarian choice in the switch dilemma and, on the contrary, more likely to choose the deontological solution in the footbridge dilemma (Research question 1). These findings are in line with our expectations and with previous studies showing similar patterns of responses allowing harm to others to utilitarian ends more in impersonal than in personal dilemmas, among 3-year-old children (Pellizzoni et al., 2010), adolescents (Dahl et al., 2018), and adults from different countries (Awad et al., 2020). Moreover, we found that younger participants were more prone to utilitarian responses, irrespective of the type of dilemma, and that there were gender differences depending on the type of dilemma, with males more willing to choose utilitarian solutions only in footbridge dilemma (Research question 2). These results are consistent with those emerged in previous studies in which males showed a stronger preference for utilitarian over deontological judgments (e.g., Fumagalli et al., 2010; Bartels and Pizarro, 2011; Friesdorf et al., 2015), particularly when considering "personal" moral dilemmas (Fumagalli et al., 2010) where harm requires physical force (Greene et al., 2009). Moreover, our results seem to make sense considering the wide research highlighting higher emotional responsiveness among females (Eisenberg, 2005), which could lead to giving more automatic and immediate responses evidencing aversion to causing harm to others in the context of moral dilemmas. On the other hand, males' moral evaluation is believed to be more pragmatic and adhering to abstract principles of justice (Jaffee and Hyde, 2000; see also the classical debate Kohlberg vs. Gilligan), although a recent meta-analysis (Friesdorf et al., 2015) suggested that gender differences in the preferences for utilitarian vs. deontological judgments stem from gender differences in affective reactions to causing harm rather than in cognitive evaluations of outcomes. Results regarding age-related differences are something new, as the present study is, to our knowledge, the first one to systematically investigate differences between younger and older adolescents with respect to moral judgments in sacrificial moral dilemmas. The finding that younger adolescents are more prone to the utilitarian solution, irrespective of the type of dilemma would suggest that the preference for the deontological vs. utilitarian solution has more to do with developmental changes than with the characteristics of the proposed scenario. However, younger adolescents tend to make, as well as older adolescents, more utilitarian choices in the switch dilemma than in the footbridge dilemma, in accordance with the expectation of Greene's dual-process model.

With respect to our third research question, which concerned the investigation of moral choices' correlates, we performed a

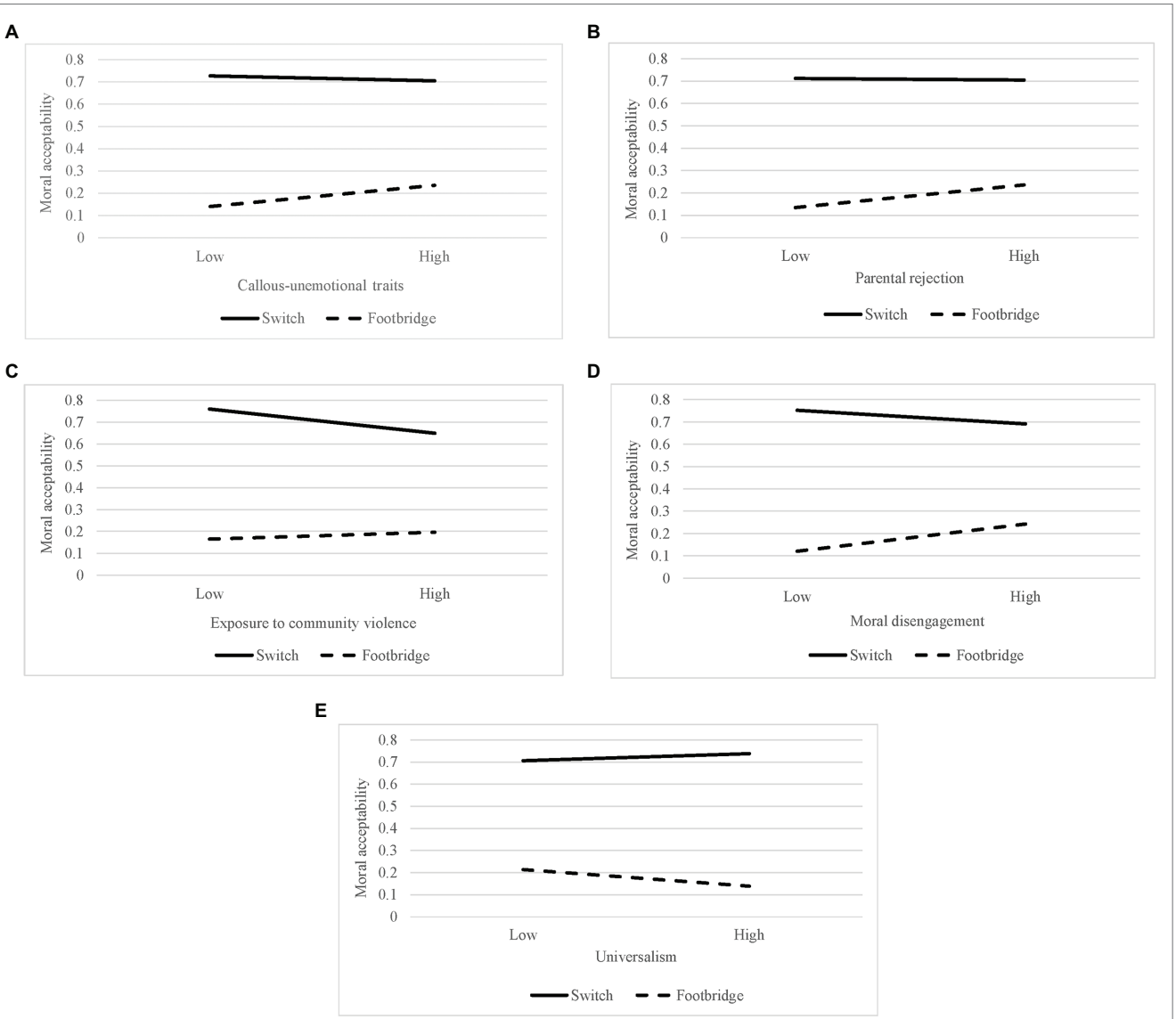


FIGURE 1 | Plots of the effects of emotional traits **(A)**, context-related factors **(B, C)**, and moral-related variables **(D, E)** depending on the type of moral scenario (switch vs. footbridge).

set of generalized linear mixed models in order to test the contribution of emotional- (i.e., callous-unemotional traits), contextual- (i.e., parental rejection and community violence witnessing), and moral-related (i.e., moral disengagement and universalism) variables in making moral judgments (i.e., deontological vs. utilitarian) in sacrificial dilemmas. Both the main effects of each variable (i.e., without take into account the type of dilemmas) and the interaction effects (i.e., testing whether the contribution of each variable varied depending on the type of dilemmas) were tested. Our first result was that adolescents higher on callous-unemotional trait were more likely to choose the utilitarian solution (i.e., push the large man onto the tracks to save five other people) in the footbridge case, while we did not find any difference in the switch case. This finding was in line with our hypotheses and with the literature evidencing that clinical psychopaths (Koenigs et al., 2012; Rosas and Koenigs, 2014), as well as subclinical individuals with psychopathic tendencies, (Glenn et al., 2010; Bartels and Pizarro, 2011; Langdon and Delmas, 2012; Gao and Tang, 2013; Djeriouat and Trémolière, 2014; Kahane et al., 2015; Patil, 2015; Balash and Falkenbach, 2018) are more willing to accept utilitarian solution when facing emotionally aversive moral dilemmas. Callous-unemotional traits, which can be considered the hallmark of the psychopathic personality (Blair, 2013), are characterized, indeed, by general disregard for others, lack of empathy and, more in general, deficient emotional activation. Therefore, it is not surprising that individuals higher on these personality traits are less responsive to the moral salience of a personal moral dilemma, as the footbridge case. Moreover, this conclusion is in line with studies evidencing a higher tendency to make

utilitarian choices in subjects with low levels of affective empathy (Koleva et al., 2014; Patil and Silani, 2014; Takamatsu and Takai, 2019; Dinić et al., 2021) or in a clinical population with brain injuries altering affective reactions (see Greene, 2014).

Then, we took into consideration contextual factors, investigating the role of two negative experiences that adolescents could have encountered within their family context and in the neighborhood: perceived parental rejection and exposure to community violence as a witness. In the present study, the two contextual dimensions showed divergent interactions with the type of dilemma. Indeed, adolescents perceiving higher parental rejection are more prone to the utilitarian choice in the footbridge dilemma, while we did not find any difference in the switch dilemma. Conversely, adolescents who are more often witnesses of violence in their neighborhood are less prone to the utilitarian solution in the switch scenario, while we did not find any difference in the footbridge situation. Although only two studies involving adults and focusing on different variables (attachment style and childhood adversity, such as physical neglect during childhood) investigated the role of family-related dimensions on moral judgment in sacrificial moral dilemmas (Koleva et al., 2014; Larsen et al., 2019), their results evidenced that dysfunctional relationships (avoidant attachment and higher physical neglect) within the family context can promote utilitarian tendencies. Moreover, there is evidence in the literature that parental rejection affects emotional responsiveness in children and hinder the normal process of internalization of moral values (Grusec et al., 2000), leading to various maladaptive outcomes including internalizing and externalizing symptoms. Therefore, parental rejection could work, in line with the basic assumption of Greene's theory, as another variable influencing, at a contextual level, the emotional and cognitive processes involved in moral decision-making. Moreover, the reduced moral responsiveness could at least in part explain why we found differences only when adolescents considered the footbridge case. With respect to exposure to community violence, to our knowledge, there is no study in the literature evaluating the association of this contextual variable with utilitarian vs. deontological choice in sacrificial moral dilemmas. However, there is a great amount of evidence highlighting the negative effects of community violence exposure on moral development. The research found that children and adolescents exposed to community violence, fail in distinguishing moral vs. conventional issues (Bacchini et al., 2013), tend to make frequent recourse to self-serving cognitive distortions (Dragone et al., 2020; Esposito et al., 2020), judge morally acceptable physically harming others in contexts of survival or revenge (Posada and Wainryb, 2008), and are more likely to condone moral transgressions when provoked or for reasons of retaliation (Ardila-Rey et al., 2009). Overall, exposure to violence has been found to substantially disrupt the moral decision-making ability as a result of impairments of several emotional (e.g., empathy), cognitive (e.g., theory of mind), and inhibitory control abilities (Zucchelli and Ugazio, 2019).

Growing up in violent communities could exert a detrimental effect on normative moral development, leading to a decreased sensitivity toward violence (Dodge et al., 2006; Huesmann and Kirwil, 2007). Therefore, harmful behaviors end up becoming normative and could also result in a sort of indifference with other's pain and suffering and, at the same time, a sort of learned helplessness that induces individuals not to interfere with the natural course of events, just like our adolescents.

Finally, we evaluated the possible role of two moral-related variables: moral disengagement and universalism, the basic human value representing the intrinsically moral goal of preserving the welfare of others. Our results showed that adolescents more prone to make use of moral disengagement mechanisms were more likely to choose utilitarian solutions in the footbridge situation, while we did not find any difference in the switch situation. Conversely, adolescents higher on universalism proved to give fewer utilitarian responses in the footbridge case, while no difference emerged in the switch case.

Although there is no study, to our knowledge, in the literature investigating the role of moral disengagement with respect to the tendency to give utilitarian vs. deontological responses in sacrificial moral dilemmas, there is little evidence of the role of beliefs in making moral judgments. In particular, as shown by Takamatsu (2019), individuals higher on social dominance orientation and, even more interesting for the present study, individuals more likely to dehumanize others were more prone to utilitarian responses. As dehumanizing beliefs are conceptually close to one of Bandura's moral disengagement mechanisms, this evidence in the literature seems to support our results. Moreover, our findings seem to make sense, considering that moral disengagement mechanisms are defined as allowing individuals to disengage moral self-sanctions from their harmful practices and utilitarian choice in sacrificial moral dilemmas, in particular in personal scenarios such as the footbridge case, requires individuals to consider acceptable harming others for the sake of a greater good. With respect to the role of universalism, although there is no study in the literature investigating the association with moral decision-making in sacrificial dilemmas, it seems to make sense that a higher endorsement of values focused on preserving human life could lead to a higher tendency to deontological choices, in particular in personal dilemmas, in which the higher moral salience of the proposed scenario makes more difficult for individuals to set aside their values. Moreover, our result receives support from the studies evidencing an association between a reduced tendency to prefer utilitarian solutions in sacrificial moral dilemmas and a higher endorsement of moral foundations, in particular of the Care/Harm foundation, underlying aversion to harmful actions (Djeriouat and Trémolière, 2014; Koleva et al., 2014; Crone and Laham, 2015) and conceptually close to universalism. Finally, consistent with our findings, Dahl et al. (2018), investigating the qualitative differences in participants' reasoning about the switch and the footbridge situation, found that unlike what happens in the switch case, in the footbridge case reasoning was more multifaceted, involving different moral considerations primarily associated with the value of life.

Strengths, Limitations, and Future Perspectives

One of the major strengths of the present study is the focus on adolescence, largely disregarded in the research investigating so far moral decision-making in sacrificial moral dilemmas,

despite its critical role for moral development. Focusing on adolescents, moreover comparing two age cohorts, allow us to understand utilitarian vs. deontological choices from a developmental perspective, evidencing whether and how moral judgments change over time. Furthermore, investigating the role of a broad range of possible correlates at the individual and contextual level and allows us to get a clearer picture of which features are more relevant with respect to the development of utilitarian vs. deontological inclinations.

On the other hand, our understanding of the developmental aspects of moral decision-making in sacrificial dilemmas is limited by the cross-sectional nature of the study, not allowing us to determine causal influences. More research involving longitudinal samples is needed to confirm and deepen our results. Another limitation is related to the use of sacrificial dilemmas as a measure of utilitarian vs. deontological inclination. Such measure is the subject of an ongoing scientific debate discussing the idea that this methodology has different weaknesses, such as treating utilitarian and deontological responses as inversely related (Conway and Gawronski, 2013), lack of manipulation of consequences and norms, that are the defining aspects of utilitarianism and deontology (see Gawronski and Beer, 2017), and directly measuring only a negative dimension of utilitarianism, called "instrumental harm" (Kahane et al., 2018). However, despite the proposal of adapting the traditional dilemma methodology (Conway and Gawronski, 2013; Gawronski et al., 2017) or replacing it with new measures (Kahane et al., 2018), the use of sacrificial moral dilemmas remains widely accepted (see Conway et al., 2018; Kahane et al., 2018), with the recommendation of taking in mind the limitations that were evidenced and that should be therefore extended to our results. In particular, our findings evidencing adolescents higher on different variables associated with maladaptation (higher callousness, more parental rejection) are more willing to sacrifice the man in the footbridge situation and seem to suggest that utilitarian choices are more likely to stem from a decreased aversion to harming others, rather than reflect a genuine concern for the greater good. Accordingly, the finding that younger adolescents tend to give more utilitarian responses, irrespective to the proposed scenario, could be read as in part reflecting the incomplete maturation of nervous system, resulting in a still incomplete development of executive. functions However, the criticisms we have mentioned above require caution, in view of the difficulty to ascertain the underlying mechanism of the adolescents' choices. More research is needed to shed light on this issue.

Moreover, future research should clarify how aversive experiences within the family or community exert their influence on adolescents' moral decision-making in sacrificial dilemmas and, even more basically, which are the main characteristics of the adolescents choosing a deontological solution in the switch case. It would be also interesting deepen the role of cognitive variables, since previous research has evidenced that the reversal of moral preferences that can be observed when individuals face different moral scenarios, as in the switch and footbridge case, may occur because utilitarian moral judgments are cognitively too demanding (Da Silva et al., 2016). Finally, future studies could be useful to deepen the practical implications of results regarding the utilitarian vs. deontological choices and to explore the possible use of sacrificial dilemma as a tool to increase moral skills (Seider, 2009).

DATA AVAILABILITY STATEMENT

The datasets for this study can be found in the **Supplementary Material**.

ETHICS STATEMENT

The study was reviewed and approved by the Ethical Committee of the Department of Humanties, University of Naples "Federico II" (project identification code: 2/2020). Written informed consent to participate in this study was provided by the participants' parents, or legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

DB, GDA, and MD contributed to conception and design of the study. GDA organized the database. GDA, MD, and CE performed the statistical analysis and wrote sections of the manuscript. DB wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.770891/full#supplementary-material

REFERENCES

Ardila-Rey, A., Killen, M., and Brenick, A. (2009). Moral reasoning in violent contexts: displaced and non-displaced Colombian children's evaluations of moral transgressions, retaliation, and reconciliation. *Soc. Dev.* 18, 181–209. doi: 10.1111/j.1467-9507.2008.00483.x

Armstrong, J., Friesdorf, R., and Conway, P. (2019). Clarifying Gender Differences in Moral Dilemma Judgments: The Complementary Roles of Harm Aversion and Action Aversion. *Soc. Psychol. Pers. Sci.* 10, 353–363. doi: 10.1177/1948550618755873

Awad, E., Dsouza, S., Shariff, A., Rahwan, I., and Bonnefon, J. F. (2020). Universals and variations in moral decisions made in 42 countries by 70,000 participants. *Proc. Natl. Acad. Sci.* 117, 2332–2337. doi: 10.1073/pnas.1911517117

Bacchini, D., Affuso, G., and De Angelis, G. (2013). Moral vs. non-moral attribution in adolescence: environmental and behavioural correlates. *Eur. J. Dev. Psychol.* 10, 221–238. doi: 10.1080/17405629.2012.744744

Bacchini, D., De Angelis, G., and Dragone, M. (2018). Empathy and utilitarian judgment in incarcerated and community control adolescents. *Crim. Justice Behav.* 45, 1093–1109. doi: 10.1177/0093854818781438

Balash, J., and Falkenbach, D. M. (2018). The ends justify the meanness: An investigation of psychopathic traits and utilitarian moral endorsement. *Pers. Individ. Differ.* 127, 127–132. doi: 10.1016/j.paid.2018.02.009

Bandura, A. (1986). *Social Foundations of Thought and Action*. Englewood Cliffs, NJ: Prentice Hall.

Bartels, D. M., and Pizarro, D. A. (2011). The mismeasure of morals: antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition* 121, 154–161. doi: 10.1016/j.cognition.2011.05.010

Blair, R. J. R. (2013). Psychopathy: cognitive and neural dysfunction. *Dialogues Clin. Neurosci.* 15, 181–190. doi: 10.31887/2FDCNS.2013.15.2/2Frblair

Bleske-Rechek, A., Nelson, L. A., Baker, J. P., Remiker, M. W., and Brandt, S. J. (2010). Evolution and the trolley problem: people save five over one unless the one is young, genetically related, or a romantic partner. *J. Soc. Evol. Cult. Psychol.* 4:115. doi: 10.1037/h0099295

Capanna, C., Vecchione, M., and Schwartz, S. H. (2005). La misura dei valori. Un contributo alla validazione del Portrait Value Questionnaire su un campione italiano [The measurement of values. A contribution to the validation of the Portrait Values Questionnaire on an Italian sample]. *Boll. di Psicol. Appl.* 246, 29–41.

Caprara, G. V., Barbaranelli, C., Pastorelli, C., Iafrate, C., Beretta, M., Steca, P., et al. (2006). La misura del disimpegno morale nel contesto delle trasgressioni dell'agire quotidiano [The measure of the moral disengagement in the context of the transgressions of daily living.]. *G. Ital. di Psicol.* 33, 83–106. doi: 10.1421/21961

Capraro, V., and Sippel, J. (2017). Gender differences in moral judgment and the evaluation of gender-specified moral agents. *Cogn. Process.* 18, 399–405. doi: 10.1007/s10339-017-0822-9

Cima, M., Tonnaer, F., and Hauser, M. D. (2010). Psychopaths know right from wrong but don't care. *Soc. Cogn. Affect. Neurosci.* 5, 59–67. doi: 10.1093/scan/nsp051

Ciucci, E., Baroncelli, A., Franchi, M., Golmaryami, F. N., and Frick, P. J. (2014). The association between callous-unemotional traits and behavioral and academic adjustment in children: further validation of the inventory of callous-unemotional traits. *J. Psychopathol. Behav. Assess.* 36, 189–200. doi: 10.1007/s10862-013-9384-z

Conway, P., and Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision-making: A process dissociation approach. *J. Pers. Soc. Psychol.* 104, 216–235. doi: 10.1037/a0031021

Conway, P., Goldstein-Greenwood, J., Polacek, D., and Greene, J. D. (2018). Sacrificial utilitarian judgments do reflect concern for the greater good: clarification via process dissociation and the judgments of philosophers. *Cognition* 179, 241–265. doi: 10.1016/j.cognition.2018.04.018

Cosmides, L., and Tooby, J. (2004). "Knowing thyself: The evolutionary psychology of moral reasoning and moral sentiments," in *Business, Science, and Ethics. The Ruffin Series in Business Ethics. Vol. 1*. eds. R. Freeman and P. H. Werhane (Charlottesville, VA: Society for Business Ethics), 93–128.

Crone, D. L., and Laham, S. M. (2015). Multiple moral foundations predict responses to sacrificial dilemmas. *Pers. Individ. Differ.* 85, 60–65. doi: 10.1016/j.paid.2015.04.041

Cushman, F., Young, L., and Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment: testing three principles of harm. *Psychol. Sci.* 17, 1082–1089. doi: 10.1111/j.1467-9280.2006.01834.x

Da Silva, S., Matsushita, R., and De Sousa, M. (2016). Utilitarian Moral Judgments Are Cognitively Too Demanding. *OALib Journal.* 3, 1–9. doi: 10.4236/OALIB.1102380

Dahl, A., Gingo, M., Uttich, K., and Turiel, E. (2018). Moral reasoning about human welfare in adolescents and adults: judging conflicts involving sacrificing and saving lives. *Monogr. Soc. Res. Child Dev.* 83, 1–109. doi: 10.1111/mono.12374

Detert, J. R., Treviño, L. K., and Sweitzer, V. L. (2008). Moral disengagement in ethical decision making: a study of antecedents and outcomes. *Int. J. Appl. Psychol.* 93:374. doi: 10.1037/0021-9010.93.2.374

Dickinson, D. L., and Masclet, D. (2019). Using ethical dilemmas to predict antisocial choices with real payoff consequences: An experimental study. *J. Econ. Behav. Organ.* 166, 195–215. doi: 10.1016/j.jebo.2019.08.023

Dinić, B. M., Milosavljević, M., and Mandarić, J. (2021). Effects of dark tetrad traits on utilitarian moral judgement: The role of personal involvement and familiarity with the victim. *Asian J. Soc. Psychol.* 24, 48–58. doi: 10.1111/ajsp.12422

Djeriouat, H., and Trémolière, B. (2014). The dark triad of personality and utilitarian moral judgment: The mediating role of honesty/humility and harm/care. *Pers. Individ. Differ.* 67, 11–16. doi: 10.1016/j.paid.2013.12.026

Dodge, K. A., Coie, J. D., and Lynam, D. (2006). "Aggression and antisocial behavior in youth," in *Handbook of Child Psychology: Social, Emotional, and Personality Development. 6th Edn*. eds. N. Eisenberg, W. Damon and R. M. Lerner (Hoboken, NJ: John Wiley), 719–788.

Dragone, M., Esposito, C., De Angelis, G., Affuso, G., and Bacchini, D. (2020). Pathways linking exposure to community violence, self-serving cognitive distortions and school bullying perpetration: A three-wave study. *Int. J. Environ. Res. Public Health* 17:188. doi: 10.3390/ijerph17010188

Eisenberg, N. (2005). "The development of empathy-related responding," in *Moral Motivation through the Life Span*. eds. G. Carlo and C. P. Edwards (Lincoln, NE: University of Nebraska Press), 73–117.

Esposito, C., Affuso, G., Dragone, M., and Bacchini, D. (2020). Effortful control and community violence exposure as predictors of developmental trajectories of self-serving cognitive distortions in adolescence: a growth mixture modeling approach. *J. Youth Adolesc.* 49, 2358–2371. doi: 10.1007/s10964-020-01306-x

Esposito, C., Bacchini, D., Eisenberg, N., and Affuso, G. (2017). Effortful control, exposure to community violence, and aggressive behavior: exploring cross-lagged relations in adolescence. *Aggress. Behav.* 43, 588–600. doi: 10.1002/ab.21717

Foot, P. (1967). The problem of abortion and the doctrine of the double effect. *Oxford Rev.* 5, 5–15. doi: 10.1093/0199252866.003.0002

Friesdorf, R., Conway, P., and Gawronski, B. (2015). Gender differences in responses to moral dilemmas: a process dissociation analysis. *Personal. Soc. Psychol. Bull.* 41, 696–713. doi: 10.1177/0146167215575731

Fumagalli, M., Ferrucci, R., Mameli, F., Marceglia, S., Mrakic-Sposta, S., Zago, S., et al. (2010). Gender-related differences in moral judgments. *Cogn. Process.* 11, 219–226. doi: 10.1007/s10339-009-0335-2

Gao, Y., and Tang, S. (2013). Psychopathic personality and utilitarian moral judgment in college students. *J. Crim. Justice.* 41, 342–349. doi: 10.1016/j.jcrimjus.2013.06.012

Gawronski, B., Armstrong, J., Conway, P., Friesdorf, R., and Hütter, M. (2017). Consequences, norms, and generalized inaction in moral dilemmas: The CNI model of moral decision-making. *J. Pers. Soc. Psychol.* 113, 343–376. doi: 10.1037/pspa0000086

Gawronski, B., and Beer, J. S. (2017). What makes moral dilemma judgments "utilitarian" or "deontological"? *Soc. Neurosci.* 12, 626–632. doi: 10.1080/17470919.2016.1248787

Gleichgerrcht, E., Torralva, T., Rattazzi, A., Marenco, V., Roca, M., and Manes, F. (2013). Selective impairment of cognitive empathy for moral judgment in adults with high functioning autism. *Soc. Cogn. Affect. Neurosci.* 8, 780–788. doi: 10.1093/scan/nss067

Glenn, A. L., Koleva, S., Iyer, R., Graham, J., and Ditto, P. H. (2010). Moral identity in psychopathy. *Judgm. Decis. Mak.* 5, 497–505.

Glenn, A. L., Raine, A., and Schug, R. A. (2009). The neural correlates of moral decision-making in psychopathy. *Mol. Psychiatry* 14, 5–6. doi: 10.1038/mp.2008.104

Greene, J. D. (2007). "The secret joke of Kant's soul," in *Moral psychology, Vol. 3: The neuroscience of morality: Emotion, brain disorders, and development*. ed. W. Sinnott-Armstrong (Cambridge, MA: The MIT Press), 35–79.

Greene, J. D. (2014). "The cognitive neuroscience of moral judgment and decision making," in *The cognitive neurosciences*. eds. M. S. Gazzaniga and G. R. Mangun (Cambridge, MA: The MIT Press), 1013–1023. doi: 10.7551/mitpress/9988.003.0017

Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., and Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition* 111, 364–371. doi: 10.1016/j.cognition.2009.02.001

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., and Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44, 389–400. doi: 10.1016/j.neuron.2004.09.027

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., and Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science* 293, 2105–2108. doi: 10.1126/science.1062872

Grusec, J. E., Goodnow, J. J., and Kuczynski, L. (2000). New directions in analyses of parenting contributions to children's acquisition of values. *Child Dev.* 71, 205–211. doi: 10.1111/1467-8624.00135

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychol. Rev.* 108, 814–834. doi: 10.1037/0033-295X.108.4.814

Haidt, J. (2012). *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. New York, NY: Pantheon Books.

Haidt, J. (2013). Moral psychology for the twenty-first century. *J. Moral Educ.* 42, 281–297. doi: 10.1080/03057240.2013.817327

Hauser, M., Cushman, F., Young, L., Kang-Xing Jin, R., and Mikhail, J. (2007). A dissociation between moral judgments and justifications. *Mind Lang.* 22, 1–21. doi: 10.1111/j.1468-0017.2006.00297.x

Huesmann, L. R., and Kirwil, L. (2007). "Why observing violence increases the risk of violent behavior in the observer," in *The Cambridge Handbook of Violent Behavior and Aggression*. eds. D. J. Flannery, A. T. Vazsonyi and I. D. Waldman (Cambridge, UK: Cambridge University Press), 545–570.

Istituto Nazionale di Statistica (ISTAT) (2018). Rapporto Bes 2018: il benessere equo e sostenibile in Italia [Bes 2018. The Equitable and Sustainable Welfare in Italy]. Available at: https://www.istat.it/it/files/2018/12/Bes_2018.pdf (Accessed August 26, 2021).

Jaffee, S., and Hyde, J. S. (2000). Gender differences in moral orientation: a meta-analysis. *Psychol. Bull.* 126:703. doi: 10.1037/0033-2909.126.5.703

JASP Team (2020). JASP (Version 0.14.1) [Computer software]. Available online at: https://jasp-stats.org/faq/how-do-i-cite-jasp/ (Accessed August 26, 2021).

Johnson, J. F., and Connelly, S. (2016). Moral disengagement and ethical decision-making. The moderating role of trait guilt and shame. *J. Pers. Psychol.* 15, 184–189. doi: 10.1027/1866-5888/a000166

Kahane, G., Everett, J. A., Earp, B. D., Caviola, L., Faber, N., Crockett, M., et al. (2018). Beyond sacrificial harm: A two-dimensional model of utilitarian psychology. *Psychol. Rev.* 125, 131–164. doi: 10.1037/rev0000093

Kahane, G., Everett, J. A., Earp, B. D., Farias, M., and Savulescu, J. (2015). 'Utilitarian' judgments in sacrificial moral dilemmas do not reflect impartial concern for the greater good. *Cognition* 134, 193–209. doi: 10.1016/j.cognition.2014.10.005

Kimonis, E. R., Frick, P. J., Skeem, J. L., Marsee, M. A., Cruise, K., Munoz, L. C., et al. (2008). Assessing callous–unemotional traits in adolescent offenders: validation of the inventory of callous–unemotional traits. *Int. J. Law Psychiatry* 31, 241–252. doi: 10.1016/j.ijlp.2008.04.002

Koenigs, M., Kruepke, M., Zeier, J., and Newman, J. P. (2012). Utilitarian moral judgment in psychopathy. *Soc. Cogn. Affect. Neurosci.* 7, 708–714. doi: 10.1093/scan/nsr048

Kohlberg, L. (1968). The child as a moral philosopher. *Psychol. Today* 2, 25–30.

Kohlberg, L. (1969). "Stage and sequence: The cognitive-developmental approach to socialization" in *Handbook of Socialization Theory and Research*. ed. D. Goslin (Chicago, IL: Rand McNally), 380–437.

Koleva, S., Selterman, D., Iyer, R., Ditto, P., and Graham, J. (2014). The moral compass of insecurity: anxious and avoidant attachment predict moral judgment. *Soc. Psychol. Pers.* 5, 185–194. doi: 10.1177/1948550613490965

Langdon, R., and Delmas, K. (2012). "Moral reasoning and psychopathic tendencies in the general community" in *Macquarie Monographs in Cognitive Science. Emotions, Imagination, and Moral Reasoning*. eds. R. Langdon and C. Mackenzie (New York, NY: Psychology Press), 91–118.

Larsen, E. M., Ospina, L. H., Cuesta-Diaz, A., Vian-Lains, A., Nitzburg, G. C., Mulaimovic, S., et al. (2019). Effects of childhood trauma on adult moral decision-making: clinical correlates and insights from bipolar disorder. *J. Affect. Disord.* 244, 180–186. doi: 10.1016/j.jad.2018.10.002

Mikhail, J. M. (2000). *Rawls' Linguistic Analogy: A Study of the 'Generative Grammar' Model of Moral Theory Described by John Rawls in 1A Theory of Justice'*. Ithaca (NY): Cornell University.

Patil, I. (2015). Trait psychopathy and utilitarian moral judgement: The mediating role of action aversion. *Cogn. Psychol.* 27, 349–366. doi: 10.1080/20445911.2015.1004334

Patil, I., and Silani, G. (2014). Reduced empathic concern leads to utilitarian moral judgments in trait alexithymia. *Front. Psychol.* 5:501. doi: 10.3389/fpsyg.2014.00501

Paxton, J. M., Ungar, L., and Greene, J. D. (2012). Reflection and reasoning in moral judgment. *Cogn. Sci.* 36, 163–177. doi: 10.1111/j.1551-6709.2011.01210.x

Pellizzoni, S., Siegal, M., and Surian, L. (2010). The contact principle and utilitarian moral judgments in young children. *Dev. Sci.* 13, 265–270. doi: 10.1111/j.1467-7687.2009.00851.x

Posada, R., and Wainryb, C. (2008). Moral development in a violent society: Colombian children's judgments in the context of survival and revenge. *Child Dev.* 79, 882–898. doi: 10.1111/j.1467-8624.2008.01165.x

Pujol, J., Batalla, I., Contreras-Rodríguez, O., Harrison, B. J., Pera, V., Hernández-Ribas, R., et al. (2012). Breakdown in the brain network subserving moral judgment in criminal psychopathy. *Soc. Cogn. Affect. Neurosci.* 7, 917–923. doi: 10.1093/scan/nsr075

Roccas, S., Sagiv, L., Schwartz, H. S., and Knafo, A. (2002). The big five personality factors and personal values. *Personal. Soc. Psychol. Bull.* 1, 789–801. doi: 10.1177/0146167202289008

Rohner, R. P., Khaleque, A., and Cournoyer, D. E. (2005). Parental acceptance-rejection: theory, methods, cross-cultural evidence, and implications. *Ethos* 33, 299–334. doi: 10.1525/eth.2005.33.3.299

Roose, A., Bijttebier, P., Decoene, S., Claes, L., and Frick, P. J. (2010). Assessing the affective features of psychopathy in adolescence: A further validation of the inventory of callous and unemotional traits. *Assessment* 17, 44–57. doi: 10.1177/1073191109344153

Rosas, A., and Koenigs, M. (2014). Beyond "utilitarianism": maximizing the clinical impact of moral judgment research. *Soc. Neurosci.* 9, 661–667. doi: 10.1080/17470919.2014.937506

Schwartz, S. H. (1992). "Universals in the content and structure of values: theoretical advances and empirical tests in 20 countries," in *Advances in Experimental Social Psychology. Vol. 25*. ed. M. P. Zanna (New York, NY: Academic Press), 1–65.

Schwartz, S. H. (2007). Universalism values and the inclusiveness of our moral universe. *J. Cross-Cult. Psychol.* 38, 711–728. doi: 10.1177/0022022107308992

Schwartz, S. H., Melech, G., Lehmann, A., Burgess, S., Harris, M., and Owens, V. (2001). Extending the cross-cultural validity of the theory of basic human values with a different method of measurement. *J. Cross-Cult. Psychol.* 32, 519–542. doi: 10.1177/0022022101032005001

Seider, S. (2009). The trouble with teaching ethics on trolley cars and train tracks. *J. Moral Educ.* 38, 219–236. doi: 10.1080/03057240902792744

Smetana, J. G. (2006). "Social-cognitive domain theory: consistencies and variations in children's moral and social judgments," in *Handbook of Moral Development*. eds. M. Killen and J. G. Smetana (Mahwah, NJ: Erlbaum), 119–154.

Smetana, J. G. (2013). "Moral development: The social domain theory view," in *Oxford Handbook of Developmental Psychology. Body and Mind. Vol. 1*. ed. P. D. Zelazo (New York, NY: Oxford University Press), 832–866.

Stey, P. C., Lapsley, D., and McKeever, M. O. (2013). Moral judgement in adolescents: age differences in applying and justifying three principles of harm. *Eur. J. Dev. Psychol.* 10, 206–220. doi: 10.1080/17405629.2013.765798

Takamatsu, R. (2019). Personality correlates and utilitarian judgments in the everyday context: psychopathic traits and differential effects of empathy, social dominance orientation, and dehumanization beliefs. *Pers. Individ. Differ.* 146, 1–8. doi: 10.1016/j.paid.2019.03.029

Takamatsu, R., and Takai, J. (2019). With or Without empathy: primary psychopathy and difficulty in identifying feelings predict utilitarian judgment in sacrificial dilemmas. *Ethics Behav.* 29, 71–85. doi: 10.1080/10508422.2017.1367684

Thomson, J. J. (1985). The trolley problem. *Yale Law J.* 94:5. doi: 10.2307/796133

Turiel, E. (1989). Domain-specific social judgments and domain ambiguities. *Merrill-Palmer Q.* 35, 89–114.

Turiel, E. (2006). "Thought, emotions and social interactional processes in moral development," in *Handbook of Moral Development*. eds. M. Killen and J. G. Smetana (Mahwah, NJ: Erlbaum), 1–30.

Zucchelli, M. M., and Ugazio, G. (2019). Cognitive-emotional and inhibitory deficits as a window to moral decision-making difficulties related to exposure to violence. *Front. Psychol.* 10:1427. doi: 10.3389/fpsyg.2019.01427

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

CONCEPTUAL ANALYSIS
published: 16 December 2021
doi: 10.3389/fpsyg.2021.756654

Conceptualizations of Knowledge in Structuring Approaches to Moral Development: A Process-Relational Approach

Jeremy I. M. Carpendale, Vicki L. Parnell and Beau Wallbridge*

Department of Psychology, Simon Fraser University, Burnaby, BC, Canada

Like other aspects of child development, views of the nature and development of morality depend on philosophical assumptions or worldviews presupposed by researchers. We analyze assumptions regarding knowledge linked to two contrasting worldviews: Cartesian-split-mechanistic and process-relational. We examine the implications of these worldviews for approaches to moral development, including relations between morality and social outcomes, and the concepts of information, meaning, interaction and computation. It is crucial to understand how researchers view these interrelated concepts in order to understand approaches to moral development. Within the Cartesian-split-mechanistic worldview, knowledge is viewed as representation and meaning is mechanistic and fixed. Both nativism and empiricism are based in this worldview, differing in whether the source of representations is assumed to be primarily internal or external. Morality is assumed to pre-exist, either in the genome or the culture. We discuss problems with these conceptions and endorse the process-relational paradigm, according to which knowledge is constructed through interaction, and morality begins in activity as a process of coordinating perspectives, rather than the application of fixed rules. The contrast is between beginning with the mind or beginning with social activity in explaining the mind.

Keywords: moral development, worldviews, knowledge, nativism, empiricism, constructivism, process-relational

OPEN ACCESS

Edited by:
Simona Carla Silvia Caravita, University of Stavanger, Norway

Reviewed by:
Tobias Krettenauer, Wilfrid Laurier University, Canada
Giuseppe Lo Dico, Catholic University of the Sacred Heart, Italy

****Correspondence:***
Jeremy I. M. Carpendale
jcarpend@sfu.ca

Specialty section:
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

Received: *10 August 2021*
Accepted: *22 November 2021*
Published: *16 December 2021*

Citation:
Carpendale JIM, Parnell VL and Wallbridge B (2021) Conceptualizations of Knowledge in Structuring Approaches to Moral Development: A Process-Relational Approach. Front. Psychol. 12:756654. doi: 10.3389/fpsyg.2021.756654

INTRODUCTION

Morality permeates our lives. It is embedded in our way of life, in ways that range from how we treat each other in everyday interaction, to broader social and political levels concerning injustice and inequality. The social fabric of our cultural worlds is woven around how we treat each other, and concern for others and their dignity is embedded in the pragmatic structures of human communication. Morality extends to larger scale social structures and decisions that affect others' lives and involves concern for others' welfare and well-being and obligations concerning what is right and wrong (Dahl and Killen, 2018). We face decisions that affect others at many levels, and we now have increasing awareness of how our actions affect others around the world through climate change. Thus, morality is a central aspect of being human and living with others.

The importance of morality in human life raises the question of how children become moral. Accounting for this aspect of child development depends on understanding the philosophical assumptions regarding knowledge and meaning which are the starting points for developmental research (Jopling, 1993). If we want to understand how children come to think about the

world, including moral aspects of experience, we cannot ignore how they come to know the world, and this requires an epistemological analysis concerning views about the nature and development of knowledge (Chapman, 1999). We analyze conceptions of knowledge assumed by researchers that are linked to two worldviews: Cartesian-split-mechanistic and process-relational (Overton, 2015). We explore how the related concepts of information, meaning, and interaction are conceptualized differently from the perspectives of these two worldviews, and play a role in the ways that morality is understood.

Two contrasting approaches to morality that tend to be taken for granted are the nativist view that moral norms are primarily explained as pre-existing in humans' biological nature, versus the empiricist view that moral norms pre-exist in the social world and are imposed on children from previous generations. These are generally the only options considered, so that if researchers argue against nativism it is assumed that they are arguing for empiricism. Most researchers, whether they emphasize nativism or empiricism, acknowledge that morality is some synthesis of biological and cultural (e.g., Bloom, 2010; Mikhail, 2020). Clearly it is necessary to include biological and evolutionary factors as well as social and cultural dimensions in understanding moral development, but it is possible to explicate the role of biology either from a gene centered approach, as in nativism, or an alternative, developmental systems approach (Griffiths and Tabery, 2013) that we outline, consistent with a process-relational perspective.

We argue that it is not just a matter of knowing where to draw the line in a "middle ground" between nativism and empiricism. This does not solve the problem because nature and nurture are still taken as pre-existing; information is assumed to pre-exist, either in the genes or in the environment. Both approaches are problematic, in that rather than explaining moral norms they explain them away by reducing them to either biological determinism or conformity to culturally imposed rules. In both views the passive individual is caused to act so there is no moral agent making choices to act morally, and thus neither approach cannot explain the development of a sense of moral obligation (Carpendale et al., 2010).

Morality is a complex aspect of human life that can be discussed in terms of cognition, emotions, and action. Various theorists focus on different aspects of this interrelated system and conceptualize the processes differently. For instance, in reacting to Kohlberg's emphasis on reasoning others have more recently swung to focus on emotions. However, their role in morality is conceptualized in radically different ways. One claim is that moral decisions are due to evolved emotional responses (Haidt, 2001). We recognize the important role of emotions, but we consider their role in structuring the relationships in which morality develops. From the developmental systems perspective we propose, although these aspects of morality can be abstracted, they are interrelated and bidirectionally interwoven over development. Morality concerns the coordination of actions with others, and cognition, emotions, and action are aspects of the interaction with others in which children develop morality.

From our perspective on understanding the links between morality and social consequences, morality emerges within lives lived with others, and thus, within social consequences. From the action-based perspective we endorse, morality emerges from social and emotional interaction, so social consequences underlie the development of morality. Once morality begins to emerge as a way of understanding and thinking about interaction with others, it then plays a role in individuals' choices about actions. Skills in understanding and making moral choices can influence future moral action in a complex bidirectional manner that plays out over developmental time.

In contrast to the approaches we criticize, we suggest that explaining the development of moral thinking and action should begin with interpersonal relations from the perspective of a process-relational worldview. We trace the implications for minds and morality of the view that "We are what we are through our relations with others" (Mead, 1934, p. 379). We propose that moral norms emerge in intersubjective experience through cooperative interaction among equals. We outline a process-relational approach to explaining the emergence of moral norms, beginning in infancy within the human social and emotional developmental system and extending through childhood. Interpersonal agreement is made possible in relationships based on mutual affection and mutual respect. Within such cooperative relationships among equals, practical morality emerges before children are able to articulate and then think about moral conflicts. In this way morality has its developmental roots in interactivity, and as children master a language they can then use it to reflect on moral issues and make decisions. Thus, morality is understood as beginning in social and emotional activity before children are able to think about and reflect on moral issues. What emerges is not a set of moral rules but rather a method or process for reaching moral decisions (Piaget, 1932/1965; Mead, 1934; Carpendale et al., 2013)[1].

We first outline and critique current nativist explanations of moral development, focusing primarily on Moral Foundations Theory (Graham et al., 2013). We then explore how the nativist-empiricist debate is embedded in the Cartesian-split-mechanistic worldview, and instead we suggest a developmental systems perspective. Then we consider how conceptions of meaning and knowledge are influenced by worldviews. Finally, as an

[1]Describing and explaining the emergence of more complex forms of organization is central to developmental psychology. However, the possibility of emergence is highly controversial in philosophy (O'Connor, 2020). One barrier to the possibility of emergence is the assumption that all definition permits back-translation. Bickhard (2009) claims, however, that this is false because implicit definition does not. A second barrier to the possibility of emergence is Jaegwon Kim's argument that emergence is not possible because if only particles possess causal power, then the organization of particles cannot possess causal power beyond the particles themselves. According to contemporary physics, however, rather than particles, the fundamental constituents of the world are dynamic quantum fields, which are processes. Bickhard (2009) argues that in contrast to a substance/particle metaphysics, within a process metaphysics organization can have causal power and thus emergence is possible. Many sciences are moving toward a process perspective (Griffiths and Stotz, 2000; Bickhard, 2009), and that is what we are suggesting. We suggest that morality does not exist at the level of the individual, it emerges from the coordination of conflicting perspectives, which is possible within cooperative relationships based on mutual affection and respect. Thus, what emerges is a pattern of organization that exists at the level of the coordination of the views of the people involved (Boom, 2004; Bickhard, 2008, 2009, 2011).

alternative, we outline an approach to moral development from the perspective of the process-relational worldview[2].

BIOLOGY AND ENVIRONMENT IN CONCEPTUALIZATIONS OF INTERACTION AND INFORMATION

There are many recent claims that some aspects of morality are innate (e.g., Hauser, 2006a,b; Hamlin et al., 2007; Mikhail, 2007, 2020; Bloom, 2010, 2012; Hamlin, 2013; Margolis and Laurence, 2013; Warneken, 2016). Bloom (2010, p. 46), for example, claimed that humans, "have a rudimentary moral sense from the very start of life.. . . Some sense of good and evil seems to be bred in the bone." There is ongoing debate regarding this claim (e.g., Prinz, 2009; Sterelny, 2010), and elsewhere we have criticized nativist approaches making claims that infants are born with innate principles of fairness (Carpendale et al., 2021), an innate moral core (Carpendale et al., 2013), and innate altruism (Carpendale et al., 2015; Carpendale and Lewis, 2021). Here we focus our critical attention primarily on one highly cited approach and we check the foundations of Moral Foundation Theory (Graham et al., 2013).

According to Moral Foundations Theory, "genes (collectively) write the first draft into neural tissue, beginning *in utero* but continuing throughout childhood. Experience (cultural learning) revises the draft during childhood, and even (to a lesser extent) during adulthood"(Graham et al., 2013, p. 61). In stating their claim that morality is innate, Graham et al. (2013, p. 100) assert that, "the discussion should focus on how exactly moral knowledge is innate, not whether it is." However, despite this strong claim, they and others do not discuss "how exactly moral knowledge is innate" (Dahl et al., 2021). In phrasing the problem in terms of *how* moral knowledge is innate, it seems that what is being proposed is an account of how knowledge could be innate, referring to the biological processes involved in getting from genes to justice and from neurons to norms. We suggest that an attempt to explicate *how* this is claimed to occur would make it clear that a direct route is in fact not consistent with current biology.

With its reference to the role of genes, nativism could mistakenly be taken for a biological approach, and to criticize nativism is thus to criticize the role of biological factors and instead argue for social factors. That is, it could be assumed that the only alternative to nativism's claim that knowledge pre-exists in biology is empiricism and social factors. But this buys into the very dichotomy we reject. In discussing flaws in nativism we are not arguing for empiricism. In fact, nativism and empiricism share many of the same problematic foundations of knowledge as pre-existing, either in the individual or in the society.

In fact, nativism does not amount to taking biology seriously, and does not actually rest on biological knowledge. Although Graham et al. (2013) claim that genes write into neural tissue, they fail to provide a reference to the biological literature regarding this claimed process[3]. Nativists seem to have a tendency to neglect citing biological research; this is consistent with the tradition following from Chomsky (2007), who was so compelled by his logical arguments that evidence from psychology or biology was not considered relevant. In fact, a number of biologists and geneticists have attempted to provide psychologists with some rudimentary idea of how genes work in the hope that they might improve understanding of the role of genes and environment in development (e.g., Fisher, 2006; Meaney, 2010), as well as how neural pathways are formed through experience (e.g., Mareschal et al., 2007; Stiles, 2009; Stiles et al., 2015). Genes do not write into neural tissue in the sense of forming neural connectivity. There is a long "tortuous" process through which genes are involved in a system resulting in human development (e.g., Fisher, 2006).

Genes are fairly inert molecules involved in the production of chains of amino acids as a step in the production of proteins, but the chains still need to be folded up within the cell. Although genes are crucial in development, it is problematic to assume that information exists at the level of genes or a "genetic program" because how genes function depends on many other factors in the environment, beginning at the level of the cell and extending to social interaction. The effect of a gene can vary as widely as promoting cell life or leading to cell death, depending on what co-factors exist in the cytoplasm (Meaney, 2010). Thus, the foundations of MFT appear to be based on implausible biological assumptions. Furthermore, the incredible complexity of neural interconnectivity that makes human lives possible is gradually shaped through experience (e.g., Stiles, 2009; Stiles et al., 2015).

Although clearly it is necessary to include biological and evolutionary factors as well as social and cultural factors in understanding moral development, the process-relational approach we propose is fundamentally different from a mixture of nativism and empiricism. The approach we endorse does of course consider the human developmental system consisting of evolved biological characteristics of infants and parents, set within cultural contexts. However, there are two fundamentally different ways in which this interaction can be conceptualized: either through a gene-centered approach according to which

[2]We lack the space to critically examine all of the aspects of nativism that we see as problematic, nor can we examine all of the versions of nativism and their roots in the history of ideas (see Allen and Bickhard, 2013). But it is important to explicate some of the underlying assumptions of nativism and empiricism in an attempt to avoid the assimilation of our approach to the Procrustean bed of this worldview which contains only these two options. If we criticize nativism it might be assumed that we therefore must be endorsing empiricism. Instead, in taking an action-based approach we step off the pendulum between the two options and work from the perspective of a different worldview (Allen and Bickhard, 2013).

[3]Graham et al. (2013) might respond that they were just using a metaphor, but they still need some actual biological process for getting from genes to neural interconnectivity that they claim is "*organized in advance of experience*" (p. 61, emphasis is original). They might respond that this is a problem for biologists not for them. However, the problem they give to biologists should at least be consistent with what is currently known about genetics and neural development (e.g., Stiles, 2009; Meaney, 2010; Stiles et al., 2015). Even if genes could "write . . . into neural tissue" by specifying neural interconnectivity it is not clear how many genes would be required to pre-specify the trillions of synaptic connections in human brains, and many species have more genes than humans. Infants are born with exuberant connectivity that is "pruned" away after birth in an ongoing dynamic and interactive process of brain development that does not seem consistent with Graham et al's. (2013) claims (Mareschal et al., 2007; Stiles et al., 2015).

these are considered separate, pre-existing factors that then interact, or from a developmental systems perspective that does without the dichotomy (e.g., Gottlieb, 2007; Griffiths and Tabery, 2013). We have discussed flaws with the gene-centered approach that assumes genes carry information in a genetic program; instead, we suggest that genes are one crucial factor in a system of bidirectionally interacting factors. Thus, in contrast to nativism, process-relational approaches take biology seriously through its role in human developmental systems. Rather than thinking of biology and environment as separable factors that interact, if we look closely at development, we see a system from which these interwoven factors are abstracted. Biology cannot be separated from environments because organisms create and elicit their environments through their characteristics, and, in turn, biology is structured by experience. This process begins with infants' action tendencies and sensory abilities in early development, which elicit the environment in which they develop in a bi-directional manner.

Regularity in developmental outcomes could be assumed to be evidence for nativists' claims of knowledge pre-existing at the genetic level. However, such regularity does not mean that it must be due to something pre-existing, such as genes. To illustrate this point, consider the example of mature forests in particular climate zones. The type of forest tends to be a consistent and regular outcome over many decades. This is the result of the interaction of multiple factors, including the characteristics of the species involved such as their tolerance for shade and so on. For example, a mature forest in the Pacific northwest tends to be Western Red Cedar and Douglas Fir, even though the information for that outcome does not pre-exist anywhere. Instead, it is the natural outcome of ecological succession within particular conditions (Griffiths and Stotz, 2000).

Rather than assuming pre-existing information encoded in genes, we argue that regularities in the development of moral knowledge can emerge through typical human developmental systems. We now turn from criticizing claims about getting from genes to the first draft of moral cognition, to analyzing the nature of this cognition.

CONCEPTIONS OF KNOWLEDGE AND MEANING IN CONTRASTING WORLDVIEWS

The moral nativists we are focusing on assume the computational/representational framework. According to this computational theory of mind, thinking—in this case, about morality—is based on computation involving mental representations that are linked to the world (e.g., Graham et al., 2013; Mikhail, 2020). Moral Foundations Theory

> *"proposes that the human mind is organized in advance of experience so that it is prepared to learn values, norms, and behaviors related to a diverse set of recurrent adaptive social problems. . .. We think of this innate organization as being implemented by sets of related modules which work together to guide and constrain responses to each particular problem"* (Graham et al., 2013, p. 63).

For example, it is assumed that infants are born with innate principles such as a principle of fairness in the "first draft" of moral cognition (e.g., Bian et al., 2018; Buyukozer Dawkins et al., 2019). From this perspective, thinking is computation based on mental representations that are meaningful because they are linked to the world.

An issue with this view, however, is the problem of how such representations are linked to the world in a way that makes them meaningful. As Wittgenstein (1953/2009) has argued, representations cannot carry their own meaning because any representation could be interpreted in multiple ways (e.g., Heil, 1981; Goldberg, 1991). Thus, the only way to bring meaning into the computational/representational framework is to implicitly assume a homunculus (Heil, 1981; Kenny, 1991). That is, what is required is something in the system like a small person that attributes meaning to the representation, just as a person must attribute meaning to the input and output of a computer. This, of course, is problematic because it just puts off rather than provides an explanation (for further criticism of the computational theory of mind see e.g., Heil, 1981; Carpendale et al., 2021; Carpendale and Lewis, 2021).

As Hobson (2002, p. 14) noted, "computers don't understand anything, nor do they care." To further spell out Hobson's point, computers don't understand *because* they don't care. A computational system is not linked to the world in ways that can involve significance and meaning. Morality, however, is necessarily based on the emotional significance of social actions and outcomes for the people involved. It involves coordinating conflicting goals.

A mechanical process is not based on meaning, so it is not possible to get from computation to caring. To claim that a computer program can contain principles of fairness is like assuming that an automatic door opener is polite. The door works through a passive mechanical process. If it fails to open it is not rude, just broken. In contrast, persons open doors for others because they are recognized as persons with goals and all that entails. Similarly, although it is possible to design a mechanism to divide resources equally, this does not mean that the machine is moral and is applying a principle of fairness. If it fails to operate properly it is not immoral, just defective. If morality is reduced to the mechanical computation that genes write into neural tissue, then it is causal not normative, and within such an approach the normative dimension of morality is not explained; rather, it is explained away. Additionally, this sort of an approach risks self-contradiction if theorists claim to be involved in the normative enterprise of science yet there is no place for reasons in their theory (Habermas, 1983/1990; Carpendale et al., 2010). This would be like "sawing off the branch on which one sits" (Bennett and Hacker, 2003, pp. 376–377). We have analyzed the influence of how information is conceptualized in the context of genes encoding information, and how meaning is conceptualized as fixed in the context of the computational view of the mind. Both conceptualizations of information and meaning in this Cartesian-split-mechanistic worldview are derived from the theory of knowledge as representation. Both nativism and empiricism, although apparently different approaches, actually begin from the same Cartesian-split-mechanistic worldview (Overton, 2015),

and the theory of knowledge as pre-existing representation that is located either primarily in the child and explained through biology or in the social world and imposed on the child.

This conception of knowledge is what Piaget (1970) referred to as a copy theory and Dewey (1960) labeled the spectator notion of knowledge. Both descriptions bring out the point that knowledge is viewed as solely based on perception (Piaget, 1970/1972). According to this perspective, developing knowledge is viewed as forming a copy of the world. But because it is assumed that we have no direct access to the world, the only way to check our representation is by forming another copy (Wittgenstein, 1953/2009). Thus, we cannot tell if our representation of the world that is meant to reflect our knowledge of the world is accurate because we cannot directly compare it to the world. If we cannot become aware of errors then we cannot learn, and an approach that cannot account for learning is fundamentally flawed (Bickhard, 2009). Although such an approach is meant to explain knowledge, in fact, it already presupposes knowledge (e.g., Dewey, 1960; Piaget, 1970; Chapman, 1999; Carpendale and Lewis, 2004, 2006, 2021; Bibok et al., 2008; Bickhard, 2009).

We have argued that the worldview on which nativism and empiricism are based is problematic, and is based on a flawed view of knowledge as representation, which results in flawed conceptualizations of information and meaning. An alternative approach to knowledge, consistent with the process-relational worldview, is an emergent constructivism according to which knowledge develops through learning the interactive potential of the world. This world includes other people, which adds a normative dimension to this process. An infant is an agent with needs and emotions who learns how the world responds to her actions in positive or negative ways. Thus, meaning emerges along with anticipations as the child learns about the world and what she can do with it (Piaget, 1936/1952), and perception is seeing the world in terms of the potential for interaction (e.g., Chapman, 1999). An understanding of fairness is based on the meaning actions have in terms of the emotional consequences for the people involved, and others' reactions must be valued. We now turn to explicating a view of moral development based on this perspective.

A PROCESS-RELATIONAL APPROACH TO MORAL DEVELOPMENT

Moral norms concern what is right and wrong. They are not a part of the physical world, so how do they arise (Brandom, 1994)? We don't observe an understanding of commitment, obligation, and right and wrong in other species. What, then, is the source of moral norms? Children grow up in cultures with moral norms that are imposed on them. They may come to understand and accept these moral rules, or perhaps challenge and possibly change such norms. To account for this and explain how it is that rather than passively accept rules children may challenge and attempt to change cultural norms, we argue that moral norms and right and wrong emerge at the level of action and interaction. Assumptions about the source of moral norms are linked to assumptions about their nature. That is, if they are causal through being determined by biology, or external and imposed from the outside, then it could be questioned whether they are actually morality because the individual is being compelled to act (Wright, 1982a,b). Thus, this way of thinking seems to define moral norms out of existence (Carpendale et al., 2010). Although neurons are necessary for morality, they do not cause norms and explain moral development. Just as we do not get an answer for why two plus two is four at the level of neural activity (Piaget, 1971, p. 49), morality does not arise at the level of the individual and biological activity.

The process-relational approach we propose begins from activity, and thus fits with Piaget's (1932/1965) still overlooked view that children first work out a way of interacting with each other in their practical activity through coordinating their actions with others. Piaget began from practical interaction, within a particular form of relationship that is based on mutual affection and mutual respect. In these relationships, children enjoy the interaction and want to continue it and thus have to work out a way of getting along with each other and coordinating their sometimes conflicting goals. This is a form of interaction that is best suited to reaching mutual understanding because equals feel an obligation to listen to each other as well as explain their own position (Piaget, 1932/1965). There is already a form of morality in the "constitutive rules" that structure such relationships because the individuals involved treat each other as persons and listen to others as well as explain their own perspective. Thus, "morality is the logic of action" (Piaget, 1932/1965, p. 398). Within such interaction it is possible to formulate "constituted rules" concerning how to coordinate their action (Piaget, 1932/1965; Carpendale, 2009).

This outcome is a coordination of everyone's interests, and it must be based on caring about each other as a foundation in structuring the interaction. Caring is not something that is added later or reached through reasoning. Instead, concern for others and not just taking everything for oneself is part of the foundation of the human developmental system. This involves treating others as persons—someone, not something (Spaemann, 2006). This interaction is based on care, affection, and enjoyment of interaction. Language is used in order to explain oneself and listen to others in coordinating conflicting goals.

Once a form of morality has emerged at the practical, lived level, Piaget (1932/1965) then suggested that a gradual process of "conscious realization" occurs, through which children became able to verbally articulate the ways of interacting that are already present in their activity. Here language plays an important role, first, in the process of reaching mutual understanding within interaction among equals and achieving a solution to conflicts that is agreeable to all. Second, language again plays a role in the process of articulating and reflecting on that earlier achieved competence in the way children treat each other. From this perspective morality does not begin in the structure of the mind. Instead, it begins in social activity and emerges in the coordination of action through experiencing the consequences of one's actions. The mind is structured through activity and cannot be structured prior to experience.

From this perspective, morality is a process. There is nothing "objective" about this in a foundational or fixed sense of certainty outside of human interaction. Morality is also not "subjective" in the sense of just being based on personal whims. Instead, it is intersubjective in the sense of arising through coordination with others. It is the logic of interaction that arises in the coordination of action with others (Piaget, 1932/1965). And in this sense, it is an aspect of communication and cooperation through being based on valuing all other viewpoints as equal. This begins by being rooted in relations based in mutual affection and caring for each other. Children first work out a practical morality as a way of coordinating their action with others because it is more enjoyable to play that way. Adults can play a role in facilitating relationships that are best suited to reaching understanding. But this is different from adults imposing rules that children don't understand.

Similar ideas are present in the work of other theorists. The idea of coordinating conflicting perspectives was also present in Mead's (1934, p. 389) argument for a moral process that involves considering all perspectives involved—"the method of morality." Mead (1934, p. 379) derived the universality of moral judgment from our social nature, "from the fact that we take the attitude of the entire community, of all rational beings"... "that is, everyone who can rationally appreciate the situation agrees." Moral norms should be grounded in good reasons that cannot be rationally rejected by anyone involved (Mead, 1934). Kohlberg had a similar perspective with his idea of moral development as ideal role taking, and his notion of "moral musical chairs" consisting of taking all the perspectives involved in a moral dilemma. But this was overshadowed by Kohlberg's unfortunate adoption of a problematic view of Piaget's stages (Carpendale, 2000). Habermas (1983/1990) argued that aspects of this process of considering all relevant points of view are embedded in the structure of conversation and argumentation, and thus engaging in these activities presupposes morality (see also Forst, 2005).

Piaget's (1932/1965) account focuses on interaction among school aged children, but even to get to this point in development we suggest that it is already possible to see the beginning of morality in the preconditions for interaction and communication (Winch, 1972). Infants' biological embodiment, such as being helpless at birth and thus requiring care, structures the social and emotional system in which they develop (Portmann, 1944/1990; Carpendale and Lewis, 2021). Here we focus on the moral aspects of the human developmental system. The roots of morality are already emerging in the way in which caregivers respond to their infants as persons. Even in the first few months of life mothers find it difficult to treat their babies as objects when researchers ask them to hold a "still face" rather than respond normally to their young infants. If infants are accustomed to the enjoyable interaction, they often try to elicit it by smiling if it is missing. Although caregivers' difficulty in not responding to their babies is considered an obstacle in conducting research using the "still face" paradigm, from our perspective it is a finding revealing the way caregivers respond to their infants as persons (Mcquaid et al., 2009). This reciprocal responding to each other may be a source of the expectation for turn taking in interaction. Treating others as someone not something is already present in this early interaction. A foundational component for human declarative communication is the development of children's enjoyment in participating in interaction as a goal in itself. This development can be charted from enjoying attention to self, then to what the self does, and to the activity of showing and giving objects, as well as also to interest and enjoyment in participating in adults' activities (Bates et al., 1975; Rheingold, 1982; Reddy, 2003).

We begin from children's activity within a developmental system, thus our approach could be mistaken for behaviorism. But our goal is to explain psychological development as arising from interaction and communication, rather than assume the mind as presupposed and therefore not explained (e.g., Mead, 1934). Behaviorism is restricted to passive association and cannot account for meaning. Thus, behaviorism is situated within the Cartesian-split-mechanistic worldview that we have criticized. Instead, from an action-based process approach, infants learn about the world through their experience within which they come to anticipate outcomes of their actions. They perceive the world in terms of potential for interaction (e.g., Chapman, 1999), and through this process the world becomes meaningful to them. Instead of a passive association, any association is due to the meaningful relationship the child forms. From an action-based perspective, the child is not thought of as learning a response to a stimulus. Instead, the child is "coming to organize his activity in a particular way (which he can extend to other contexts) and coordinating this activity with the corresponding acts of the mother" (Clark, 1978, p. 240). The child is not mechanistically and passively forming associations between meaningless unrelated stimuli. Instead, she is active with goals, needs and interests, as she learns about the potential for action on the physical world. Infants' interaction with the world becomes more complex as they come to anticipate the outcomes of their actions. When infants begin employing an action as a means to attain a goal, we can say they are acting with intention, typically beginning at about 8 months of age (Piaget, 1936/1952). Toward the middle of their second year, toddlers begin to coordinate action schemes implicitly or mentally. This is the beginning of one form of mental activity in which they can mentally anticipate the outcomes of their actions (Piaget, 1936/1952).

Unintentional communication is present in interaction even in infants' early months of life because, for example, the crying of a newborn infant has significance for caregivers. An important transition is to intentional communication beginning toward the end of the first year of life. This occurs as infants learn the significance or meaning their actions have for others. They come to anticipate social outcomes of stable social structures, or routines, such as requests, responding to questions, or sharing attention. These shared social routines form stable structures of interaction with common expectations in which words can be used (Carpendale and Lewis, 2021).

The links between communication and morality are complex and bidirectional over development. Communication in the human intentional sense and morality are both located within and emerge from forms of interaction in which individuals develop and experience themselves in relation to others (e.g., Carpendale, 2018). Communication from this perspective is not a matter of transmitting meaning attached to symbols,

words, or representations, as it is in the Cartesian-split-mechanistic worldview. Rather, communication is situated in the coordination of action and so it entails a view of selves in relations to others, that is, morality. Communication and morality are two sides to a coin, two ways of talking about human relations.

Language provides a way of talking about and reflecting on as well as understanding human activity with mental state terms. Beginning in their third year, children start to learn words referring to mental states such as want, know, forget and so on. These words refer to human activity (Carpendale and Lewis, 2015). Through experiencing others' reactions to themselves, children come to take themselves as an object. That is, to take others' perspective on their self, and thus to become able to reflect on their self, and now to have a self rather than just be a self (Mead, 1934). Their psychological language can now be used to reflect on their own experience as well as help in understanding others' experience. This is a developmental outcome of a gradual process that begins in activity and interaction with others (Wittgenstein, 1953/2009; Canfield, 2007; Racine and Carpendale, 2008). This approach to the development of children's social understanding contrasts with the causal psychological view that mental states underlie and cause outer behavior, assumed in the Cartesian perspective underlying much of the work on children's theories of mind (see Carpendale and Lewis, 2004, 2006, 2015, 2021; Racine and Carpendale, 2008). Instead, mental states are logically linked to such action and cannot be identified independently (Racine and Carpendale, 2008).

The link between morality and communication can be seen in the way that cooperation is a principle that underlies and structures typical conversation because it is possible to derive additional meaning from utterances in conversation based on the assumption that others are cooperating in conveying meaning (Grice, 1975). Furthermore, Grice viewed conversation as a special case of human cooperative interaction in general. Although it is possible to use language to deceive others, lying is only possible because truth telling is the norm (Holiday, 1988).

Concern for others and their dignity is embedded in the pragmatics of conversation and politeness (Brown and Levinson, 1987; Turnbull, 2003). The care and concern for others that structures our interaction is also a foundation for the development of moral obligation (Carpendale, 2018; Carpendale and Lewis, 2020). Language is based on moral aspects of interaction and enables further complexity in morality through understanding and coordinating with others' perspectives.

Some of the processes we have presupposed in moral development are based on individuals giving reasons to others and listening to others' reasons. But in the casual world of natural science how do we find room for reasons? Reasons have no place in the lives of other animals, even social species. Why is it that humans give reasons and expect reasons from others? From the perspective we take, reasoning emerges within social relations, within interpersonal obligations to others because they are given to others (Kitchener, 2004): "Man is a rational being because he is a social being" (Mead, 1934, p. 379). Giving reasons to others and valuing others' reasons and responding develops within interpersonal relations of caring and obligation. In some families, young children are given reasons even before they can understand them, whereas in other families they may be told to do things without reasons. Children must also learn when reasons are expected from them, often to explain actions in the context of social expectations. This requires understanding our actions in relation to others and appreciating the effect of our actions on others, which involves being able to take others' perspectives on the self. In Mead's (1934, p. 138) words, "The importance of what we term 'communication' lies in the fact that it provides a form of behavior in which the organism or the individual may become an object to himself." This allows individuals to see their actions in relations to others.

This link between communication and morality is further explicated by Spaemann (2006, pp. 14–15):

> *"To speak of oneself in the third person is to step out of the central position that all living things in nature occupy in relation to their environments, and to see oneself with other people's eyes as something 'out there'. For this one must adopt a point of view from outside of one's own organic center. Morality is possible only with this capacity for self-objectification and self-relativization; only on these terms, too, is speech possible. Speech differs from the cries of living things in nature, in that it anticipates the standpoint of the one who is to hear what is spoken. When someone says, 'I am in pain', that statement is not merely a cry by other means. The immediate expression of pain must be suppressed, in order to form a communication about the pain as an event in the world and to make that communication intelligible to another."*

This ability to take oneself as an object through appreciating others' attitudes toward the self makes it possible to consider one's own perspective and action in relation to others, and this is required in order to coordinate perspectives and arrive at moral solutions. This social process of seeing oneself in others' eyes may also be important in coming to value and integrate the moral dimensions of oneself, that is, to develop one's moral self (Krettenauer, 2013). The approach to this problem that is consistent with the process-relational framework we endorse is to conceptualize development in this area as an inter-personal social process that is based in lived interaction and crosses multiple domains of development (Krettenauer, 2013). The relationships of mutual respect in which practical morality emerges are also linked to a sense of self value and confidence, as well as moral and intellectual development (Wright, 1982b).

We have grappled with the task of explaining the development of moral norms. But, given the extent of injustice and oppression in the world, it could be argued that the development of morality does not always, or even typically, occur, so why does it go wrong? From our perspective, development occurs within a system and thus many factors can vary resulting in different outcomes. But the first step is to explain how morality is even possible. Furthermore, we have to explain how it is that we can recognize injustice when it occurs. In addition, one's conception of how things can go wrong depends on one's view of how they can go right. Thus, the first step is to explain how morality is possible.

Of course, not all relationships children experience are cooperative. Children's early relationships may necessarily tend to be more constraining, and any relationship is some mixture of constraint based on one-sided respect and inequality, and cooperation based on mutual respect and equality (Piaget, 1932/1965). Children do sometimes constrain and bully others, and, on the other hand, parents can be cooperative to varying degrees. The process we have described can run off the rails in various ways, and through this experience children learn about the social consequences of their actions (e.g., Dahl et al., 2011). There are factors in human interaction such as greed leading to oppression and inequality, and language can be used to deceive and oppress others. But there is also constant struggle against such forces as individuals recognize oppression, inequality, and lack of fairness. We are inclined to be convinced by Piaget's (1932/1965) hopeful stance that in spite of setbacks there is still the potential for gradual progressivity in societal change toward a more just world (Chapman, 1988). Based on Piaget's position, we argue that moral progress on a societal level is made possible because, as part of the foundation of moral development, caring for others means learning to care for and respect others' perspectives and to treat these perspectives as equivalent to our own. These are the constitutive rules that underlie engaging in a moral process in which norms can be constructed and changed based on negotiation and consensus within relationships of mutual affection and mutual respect. This is a social process which begins with caring for others in close relationships and can then be extended to engaging with and respecting other perspectives on a societal level. Such a process, originally rooted in close interpersonal relationships, provides one means for moral progress as more perspectives become coordinated at a societal level (Mead, 1934; Carpendale and Lewis, 2020, 2021). This progress could extend beyond the initial circle of close relationships to include more perspectives within one's culture and beyond, and could also be extended to other species and the broader biosphere[4].

DISCUSSION

Humans inhabit webs of moral obligations and commitments. We live with a sense of right and wrong, an understanding of what ought to be done in a world of norms, a space of justification. We have grappled with the issue of explaining the source of such moral norms, and accounting for how the feeling of obligation emerges in children's lives as different from conformity to social conventions. Two common approaches are that moral norms are imposed on children by previous generations or that they are innate. We have argued that by themselves these explanations are incomplete and are attempts to explain away rather than explain moral norms. Nativists reduce morality to being caused to act due to biology and empiricists reduce morality to conformity to socially imposed rules, so either, or a combination of the two, rule out the person as actually making a choice. From our perspective, both are no longer talking about morality, but are instead referring either to something within the child compelling her to act or to conformity to social rules. Although children do grow up within cultures with moral norms imposed on them, this does not explain the source of such norms and children must still come to understand and perhaps challenge and change such norms. A middle ground as a mixture of the two still does not deal with morality.

The assumption that biological factors play a role in morality must be further spelled out (Dahl et al., 2021) and this can be done either from a gene-centered perspective or a developmental systems approach. From the perspective we propose, morality and the idea of justice emerges and does not pre-exist in either societal beliefs nor in the biology of the individual. It develops reliably given certain conditions, just as a whirlpool is a structure that emerges in the flow of liquids given certain conditions, although it does not pre-exist anywhere. The individual's biological heritage results in the conditions in which the idea of justice can emerge. From this perspective, it is recognized that knowledge cannot be innate in the sense that it is directly the result of genes, but rather that there is a much more complex developmental system in which ideas about morality can develop (Piaget, 1932/1965; Carpendale, 2009; Carpendale et al., 2013). Thus, a third option in understanding the development of moral norms is the developmental systems approach within a process-relational perspective according to which biological and social factors do not simply pre-exist separately but are instead abstracted from social and emotional developmental systems in which they are intertwined and mutually create each other.

The way that the link between morality and social outcomes is conceptualized depends on the worldview researchers adopt. From the perspective we take, the goal is to trace a natural history of the development of moral norms through the increasingly complex forms of coordination emerging in dyads as children construct social and moral skills through their interaction with others. Morality concerns the coordination of action with others, and it emerges within the social consequences of children's actions at the level of intersubjective engagement with others who we respect and care for. Within cooperative relationships among equals children work out what is fair at a practical lived level. Norms are first implicit in interaction and ways of treating others, and children gradually come to consciously realize the principles that underlie their practical interaction. This process of interpersonal coordination continues in more complex ways with language as children became aware of the implicit norms that structured their interaction, and with the development of reasoning and justifications. From a process-relational perspective, communication and morality are interwoven. Morality emerges as an aspect of living with others. It is not that care and morality had to evolve as something separate, but rather caring about each other is what makes us human. It structures the human developmental system, the human social emotional cradle in which children develop (e.g., Carpendale and Lewis, 2021). Morality and communication emerge out of human

[4]Our approach contrasts with the two steps in explaining moral obligation proposed by Tomasello (2020): first, the level of interpersonal obligation, and a second level of "objective" morality. He explains the second step as involving an extension to the collective resulting through social pressure experienced from the internalized "we." In essence, this is conformity rather than morality. Although conformity is a part of human life, we suggest that this is an incomplete account of morality (Carpendale and Lewis, 2020).

relations. From an action-based perspective we begin with social activity, and the social consequences of children's action form the experience through which children develop morality. In a bi-directional manner, the understanding children develop in this process can then influence their subsequent thinking and action.

AUTHOR CONTRIBUTIONS

JC wrote the first draft. VP and BW revised and added to that draft. All authors contributed to the article and approved the submitted version.

REFERENCES

Allen, J., and Bickhard, M. (2013). Stepping off the pendulum: why only an action-based approach can transcend the nativist–empiricist debate. *Cogn. Dev.* 28, 96–133. doi: 10.1016/j.cogdev.2013.01.002

Bates, E., Camaioni, L., and Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill Palmer Q.* 21, 205–226. doi: 10.1007/978-94-009-8964-1_9

Bennett, M. R., and Hacker, P. M. S. (2003). *The Philosophical Foundations of Neuroscience*. Oxford: Blackwell.

Bian, L., Sloane, S., and Baillargeon, R. (2018). Infants expect ingroup support to override fairness when resources are limited. *Proc. Natl. Acad. Sci. U. S. A.* 115, 2705–2710. doi: 10.1073/pnas.1719445115

Bibok, M. B., Carpendale, J. I. M., and Lewis, C. (2008). "Social knowledge as social skill: an action based view of social understanding," in *Social Life and Social Knowledge: Toward a Process Account of Development*, eds U. Müller, J. I. M. Carpendale, N. Budwig, and B. Sokol (New York: Taylor Francis), 145–169.

Bickhard, M. (2008). "Are you social? The ontological and developmental emergence of the person," in *Social Life and Social Knowledge: Toward a Process Account of Development*, eds U. Müller, J. I. M. Carpendale, N. Budwig, and B. Sokol (New York: Taylor Francis), 17–42. doi: 10.1037/a0027033

Bickhard, M. (2009). The interactivist model. *Synthese* 166, 547–591. doi: 10.1007/s11229-008-9375-x

Bickhard, M. H. (2011). Some consequences (and enablings) of Process Metaphysics. *Axiomathes* 21, 3–32. doi: 10.1007/s10516-010-9130-z

Bloom, P. (2010). *The Moral Life of Babies*. New York: New York Times.

Bloom, P. (2012). "Moral nativism and moral psychology," in *The Social Psychology of Morality: Exploring the Causes of Good and Evil*, eds M. Mikulincer and P. R. Shaver (Washington: APA), 71–89. doi: 10.1037/13091-004

Boom, J. (2004). "Individualism and collectivism: a dynamic systems interpretation of Piaget's interactionism," in *Social Interaction and the Development of Knowledge*, eds J. I. M. Carpendale and U. Muller (Mahwah: Erlbaum), 67–85.

Brandom, R. B. (1994). *Making it Explicit: Reasoning, Representing and Discursive Commitment*. Cambridge: Harvard University Press.

Brown, P., and Levinson, S. (1987). *Politeness*. Cambridge: Cambridge University Press.

Buyukozer Dawkins, M. B., Sloane, S., and Baillargeon, R. (2019). Do infants in the first year of life expect equal resource allocations? *Front. Psychol.* 10:116. doi: 10.3389/fpsyg.2019.00116

Canfield, J. V. (2007). *Becoming Human: The Development of Language, Self, and Self-Consciousness*. New York: Palgrave Macmillan.

Carpendale, J. I. M. (2000). Kohlberg and Piaget on stages and moral reasoning. *Dev. Rev.* 20, 181–205.

Carpendale, J. I. M. (2009). "Piaget's theory of moral development," in *The Cambridge Companion to Piaget*, eds U. Müller, J. I. M. Carpendale, and L. Smith (Cambridge: Cambridge University Press), 270–286.

Carpendale, J. I. M. (2018). "Communication as the coordination of activity: the implications of philosophical preconceptions for theories of the development of communication," in *Advancing Developmental Science: Philosophy, Theory, and Method*, eds A. S. Dick and U. Müller (New York: Routledge, Taylor and Francis group), 145–156.

Carpendale, J. I. M., Hammond, S. I., and Atwood, S. (2013). "A relational developmental systems approach to moral development," in *Embodiment and Epigenesis: Theoretical and Methodological Issues in Understanding the Role of Biology Within the Relational Developmental System*, eds R. M. Lerner and J. B. Benson (Cambridge: Academic Press), 105–133.

Carpendale, J. I. M., Kettner, V. A., and Audet, K. N. (2015). On the nature of toddlers' helping: helping or interest in others' activity? *Soc. Dev.* 24, 357–366. doi: 10.1111/sode.12094

Carpendale, J. I. M., and Lewis, C. (2004). Constructing an understanding of mind: the development of children's social understanding within social interaction. *Behav. Brain Sci.* 27, 79–96.

Carpendale, J. I. M., and Lewis, C. (2006). *How Children Develop Social Understanding*. Oxford: Blackwell.

Carpendale, J. I. M., and Lewis, C. (2015). "The development of social understanding," in *Handbook of Child Psychology and Developmental Science*, eds L. Liben and U. Müller (Hoboken: Wiley Blackwell), 381.

Carpendale, J. I. M., and Lewis, C. (2020). Tomasello's tin man of moral obligation needs a heart. *Behav. Brain Sci.* 43, 19–20. doi: 10.1017/S0140525X19002462

Carpendale, J. I. M., and Lewis, C. (2021). *What Makes us Human? How Minds are Constructed in Relationships*. Milton Park: Routledge.

Carpendale, J. I. M., Müller, U., Wallbridge, B., Broesch, T., Cameron-Faulkner, T., and Ten Eycke, K. (2021). The development of giving in forms of object-exchange: exploring the roots of communication and morality in early interaction around objects. *Hum. Dev.* 65, 166–179. doi: 10.1159/000517221

Carpendale, J. I. M., Sokol, B., and Müller, U. (2010). "Is a neuroscience of morality possible?" in *Developmental Social Cognitive Neuroscience*, eds P. Zelazo, M. Chandler, and E. Crone (New York: Psychology Press), 289–311.

Chapman, M. (1988). Contextuality and directionality of cognitive development. *Hum. Dev.* 31, 92–106. doi: 10.1016/j.chc.2016.12.007

Chapman, M. (1999). Constructivism and the problem of reality. *J. Appl. Dev. Psychol.* 20, 31–43.

Chomsky, N. (2007). Biolinguistic explorations: design, development, evolution. *Int. J. Philos. Stud.* 15, 1–21. doi: 10.1080/09672550601143078

Clark, R. A. (1978). "The transition from action to gesture," in *Action, Gesture and Symbol: The Emergence of Language*, ed. A. Lock (New York: Academic Press), 231–257.

Dahl, A., Baxley, C. C., and Waltzer, T. (2021). The two-front forever war: moral nativism and its critics. *Hum. Dev.* 65, 180–187. doi: 10.1159/000517406

Dahl, A., Campos, J. J., and Witherington, D. C. (2011). Emotional action and communication in early moral development. *Emot. Rev.* 3, 147–157.

Dahl, A., and Killen, M. (2018). A developmental perspective on the origins of morality in infancy and early childhood. *Front. Psychol.* 9:1736. doi: 10.3389/fpsyg.2018.01736

Dewey, J. (1960). *On Experience, Nature, and Freedom*. Indianapolis: The Bobbs-Merrill Company, Inc.

Fisher, S. E. (2006). Tangled webs: tracing the connections between genes and cognition. *Cognition* 101, 270–297. doi: 10.1016/j.cognition.2006.04.004

Forst, R. (2005). Moral autonomy and the autonomy of morality: toward a theory of normativity after Kant. *Grad. Fac. Philos. J.* 26, 65–88. doi: 10.5840/gfpj20052611

Goldberg, B. (1991). "Mechanism and meaning," in *Investigating Psychology: Sciences of the Mind After Wittgenstein*, ed. J. Hyman (New York: Routledge), 48–66.

Gottlieb, G. (2007). Probablistic epigenesis. *Dev. Sci.* 10, 1–11.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., et al. (2013). Moral Foundations Theory: the pragmatic validity of moral pluralism. *Adv. Exp. Soc. Psychol.* 47, 55–130.

Grice, H. P. (1975). "Logic and conversation," in *Syntax and Semantics*, eds P. Cole and J. L. Morgan (New York: Academic Press), 41–58. doi: 10.1163/9789004368811_003

Griffiths, P. E., and Stotz, K. (2000). How the mind grows: a developmental perspective on the biology of cognition. *Synthese* 122, 29–51.

Griffiths, P. E., and Tabery, J. (2013). "Developmental systems theory: what does it explain, and how does it explain it?" in *Embodiment and Epigenesis: Theoretical and Methodological Issues in Understanding the Role of Biology Within the Relational Developmental System*, eds R. M. Lerner and J. B. Benson (Cambridge: Academic Press), 65–94.

Habermas, J. (1983/1990). *Moral Consciousness and Communicative Action*. Cambridge: The MIT Press.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychol. Rev.* 108, 814–834. doi: 10.1037/0033-295x.108.4.814

Hamlin, J. K. (2013). Moral judgment and action in preverbal infants and toddlers: evidence for an innate moral core. *Curr. Dir. Psychol. Sci.* 22, 186–193. doi: 10.1177/0963721412470687

Hamlin, J. K., Wynn, K., and Bloom, P. (2007). Social evaluation by preverbal infants. *Nature* 450, 557–559.

Hauser, M. D. (2006a). *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong*. New York: HarperCollins.

Hauser, M. D. (2006b). The liver and the moral organ. *Scan* 1, 214–220. doi: 10.1093/scan/nsl026

Heil, J. (1981). Does cognitive psychology rest on a mistake? *Mind* 90, 321–342.

Hobson, P. (2002). *The Cradle of Thought*. London: Macmillan.

Holiday, A. (1988). *Moral Powers: Normative Necessity in Language and History*. Milton Park: Routledge.

Jopling, D. (1993). "Cognitive science, other minds, and the philosophy of dialogue," in *The Perceived Self*, ed. U. Neisser (Cambridge: MIT Press), 290–309. doi: 10.1017/cbo9780511664007.017

Kenny, A. (1991). "The homunculus fallacy," in *Investigating Psychology: Sciences of the Mind After Wittgenstein*, ed. J. Hyman (London: Routledge), 155–165. (original work published 1971).

Kitchener, R. F. (2004). "Piaget's social epistemology," in *Social Interaction and The Development of Knowledge*, eds J. I. M. Carpendale and U. Müller (Mahwah: Erlbaum), 45–66.

Krettenauer, T. (2013). "Revisiting the moral self-construct: developmental perspectives on moral selfhood," in *Self-Regulation and Autonomy*, eds B. Sokol, U. F. Grouzet, and U. Müller (Cambridge: Cambridge University Press), 115–140.

Mareschal, D., Johnson, M. H., Sirois, S., Spratling, M. W., Thomas, M. S. C., and Westermann, G. (2007). *Neuroconstructivism: How the Brain Constructs Cognition*, Vol. 1. New York: Oxford University Press.

Margolis, E., and Laurence, S. (2013). In defense of nativism. *Philos. Stud.* 165, 693–718. doi: 10.1007/s11098-012-9972-x

Mcquaid, N., Bibok, M., and Carpendale, J. I. M. (2009). Relationship between maternal contingent responsiveness and infant social expectation. *Infancy* 14, 390–401. doi: 10.1080/15250000902839955

Mead, G. H. (1934). *Mind, Self and Society: From the Standpoint of a Social Behaviorist*. Chicago: University of Chicago Press.

Meaney, M. J. (2010). Epigenetics and the biological definition of gene x environment interactions. *Child Dev.* 81, 41–79. doi: 10.1111/j.1467-8624.2009.01381.x

Mikhail, J. (2007). Universal moral grammar: theory, evidence and the future. *Trends Cogn. Sci.* 11, 143–152. doi: 10.1016/j.tics.2006.12.007

Mikhail, J. (2020). "Moral intuitions and moral nativism," in *The Oxford Handbook of Moral Psychology*, eds M. Vargas and J. Doris (Oxford: Oxford University Press).

O'Connor, T. (2020). *Emergent Properties, The Stanford Encyclopedia of Philosophy*. Available online at: https://plato.stanford.edu/archives/fall2020/entries/properties-emergent/ (accessed October 5, 2021).

Overton, W. F. (2015). "Processes, relations, and relational-developmental-systems," in *Handbook of Child Psychology and Developmental Science*, 7th Edn, eds W. F. Overton and P. C. M. Molenaar (Hoboken: Wiley Blackwell), 9–62.

Piaget, J. (1936/1952). *The Origins of Intelligence in Children*. New York: International Universities Press.

Piaget, J. (1932/1965). *The Moral Judgment of the Child*. New York: The Free Press.

Piaget, J. (1970). *Genetic Epistemology*. London: Norton.

Piaget, J. (1971). *Biology and Knowledge*. Chicago: University of Chicago Press.

Piaget, J. (1970/1972). *Psychology and Epistemology: Towards a Theory of Knowledge*. New York: Viking Compass Book.

Portmann, A. (1944/1990). *A Zoologist Looks at Humankind*. New York: Columbia University Press.

Prinz, J. (2009). "Against moral nativism," in *Stich and His Critics*, eds D. Murphy and M. Bishop (Malden: Wiley-Blackwell), 167–189.

Racine, T. P., and Carpendale, J. I. M. (2008). "The embodiment of mental states," in *Developmental Perspectives on Embodiment and Consciousness*, eds W. F. Overton, U. Müller, and J. Newman (Mahwah: Erlbaum), 159–190.

Reddy, V. (2003). On being the object of attention: implications for self-other consciousness. *Trends Cogn. Sci.* 7, 397–402. doi: 10.1016/s1364-6613(03)00191-8

Rheingold, H. L. (1982). Little children's participation in the work of adults, a nascent prosocial behavior. *Child Dev.* 53, 114–125. doi: 10.2307/1129643

Spaemann, R. (2006). *Persons: The Difference Between 'Someone' and 'Something'*. New York: Oxford University Press.

Sterelny, K. (2010). Moral nativism: a skeptical response. *Mind Lang.* 25, 279–297.

Stiles, J. (2009). On genes, brains, and behavior: why should developmental psychologists care about brain development? *Child Dev. Persp.* 3, 196–202. doi: 10.1111/j.1365-2826.2011.02194.x

Stiles, J., Brown, T. T., Haist, F., and Jernigan, T. L. (2015). "Brain and cognitive development," in *Handbook of Child Psychology and Developmental Science, Cognitive Processes*, 7th Edn, eds L. Liben and U. Müller (Hoboken: Wiley Blackwell), 9–62.

Tomasello, M. (2020). The moral psychology of obligation. *Behav. Brain Sci.* 43:e56. doi: 10.1017/S0140525X19001742

Turnbull, W. (2003). *Language in Action: Psychological Models of Conversation*. East Sussex: Psychology Press.

Warneken, F. (2016). Insights into the biological foundation of human altruistic sentiments. *Curr. Opin. Psychol.* 7, 51–56.

Winch, P. (1972). *Ethics and Action*. Milton Park: Routledge and Kegan Paul.

Wittgenstein, L. (1953/2009). "Philosophical investigations," in *Philosophische Untersuchungen*, 4th Edn, Trans G. E. M. Anscombe, P. M. S. Hacker, and J. Schulte (Hoboken: Wiley-Blackwell).

Wright, D. (1982a). "Piaget's theory of moral development," in *Jean Piaget: Consensus and Controversy*, eds S. Modgil and C. Modgil (London: Holt, Rinehart & Winston), 207–217.

Wright, D. (1982b). Piaget's theory of practical morality. *Br. J. Psychol.* 73, 279–283.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 17 January 2022
doi: 10.3389/fpsyg.2021.756350

Development and Validation of the Ethnic Moral Disengagement Scale

Maria Grazia Lo Cricchio[1], Federica Stefanelli[2], Benedetta E. Palladino[2], Marinella Paciello[3] and Ersilia Menesini[2]*

[1] Department of Humanities (DiSU), University of Basilicata, Potenza, Italy, [2] Department of Education, Languages, Intercultures, Literature and Psychology, University of Florence, Florence, Italy, [3] Uninettuno Telematic International University, Rome, Italy

Research has underlined that moral disengagement processes, by which people switch off their moral values and act aggressively without experiencing guilt, are highly connected with contextual factors. However, research on situational variations in moral disengagement is limited, especially considering the associations with characteristics such as the ethnic origin of potential victims. The general aim of the present study was to develop a brief, specific measure of ethnic moral disengagement able to catch individual justification used in the case of ethnic bullying and cyberbullying, and test its validity and reliability. An eight items scale was developed and administered in study 1, in a sample of 961students attending several Italian high schools (53.5% female; Mage 15 years). Considering the results of the CFA, we modified one of the items and the scale was administered again, in a second sample of 1,229 students (49.9% female; Mage 15.62 years) in study 2. A one-factor model of ethnic moral disengagement fit the data well and internal consistency showed to be good. As an additional step, we found that the model was invariant across Italian adolescents and youths with a different ethnic or culture of origin (having at least one parent born abroad) strengthened our confidence regarding the factorial integrity of the scale. Last, the scale showed to be positively associated with ethnic bullying and cyberbullying. Generally, findings suggested that the Ethnic Moral Disengagement scale can be a useful tool for those interested in measuring moral disengagement and evaluating how it impacts bullying and cyberbullying of minority groups.

Keywords: moral disengagement, ethnicity, ethnic bullying, ethnic cyberbullying, scale development

OPEN ACCESS

Edited by:
Robert Thornberg, Linköping University, Sweden

Reviewed by:
David Álvarez-García, University of Oviedo, Spain
Christina Salmivalli, University of Turku, Finland

****Correspondence:***
Federica Stefanelli
federica.stefanelli@unifi.it

Specialty section:
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

Received: *10 August 2021*
Accepted: *10 December 2021*
Published: *17 January 2022*

Citation:
Lo Cricchio MG, Stefanelli F, Palladino BE, Paciello M and Menesini E (2022) Development and Validation of the Ethnic Moral Disengagement Scale. Front. Psychol. 12:756350. doi: 10.3389/fpsyg.2021.756350

INTRODUCTION

Globalization has increased the movement of many people from one country to another, thus promoting processes of migration (Fandrem et al., 2012). In this context, it must be considered as pivotal to gaining a better understanding of the factors which encourage positive intercultural relationships, thus reducing intolerance and discriminatory behaviors. Since mechanisms of Moral Disengagement (MD) are particularly informative with respect to discriminatory and racial behaviors (Faulkner and Bliuc, 2016; D'Errico and Paciello, 2018), the present study is aimed at contributing to the literature concerning this issue, by presenting a new scale aimed at measuring Ethnic MD (EMD).

As the moral self develops, individuals adopt standards of right and wrong that will guide their moral conduct. According to theory of moral agency Bandura's (1991, 2016), moral self-regulation

processes promote ethical conduct, and prevent unethical behaviors, by means of proactive or inhibitive mechanisms. The proactive process supports ethical conduct by regulating how behavior fits with personal and social standards, while the inhibitive process impedes negative actions by seeing them as ethically and socially punishable. So, when individuals engage in moral conduct, they may feel guilty or proud, depending on these processes of self-monitoring and judgment. However, moral self-regulation does not always work in a consistent manner (Bandura, 2015), and, under specific circumstances, certain cognitive practices lead an individual to disengage from their own moral principles, and to behave unfairly. These processes have been proposed as the mechanisms of MD, which work by restructuring the four dimensions of behavior representation, or *loci* of cognitive restructuring: behavior, agency, consequences, and victims (Bandura, 1991).

The *behavior locus* refers to the maneuvers focused on changing the meaning of harmful conducts, and it includes moral justification (the cognitive redefinition of negative behaviors as respectable), euphemistic labelling (the use of language that cognitively masks blameworthy actions as less harmful), and advantageous comparison (comparing negative behaviors with more unacceptable behaviors, thus making them appear better or less severe). The *locus of agency* refers to mechanisms aimed at avoiding personal responsibility, and it includes displacement of responsibility (viewing one's own actions as the result of social pressures and thus not under one's personal responsibility), and diffusion of responsibility (when duty is shared with others, thus reducing personal responsibility and motivation to action). The *locus of consequences* refers to processes aimed at altering one's perception of the effects of their own behavior, by disregarding or distorting its results (avoiding or cognitively minimizing the harm caused by bad acts). Finally, the *victims' locus* of redefinition refers to attempts to displace responsibility onto the victim *via* mechanisms of dehumanization (depriving victims of human qualities or attributing animalistic characteristics to them) and by attribution of blame (attributing victims the fault of injurious or provocative conducts). By using these MD processes, people can concretely switch off their moral values and act wickedly and aggressively without experiencing shame, guilt, or blameworthiness (Bandura, 1991; Paciello et al., 2008). Despite the different dimensions and *loci*, Bandura (1991) suggested that all mechanisms of MD are part of a single construct, and that MD maneuvers are only diverse ways to pursue the same and unique aim of decreasing guilt for one's detrimental conduct. This theoretical idea has been confirmed by different studies in which MD was evaluated with diverse scales, including items measuring each mechanism on MD. Findings have shown a single factor structure for this construct, when it was measured by one item for each dimension (8-item scales; Boardley and Kavussanu, 2008; Lucidi et al., 2008; Moore et al., 2012) or a common, second order, latent variable when more items were included (Bandura et al., 1996, 2001; Caprara et al., 1996; Pelton et al., 2004).

Research has shown that MD is strictly linked to aggressive behavior, including traditional bullying in schools (e.g., Bandura et al., 2001; Gini et al., 2014; Kowalski et al., 2014), and cyberbullying (e.g., Lo Cricchio et al., 2021). In particular, literature underlined that bullying perpetrators are more likely to score higher in MD than those not involved in bullying (Menesini et al., 2003; Caravita et al., 2012; Thornberg and Jungert, 2013). Moreover, bystanders with higher levels of MD are less likely to defend the victims when witnessing episodes of bullying (Gini, 2006; Obermann, 2011; Caravita et al., 2012; Thornberg and Jungert, 2014).

MD must be considered a product of the reciprocal interaction between individual and social and/or situational factors: it is not a trait or a disposition, but a process that can be selectively activated under different conditions (Bandura, 1999, 2016). Nevertheless, in a recent meta-analysis concerning the association between MD and bullying, Killer et al. (2019) concluded that there is a lack of investigation of the broader impact of these situational contexts, and underlined the need for further research into how MD and contextual variables may interact and explain aggressive and bullying behavior. Studies suggest, in fact, that the associations between MD and bullying can be affected by specific factors, such as the characteristics of the victims and the relational context (Thornberg et al., 2020). As in the case of discriminative aggression, it seems plausible to expect that the ethnic origin of potential victims may play a role.

Previous research on anti-immigrant attitudes and prejudicial bullying behaviors indicates that MD may be important to explain why some youths perpetrate aggression toward their peers with different ethnic or cultural backgrounds. For example, it has been suggested that the likelihood of harassing immigrants is increased by negative attitudes and preconceptions toward them, and by having strong beliefs that immigrants deserve any negative treatment they receive (Bayram Özdemir et al., 2020).

Within the bullying context, research suggested that bullies' perception of victims as different because of their immigration status can increase the risk of harassment (Caravita et al., 2019). In fact, the more immigrants are perceived as different and not fitting peer group norms and characteristics, the more this can cause their peers to mark them as dangerous or deviant, eliciting MD processes in which bullies justify themselves as acting to protect their group from the aberrant outsiders. Furthermore, prejudices and stereotypes might activate specific MD processes such as dehumanization of the victim, through which bullies become more compelled and disposed toward acting cruelly and harshly toward victims who are ethnically diverse (Webster and Saucier, 2015; Bandura, 2016).

However, one of the main limits of the available knowledge on these issues is related to how MD has been measured. Despite how Bandura (1977) claimed that the closer the cognitions are to the actions, the stronger the explicative power of the measure, in the majority of studies, MD has been assessed as a general disposition by using items such as those of the traditional measure of Bandura et al. (1996). These measures usually ask to express personal opinions concerning negative behaviors toward people, without considering the contextual factors, such as who these individuals are, and, in particular, how the different ethnic origins of potential victims can influence adolescents' cognition and behavior.

Even if some scholars have developed more specific measures of MD, such as for cyber (Paciello et al., 2020), civil

(Caprara et al., 2009), organizational (Moore et al., 2012), and sports contexts (Boardley and Kavussanu, 2008), to the best of our knowledge, only one study has proposed to evaluate MD in interethnic relations with *ad hoc* measures (Caravita et al., 2019). In particular, Caravita et al. (2019) used vignettes in which the target of bullying is a non-immigrant vs an immigrant new classmate. For each scenario, participants must answer sixteen items, for a total of thirty-two items. The use of vignettes can have potential strengths, but also some limitations. The main strength is related to the possibility of having direct examples of ethnic bullying episodes, but at the same time, the specificity of the situations may limit the reported reactions to those contexts, without providing a more general perception of ethnic bullying. Besides, the vignettes may be more appropriate for younger students, whereas a general brief scale may be more suitable for older students and, in general for school administration purposes, where limited time is often requested by teachers. Therefore, a brief, valid, and reliable measure of EMD would be an added value in the research field and serve scholars who wish to incorporate a specific MD measure in a multivariate investigation.

The Present Study

To sum it up, the literature indicates that MD mechanisms are highly connected with contexts and to interethnic relationships and dynamics. However, research on situational variations in MD is limited, especially considering characteristics such as the ethnic background of potential victims. Hence, the general aim of the present study was to address this gap in knowledge by developing a specific and short measure of EMD, able to catch individual justifications and interpersonal mechanisms used in cases of ethnic bullying and victimization. In developing the new scale, we followed recommendations for constructing and revising scales (e.g., Smith et al., 2000). Generally, (a) the factorial dimensionality of the instrument must be examined by factor analyses (CFA); (b) factor must demonstrate standards of reliability; (c) the newly developed instrument must be administered to a different sample from the one used when the scale was originally constructed; (d) the factor structure and the reliability of its factor must be confirmed; and (e) the newly developed scale must be validated. We followed these guiding standards in carrying out two studies. In particular, study 1 aims at developing and testing the factor structure of the scale, whereas study 2 aims at evaluating its validity, reliability and structural invariance.

STUDY 1

The first study aims at developing and evaluating the factor structure of the scale of EMD in a sample of students attending Italian high schools.

Method

Participants and Procedure

Participants were 1,311 students nested in 58 classes of 13 Secondary Schools in Italy, all of which attended Lyceum, Technical or Vocational high schools (grade 9). Before questionnaire administration, informed consent, consisting of initial approval by the School Principal and the class council, was requested. Once permission was gained from schools, informative letters were sent to all students and to their parents, explaining the study aims and requesting the parents' consent for their child's participation. 1,153 students were present at school on the day of the data collection but data were retrieved only by 961 students because 192 did not have parental authorization. Of the 961 students who filled the questionnaire, 437 (46,5% of the whole sample) were male, while 503 (53,5% of the whole sample) were female (21 students did not answer the question about gender). Students' mean age was 15 years old with a standard deviation of 0.60 (MAXage = 18 years old; MINage = 11 years old). Most of the participants were Italian, having both parents born in Italy (71.1% of the whole sample). 278 students (28.9% of the whole sample) had a different ethnic or culture of origin, having at least one parent born abroad. The students with an immigrant background came from various countries of the world, such as China (4.2 %), Albania (2.7 %), Morocco (1.3 %), Romania (1.1 %), and other countries (19.6 %).

Before collecting data, institutional ethical committee approval was obtained for the study procedure. The schools that took part in the research were recruited for a voluntary census. Specifically, the call for participation was extended to all the high schools in several Italian provinces. The study survey was administered in January 2020 by trained assistants during school hours. Of the 961 high school students who participated in the study, 509 filled the paper version of the questionnaire, while 452 filled the online one, using school computers.

Measure

Ethnic Moral Disengagement Scale

Prior to all steps, we developed an initial set of eight items to measure MD related to ethnic minority potential victims. The general references for this aim were: (a) the theory of MD proposed by Bandura (1991); and (b) the use of existing items concerning MD, such as those of the Online MD (Paciello et al., 2020). In developing the scale, each MD process was represented with one item. This initial pool of eight items was reviewed by a professional with research expertise relating to the fields of ethnicity and MD. Items were adjusted following their feedback, resulting in a final 8-item scale that is presented in **Table 1**. The initial set of eight items to measure MD related to ethnic minority potential victims was administered to participants. Each item was evaluated along a 5-point scale (strongly disagree, disagree, quite agree, agree, and totally agree).

Analytic Plan

As a preliminary step, we looked at missing values in the matrix. Thus, we tested if missing data occurred completely at random (MCAR) using Little's test analysis. Little (1998) has proposed a statistical test of the MCAR assumption, which is a chi-square test. Significant chi-square values indicate that the data are not MCAR.

After controlling for MCAR assumption, firstly, we explored items' distributions and correlations by performing a descriptive

TABLE 1 | EMD scale items and mechanisms of moral disengagement.

EMD items	MD mechanisms
1. Bullying children of different ethnicities or origins is just a way to spend time with friends	Euphemistic labeling
2. There is no reason why boys/girls of different ethnicities or origins get offended when they are teased, because this is still a way to pay attention to them.	Disregarding/Distorting consequences
3. If any boy/girl of different ethnicity or origin is treated badly by others, it is because he/she is the first to behave badly toward Italians.	Attribution of blame
4. It is right to exclude boys/girls of different ethnicity or origin to defend our culture	Moral justification
5. People of different ethnicities or origins who are mistreated usually deserve it because they are like beasts	Dehumanization
6. It is not serious to insult someone of a different ethnicity or origin since beating them would be even worse	Advantageous comparison
7. If most parents provide a bad example, it is not the children's fault if they denigrate those of a different ethnicity or origin.	Displacement of responsibility
8. Young people should not be blamed for insulting those of a different ethnicity or origin since most Italians do the same	Diffusion of responsibility

analysis. Not all the items included in the EMD scale presented symmetric distribution. Thus, we proceeded to examine the factorial structure of the EMD scale, using robust methods for the estimation of the parameters. In particular, following Bandura's theoretical model, and literature indications, we tested the predicted one-factorial structure of the EMD scale, performing a Confirmatory Factor Analysis using the R packages Lavaan (Rosseel, 2012). Specifically, we ran a monofactorial model (latent factor: EMD, items: 1, 2, 3, 4, 5, 6, 7, and 8). The model was evaluated according to the following indices: the chi-square ($\chi 2$) statistic, the root-mean-squared error of approximation (RMSEA), the comparative fit index (CFI), the Tucker-Lewis Index (TLI), and the standardized root mean squared residual (SRMR). Recommended cut-off points for these measures are 0.08 (Brown and Cudek, 1993) or 0.06 (Hu and Bentler, 1998) for RMSEA, 0.90 or 0.95 for CFI and TLI (Bollen, 1989) and 0.08 or 0.05 for SRMSR (Hu and Bentler, 1998). The cut-off used for the factor loading was 0.30 (Muthén and Muthén, 2007). Finally, to evaluate the reliability of the scales, we analyzed the internal consistency of the dimension by means of Cronbach's alpha. The analyses were conducted *via* R Studio (R Studio Team, 2020).

RESULTS

Descriptive Statistics

Descriptive statistics and bivariate associations between the eight items of the EMD scale are reported in **Table 2**.

As **Table 2** shows, not all the items included in the EMDS presented symmetric distribution. Indeed, the items Skewness indexes range from 0.44 to 2.70, while the Kurtosis indexes range from −0.78 to 7.53. All the items included in the scale are correlated with each other, but not too strongly.

Factorial Structure of the Ethnic Moral Disengagement Scale

Our data were missing completely at random as indicated by the non-significant Little's (1998) MCAR test [$\chi 2(81) = 96.75$, $p = 0.111$]. Thus, we proceeded by using the full information maximum likelihood approach (FIML) (Enders and Bandalos, 2001) for the estimation of missing data in our matrix.

The model fit indices were all satisfactory except for Chi Squared p, which is especially sensitive to sample size [$\chi 2(20) = 71.94$, $p < 0.001$]. Specifically, RMSEA, SRMR, TLI, and CFI had optimal values in the monofactorial solution (RMSEA = 0.075; SRMR = 0.041; TLI = 0.921; CFI = 0.944).

The standardized estimates are reported in **Figure 1**. Not all factor loadings were satisfactory. Indeed, while the factors loadings of the items 1, 2, 3, 4, 5, 6, and 8 were ranged from βItem1 = 0.48 to βItem5-6 = 0.72, the factor loading for the item 7 was βItem7 = 0.23 (SE = 0.139; $p < 0.001$). We used Cronbach's alpha coefficients to calculate the scale's internal consistency. EMD scale showed decent reliability [$\alpha = 0.77$; 95% CI (0.75–0.80)].

STUDY 2

The first study revealed some limitations of the scale for the assessment of EMD. In particular, Item 7 was not sufficiently adequate for measuring the latent factor EMD. Therefore, we reformulated it and we administered the EMD scale in a different, independent, sample to test its psychometric characteristics. We examined its factorial structure, internal consistency, and ethnic measurement invariance. We also evaluated the discriminant validity of the scale. To do so, we examined whether the score on its factor was associated with ethnic bullying and cyberbullying behaviors. Specifically, we expected that EMD to be positively correlated with both traditional and virtual forms of prejudicial ethnic bullying behaviors.

Participants and Procedure

Before collecting data, institutional ethical committee approval was obtained for the study procedure. We recruited a new larger sample composed of 1,636 students nested in 77 classes of 11 Secondary Schools in Italy. All of the participants attended Lyceum, Technical or Vocational high school (grades 9, or 10). Before questionnaire administration, informed consent, consisting of initial approval by the School Principal and the class council, was requested. Once permission was gained from schools, informative letters were sent to all students and to their parents, explaining the study aims, and requesting the parents' consent for their child's participation. The study survey was administered from February to March 2021. During that period, due to COVID-19 pandemic, schools were closed and students studying from home, so we had them fill the questionnaire online, under the supervision of trained assistants.

On the day of the data collection, 203 students were not following online classes. Of the remaining 1,433 high school students, 67 did not give their authorization for participation in the study, and 136 did not send their questionnaire answers,

TABLE 2 | Descriptive statistics of the items of the EMD scale (study 1).

	N	Mean	SD	Skewness	Kurtosis	2	3	4	5	6	7	8
1. Item 1	895	1.33	0.72	2.54	6.77	0.36**	0.26**	0.39**	0.34**	0.31**	0.07**	0.29**
2. Item 2	891	1.62	1.0	1.69	2.19	–	0.35**	0.35**	0.36**	0.35**	0.19**	0.32**
3. Item 3	895	1.80	0.96	1.20	1.01	–	–	0.40**	0.36**	0.38**	0.19**	0.31**
4. Item 4	889	1.37	0.76	2.32	5.20	–	–	–	0.61**	0.58**	0.11**	0.38**
5. Item 5	892	1.31	0.72	2.70	7.53	–	–	–	–	0.50**	0.13**	0.37**
6. Item 6	891	1.49	0.93	2.15	4.24	–	–	–	–	–	0.14**	0.47**
7. Item 7	887	2.50	1.25	0.44	−0.78	–	–	–	–	–	–	−26**
8. Item 8	890	1.69	1.07	1.59	1.76	–	–	–	–	–	–	–

$^{**}p < 0.01$.

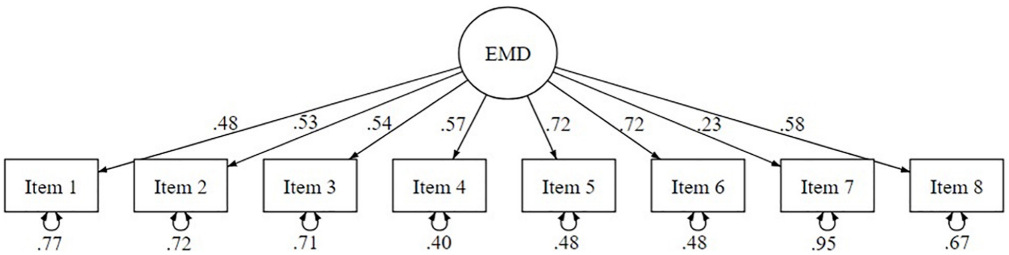


FIGURE 1 | Graphical representation of the Ethnic Moral Disengagement monofactorial model (study 1).

because of problems with their internet connections. Moreover, we decided not to consider the questionnaire responses provided by one student because he was not an adolescent (26 years old). Thus, overall, 1,229 students filled the questionnaire (49.9% female, and 50.1% male). Students' mean age was 15.62 years old (SD = 0.72; MAXage = 18 years old; MINage = 14 years old). While most of the participants were Italian, with both parents born in Italy, 275 students had an immigrant background, having at least one parent born abroad (416 students did not answer either of the questions about their parents' nationality). Specifically, not considering the missing data, 10.4% of the students' mothers come from other European countries, mostly from Albania (4.1%) and Romania (2.6%), while 9.1% come from non-European countries, mostly from Morocco (2%) and China (1.5%). On the other hand, 8.1% of the students' fathers come from other European countries, mostly from Albania (4.1%) and Romania (2%), while 10% were from non-European countries, mostly from Morocco (2.1%) and China (1.4%). In the following paragraphs, the label "with a different ethnic/culture of origin" will refer to students of whom at least one parent was born abroad. On the contrary, the label "Italian students" will refer to youths whose parents were both born in Italy.

Measure

Ethnic Moral Disengagement Scale

We administered the same 8-items scale of study 1, with only the modified version of item 7. In fact, since this item resulted as less adequate for measuring the latent factor EMD, we reformulated it with the help of an expert in the field. Maybe, the original item was not sufficiently unequivocal. Therefore, we changed it to clarify its meaning and in order to simplify its understanding. In particular, we reformed the item from "If most parents provide a bad example, it is not the children's fault if they denigrate those of a different ethnicity or origin" to "It is not the child's fault if they exclude those of a different ethnicity/origin, if most parents set a poor example."

Ethnic Bullying

We administered a modified version of the Florence Bullying Scale (Palladino et al., 2016, 2020) that ask how often, in the previous couple of months, students behaved like bullies, attacking other students with an immigrant background physically, verbally, and or indirectly (i.e., "I beat someone up because of his/her culture or country of origin"). A definition of bullying introduced the scale, consisting of four items. Each item was evaluated along a 5-point scale from "never" to "several times a week." Within our data, the scale presents acceptable internal consistency [$\alpha = 0.89$, 95% CI (0.88–0.90)].

Ethnic Cyberbullying

We used a modified version of the Florence Bullying Scale (Palladino et al., 2016, 2020) that asks how often in the previous couple of months students behaved like cyber bullies, excluding other students with a different ethnic or culture of origin from the online group, and/or taking their personal information to reuse later, and/or sending embarrassing photo or videos, and/or sending threats and insults on the Internet (i.e., "In the last 2 or 3 months, how often have you sent threats and insults on the internet to someone because of his/her culture or country of origin?). A definition of cyberbullying introduced the scale, consisting of four items. Each item was evaluated along a 5-point scale from "never" to "several times a week." Within our data,

the scale presents acceptable internal consistency [α = 0.81, 95% CI (0.79–0.82)].

Analytic Plan

As a preliminary step, we checked if missing values occurred completely at random (MCAR) using Little's test analysis (Little, 1998). Since our data were missing completely at random, we proceeded using the full information maximum likelihood approach (FIML) (Enders and Bandalos, 2001) for their estimation. After controlling for the MCAR assumption, firstly, we explored the distribution and the correlations of the items included in the EMD scale performing descriptive analysis. Not all the items presented a symmetric distribution, thus, we proceed to test our model using the robust method.

To examine the hypothesized one-factor structure of the EMD scale, we performed a Confirmatory Factor Analysis using the R packages Lavaan (Rosseel, 2012) testing a monofactorial model. The model was evaluated according to the same indexes and the same recommended cut-off used in the first study: are 0.08 (Brown and Cudek, 1993) or 0.06 (Hu and Bentler, 1998) for RMSEA, 0.90 or 0.95 for CFI and TLI (Bollen, 1989) and 0.08 or 0.05 for SRMSR (Hu and Bentler, 1998). We also evaluate the reliability of the scales, analyzing the internal consistency by means of Cronbach's alpha.

As a second step, we tested for measurement invariance to verify whether the instrument has the same psychometric properties across the majority (i.e., Italians) and the minority (i.e., students with different ethnic or culture of origin). We followed the procedures described by Meredith (1993), and Widaman and Reise (1997). The sequence of invariance testing starts from the configural invariance, which involves running a model in which all parameters are estimated freely (A configural – 1st level). At this step of measurement invariance, only the similarity across groups of the overall parameters' pattern is evaluated. This provides indications about the ability of the original model to fit the data in each group (here, Italian students and students with different ethnic or culture of origin) without invariance constraints. The invariance measure proceeds step by step, comparing increasingly restricted models. The 2nd level of invariance involves constraining factor loadings over the groups as invariants (B metric–2nd level). The third level of invariance involves a stricter model in which both factorial loadings and intercepts are constrained across groups (C scalar–3rd level). The fourth level of invariance is tested at residual variance invariance (D strict—4th level). Finally, the fifth level of invariance involves a model in which both factor loadings, intercepts, residual variance, and factor variance are constrained to be equal across groups (E factor variance—5th level). To summarize, each level of invariance involves an even more restricted model. Each one of these models, from the least (1st level) to the most restrictive one (5th level), is nested in the original model. Moreover, we tested a very strict model, in which the means equality was also imposed across the two different ethnic groups (F latent mean 6th level) (Vandenberg and Lance, 2000).

The goal of each level of the measurement invariance is to make the model fit, not to worsens it by constraining parameters equally across groups. Thus, at each level of the analysis, we tested whether subtracting parameters worsened the model fit by controlling the change in the fit indices RMSEA and CFI. It has been suggested (Chen, 2007) that support for invariance across groups requires that at each step of the analysis the CFI is not worse more than—0.01 across models and RMSEA is no worse than 0.015 across models. We also considered the Akaike Information criterion (AIC) and Bayesian Information Criterion (BIC) in testing for the evidence of invariance (Vrieze, 2012): lower AIC and BIC value indicates a better trade-off between fit and complexity.

Finally, we evaluated the discriminant validity of the scale, checking whether the EMD scale was correlated with ethnic bullying and ethnic cyberbullying behaviors. The analyses were conducted by R Studio (R Studio Team, 2020).

RESULTS

Descriptive Statistics

Descriptive statistics and bivariate associations between the eight items of the EMD scale are reported in **Table 3**.

As **Table 3** shows, not all the items included in the EMDS presented symmetric distribution. Indeed, the items Skewness indexes range from 0.80 to 2.65, while the Kurtosis indexes range from −0.10 to 6.75. All the items included in the scale are correlated with each other, but not too strongly.

Factorial Structure of the Ethnic Moral Disengagement Scale and Its Discriminant Validity

STEP 1—CFA of the EMD scale—Our data were missing completely at random as indicated by the non-significant Little's (1998) MCAR test, $\chi 2(10) = 12$, $p = 0.284$. Thus, after the estimation of missing data, we proceed to testing a monofactorial measurement model (CFA). All CFA model fit indices were satisfactory except for Chi Squared p, which is especially sensitive to sample size [$\chi 2(20) = 68.84$, $p = 0.000$]. Specifically, RMSEA, SRMR, and CFI had optimal values (RMSEA = 0.046; SRMR = 0.031; TLI = 0.951; CFI = 0.965). Moreover, all factor loadings estimated for the monofactorial model varied from βItem7 = 0.42 to βItem5 = 0.82. After its reformulation, Item 7, which had not shown satisfactory saturation in the first study, adequately saturated the latent factor EMD (β = 0.42; SE = 0.090; $p < 0.001$). Using Cronbach's alpha coefficients to calculate the scale's internal consistency, we found that the EMD scale showed a decent reliability: α = 0.84, 95% CI (0.85–0.86). The standardized estimates are reported in **Figure 2**.

STEP 2- Ethnic EMD scale measurement invariance—In **Table 4** are reported the model's fit indices for the comparison from the less restricted model (A–Configural Invariance: all parameters that are freely estimated) to the more constrained one (F–Latent Mean Invariance).

The initial model A, that assessed configural invariance (Model A), resulted in an acceptable fit, as well as the second model B, testing the full metric invariance (Model B). Given

TABLE 3 | Descriptive statistics of the items of the EMD scale (study 2).

	N	Mean	SD	Skewness	Kurtosis	2	3	4	5	6	7	8
1. Item 1	1,172	1.29	0.72	2.61	6.75	0.41**	0.38**	0.46**	0.48**	0.44**	0.25**	0.36**
2. Item 2	1,172	1.47	0.92	2.06	3.66	–	0.39**	0.45**	0.49**	0.43**	0.28**	0.4**
3. Item 3	1,171	1.67	0.90	1.44	1.98	–	–	0.53**	0.49**	0.45**	0.31**	0.39**
4. Item 4	1,171	1.36	0.77	2.26	4.84	–	–	–	0.68**	0.62**	0.27**	0.49**
5. Item 5	1,171	0.130	0.75	2.65	6.73	–	–	–	–	0.62**	0.29**	0.05**
6. Item 6	1,171	1.39	0.82	2.35	5.35	–	–	–	–	–	0.34**	0.46**
7. Item 7	1,171	1.98	1.04	0.80	−0.10	–	–	–	–	–	–	0.42**
8. Item 8	1,171	1.59	0.91	1.53	1.73	–	–	–	–	–	–	–

$^{**}p < 0.01$.

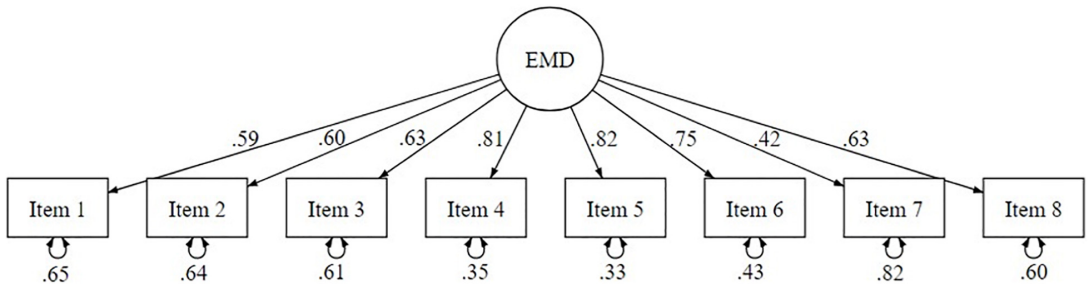


FIGURE 2 | Graphical representation of the Ethnic Moral Disengagement monofactorial model (study 2).

TABLE 4 | Tests results for measurement invariance of EMD scale across ethnicity (Italians *N* = 904; Students with an immigrant background *N* = 275).

EMD models		Compared model	χ^2 (df)	RMSEA	ΔRMSEA	CFI	ΔCFI	AIC	ΔAIC	BIC	ΔBIC
A	Configural Invariance		180.02 (40)	0.068		0.969		19,707.83		19,950.02	
B	Metric Invariance	A	184.43 (47)	0.060	−0.008	0.972	0.003	19,698.24	−9.58	19,905.12	−44.91
C	Scalar Invariance	B	189.19 (54)	0.055	−0.005	0.972	0.000	19,688.99	−9.24	19,860.55	−44.56
D	Strict Invariance	C	248.35 (62)	0.057	0.002	0.965	−0.007	19,732.15	43.16	19,863.35	2.79
E	Variance Invariance	D	255.73 (63)	0.058	0.001	0.964	−0.001	19,737.53	5.37	19,863.68	0.33
F	Latent Mean Invariance	E	258.53 (64)	0.058	0.000	0.963	−0.001	19,738.34	0.81	19,859.44	−4.24

that Model B leads to an acceptable CFI and RMSEA change compared to the configural invariance model (Model A), Metric Invariance was confirmed. For the third step of the measurement invariance test, the full scalar invariant model (Model C) resulted in an acceptable fit. Since Model C not too worsens Model B fit indices, Scalar invariance across ethnic backgrounds was confirmed. The fourth and fifth steps also, respectively, testing full strict invariance (Model D) and the factor variance invariance (Model E), yielded acceptable fits. Given that adding restrictions to the models, CFI and RMSEA fit indices do not particularly change, both the full strict invariance and the variance invariance across ethnic backgrounds were confirmed. Finally, the last step of invariance, testing latent mean invariance (Model F) showed an acceptable fit without significantly changing Model E fit indices. We may conclude that also latent mean invariance was also confirmed across the ethnic backgrounds.

STEP 3- Discriminant Validity—**Table 5** shows the Pearson's r correlation coefficients between EMD and Ethnic Bullying and between EMD and Ethnic Cyberbullying. EMD results significantly and positively correlated with both behaviors (Ethnic Bullying $r = 0.160$; $p < 0.001$; Ethnic Cyberbullying $r = 0.185$; $p < 0.001$).

GENERAL DISCUSSION

Explanations for reprehensible conduct may reside in specific cognitive processes, which have been referred to as MD mechanisms, and which explain why common people are able to engage in unethical conduct, without experiencing apparent guilt (Bandura, 1991). The use of MD has been documented in several contexts, and it has been highlighted that it plays an important role in antisocial and aggressive behavior (Bandura et al., 1996, 2001; Menesini et al., 2003; Osofsky et al., 2005; Paciello et al., 2008; Fida et al., 2015). Despite its well-known importance for explaining aggressive conduct, such as bullying

TABLE 5 | Correlations between EMD and ethnic bullying and ethnic cyberbullying.

	N	Mean (SD)	2.	3.
1. EMD	1,170	2.41 (0.34)	0.160*	0.185*
2. Ethnic Bullying	1,227	1.40 (0.11)		
3. Ethnic Cyberbullying	1,227	1.40 (0.10)		

$^{*}p < 0.001$.

and cyberbullying (Gini et al., 2014; Lo Cricchio et al., 2021), the understanding of how MD operates within intercultural contexts remains at an early phase. Since MD mechanisms may be related to specific contextual characteristics, such as the ethnic or cultural origin of the potential victims, the investigation of EMD can be fundamental to prevent intolerance and discriminatory behaviors. However, to our knowledge, not much research has considered and measured the role of MD in the specific context of ethnic victims of bullying and cyberbullying episodes.

The aim of the present study was to develop a short, reliable, and valid scale for adolescents to assess MD in the case of ethnic minority potential victims. The conceptual referent theory of MD proposed by Bandura (1991) guided the development of the EMD scale, together with the use of some items concerning general MD. Each MD process was represented with one item. The initial pool of items was reviewed by an expert with research expertise related to the fields of ethnicity and MD. This guaranteed that items adequately represented the mechanisms they are planned to measure, and they were clearly phrased, brief and unequivocal. The scale was adjusted on the basis of this feedback, resulting in a final set of 8-items on a scale that was administered in two studies. In particular, study 1 was aimed at testing the initial hypothesized one-factor structure of the scale. However, study findings revealed some limitations of the scale deriving from the adequacy of Item 7 for measuring the latent factor of EMD. Therefore, we reformulated this item and administered this second version of the EMD scale in study 2 to test its factorial structure, internal consistency, and ethnic measurement invariance.

The results confirmed a one-factor model of EMD fitting the data well, with all fit indices being acceptable, the scale being internally consistent and reliable, and all items loading highly and signed onto the factor. These findings supported the presence of a single EMD factor, indicating that all mechanisms of MD are part of one general construct. This is consistent with Bandura's (1991) theorizing that the MD maneuvers are different methods of accomplishing the same task: to disengage moral limitations from harmful behavior and decrease guilt for such conducts. In addition, these results are in line with previous studies, such as that of Boardley and Kavussanu (2008), which found that even though the items of some scales describe different mechanisms, there is evidence for a one-dimensional structure of the MD processes. So, even when we consider specific ethnic aspects of MD, all items referring to the eight theoretical mechanisms can be perceived as components of a unique common dimension that makes people inclined to use mechanisms of MD in interactions with people with a different ethnic background. Additionally, this study provides evidence of the internal consistency of the scale, which showed to be good, confirming the conceptual sense of the factor.

The invariance of the model across Italian adolescents and adolescents of different ethnicities/cultures of origin (i.e., having at least one parent born abroad) was supported through the examination of unconstrained and constrained models in the second sample of Study 2. In particular, we tested invariance constraining the latent factor means to be equal across the two groups. Results indicated that the scale works in the same way with students with or without immigrant background (i.e., majority vs minority), strengthening our confidence regarding the factorial integrity of the scale.

In order to assess the discriminant validity of the measure, we examined the links between the EMD factor and ethnic behavior of bullying, and cyberbullying. Results were in the expected direction. EMD showed to be positively associated with bullying peers of a different ethnicity or culture of origin. Similarly, a positive correlation was found between ethnic MD and online bullying of ethnic victims. Generally, these findings are in line with previous research, in which general MD has been positively associated with higher risk of engagement in ethnic victimization (Bayram Özdemir et al., 2020) and online racist form of harassment (Faulkner and Bliuc, 2016). However, it is important to note that correlations in our study were lower than expected. These results seem in line with those of Caravita et al. (2019), who found a lower level of specific MD when the victim is an immigrant peer rather than when he/she is a member of the autochthonous group. It is possible that the higher likelihood that immigrant people have of being victims of bullying episodes over time, both in cyberspace and in real life, has increased young people's perception of ethnic bullying as more normative, and consequently, this has reduced individual's need to justify (using MD maneuvers) these types of misconducts. Despite this, the ability of the EMD scale to be linked to these theoretically related constructs supports its usefulness in future research, in which ethnic MD's role in ethnic bullying may be studied in association with other contextual and situational factors, such as the normativity of ethnic bullying behaviors.

One of the major benefits of the EMD scale is its brevity: the scale has both pragmatic power as well as sound psychometric properties. In light of these findings, the EMD scale appears to be a useful tool for those interested in measuring MD and predicting the occurrence of unethical or wrong behaviors toward victims belonging to minority groups.

Study Limitations, Future Directions, and Conclusion

Some limitations should be noted. First, the items were developed to be used with adolescent samples, therefore the measure is appropriate for this specific demographic. However, further research into the psychometric properties of the EMD scale with more diverse age populations is encouraged. Secondly, both samples of the studies were recruited from schools in Italy. Keeping in view the scope of this study, the samples were adequate. However, for future research it would be beneficial to include samples from other countries so as to increase its generalizability and external validity. Third, the correlational design did not permit us to examine the longitudinal trajectories of the EMD scores. Future studies may pursue the aim of evaluating the stability

of the measure over time. Last, we are aware that, despite the items of the EMD scale are derived from existing and validated measures, and are in line with the aim of measuring MD processes, there is the risk of legitimizing and/or reinforcing some prejudices toward immigrants. In light of these ethical concerns, our recommendation is to administer the scale together with others, which could highlight opposite attitudes and behaviors, such as tolerance toward diversity. Additionally, as previously stated, the scale has been developed in the Italian context, so it suits the language use of Italian adolescents. However, researchers from other countries, before using the scale, need to be aware of cultural and language differences and peculiarities in how adolescents talk in their everyday school contexts. As a consequence, in fact, there might be the need to adjust some of the items of the scale to better adapt to their specific ethical standards and beliefs.

Notwithstanding these limitations, the obtained findings start to shed light on the intricate aspects of ethnic MD as well as indicate that the EMD scale has considerable promise to be considered a useful measure to assess the related process in order to identify and prevent discriminative forms of aggression. It is easy to administer and it might attract a wide range of scientists, teachers, and educators who could take advantage from employing such a measure, which shows a balance between shortness and psychometric demandingness. In conclusion, we deem that future investigations on the adolescents' ethnic MD are necessary and the EMD scale can be particularly helpful in this research.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of the University of Florence. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

BP, ML, and EM contributed to the conception and design of the study. FS organized the database and performed the statistical analysis. ML and FS wrote the first draft of the manuscript. BP, EM, and MP contributed to the manuscript revision. All authors read and approved the submitted version of the manuscript.

FUNDING

The founder is PRIN: Progetti di ricerca di Rilevante Interesse Nazionale (Research Projects of Relevant National Interest)—Call 2017. Award Number: 20173E3Z7W_003 approved with "SH3"—n. 226 of 12/02/2019.

REFERENCES

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychol. Rev.* 84, 191–215. doi: 10.1037/0033-295X.84.2.191

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organ. Behav. Hum. Decis. Process.* 50, 248–287. doi: 10.1016/0749-5978(91)90022-L

Bandura, A. (1999). "Social cognitive theory of personality," in *Handbook Of Personality: Theory And Research*, eds L. A. Pervin and O. P. John (New York, NY: Guilford Press), 154–196.

Bandura, A. (2015). On deconstructing commentaries regarding alternative theories of self-regulation. *J. Manag.* 41, 1025–1044. doi: 10.1177/0149206315572826

Bandura, A. (2016). *Moral Disengagement: How People Do Harm And Live With Themselves*. New York, NY: Worth Publishers.

Bandura, A., Barbaranelli, C., Caprara, G. V., and Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *J. Pers. Soc. Psychol.* 71, 364–374. doi: 10.1037/0022-3514.71.2.364

Bandura, A., Barbaranelli, C., Caprara, G. V., and Pastorelli, C. (2001). Self-efficacy beliefs as shapers of children's aspirations and career trajectories. *Child Dev.* 72, 187–206. doi: 10.1111/1467-8624.00273

Bayram Özdemir, S., Giles, C., and Özdemir, M. (2020). Differences and similarities between perpetrators of ethnic and non-ethnicity-based victimization. *J. Youth Adolesc.* 49, 1805–1820. doi: 10.1007/s10964-020-01271-5

Boardley, I. D., and Kavussanu, M. (2008). The moral disengagement in sport scale–short. *J. Sports Sci.* 26, 1507–1517. doi: 10.1080/02640410802315054

Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociol. Methods Res.* 17, 303–316. doi: 10.1177/0049124189017003004

Brown, M. W., and Cudek, R. (1993). "Alternative ways of assessing model fit," in *Testing Structural Equation Models*, eds K. A. Bollen and J. S. Long (Newbury Park, CA: Sage), 136–162.

Caprara, G. V., Bandura, A., Barbaranelli, C., and Vicino, S. (1996). La misura del disimpegno morale [the assessment of moral disengagement]. *Rass. Psicol.* 13, 93–105.

Caprara, G. V., Fida, R., Vecchione, M., Tramontano, C., and Barbaranelli, C. (2009). Assessing civic moral disengagement: dimensionality and construct validity. *Pers. Individ. Dif.* 47, 504–509.

Caravita, S. C. S., Di Blasio, P., Strohmeier, D., and Salmivalli, C. (2019). Bullying immigrant versus non-immigrant peers: moral disengagement and participant roles. *J. Sch. Psychol.* 75, 119–133. doi: 10.1016/j.jsp.2019.07.005

Caravita, S. C. S., Gini, G., and Pozzoli, T. (2012). Main and moderated effects of moral cognition and status on bullying and defending. *Aggress. Behav.* 38, 456–468. doi: 10.1002/ab.21447

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct. Equ. Model. A Multidiscip. J.* 14, 464–504. doi: 10.1080/10705510701301834

D'Errico, F., and Paciello, M. (2018). Online moral disengagement and hostile emotions in discussions on hosting immigrants. *Internet Res.* 28, 1313–1335. doi: 10.1108/IntR-03-2017-0119

Enders, C. K., and Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Struct. Equ. Model.* 8, 430–457.

Fandrem, H., Strohmeier, D., and Jonsdottir, K. A. (2012). Peer groups and victimisation among native and immigrant adolescents in Norway. *Emot. Behav. Diffic.* 17, 273–285. doi: 10.1080/13632752.2012.704308

Faulkner, N., and Bliuc, A.-M. (2016). 'It's okay to be racist': moral disengagement in online discussions of racist incidents in Australia. *Ethn. Racial Stud.* 39, 2545–2563. doi: 10.1080/01419870.2016.1171370

Fida, R., Paciello, M., Tramontano, C., Fontaine, R. G., Barbaranelli, C., and Farnese, M. L. (2015). An integrative approach to understanding counterproductive work behavior: the roles of stressors, negative emotions, and moral disengagement. *J. Bus. Ethics* 130, 131–144. doi: 10.1007/s10551-014-2209-5

Gini, G. (2006). Social cognition and moral cognition in bullying: what's wrong? *Aggress. Behav.* 32, 528–539. doi: 10.1002/ab.20153

Gini, G., Pozzoli, T., and Hymel, S. (2014). Moral disengagement among children and youth: a meta-analytic review of links to aggressive behavior. *Aggress. Behav.* 40, 56–68. doi: 10.1002/ab.21502

Hu, L. T., and Bentler, P. M. (1998). Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. *Psychol. Methods* 3:424. doi: 10.1037/1082-989X.3.4.424

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

Kowalski, R. M., Giumetti, G. W., Schroeder, A. N., and Lattanner, M. R. (2014). Bullying in the digital age: a critical review and meta-analysis of cyberbullying research among youth. *Psychol. Bull.* 140, 1073–1137. doi: 10.1037/a0035618

Little, R. J. A. (1998). "Missing data," in *Encyclopedia of Biostatistics*, eds P. Armitage and T. Colton (Chichester: John Wiley and Sons).

Lo Cricchio, M. G., García-Poole, C., te Brinke, L. W., Bianchi, D., and Menesini, E. (2021). Moral disengagement and cyberbullying involvement: a systematic review. *Eur. J. Behav. Dev.* 18, 271–311. doi: 10.1080/17405629.2020.1782186

Lucidi, F., Zelli, A., Mallia, L., Grano, C., Russo, P. M., and Violani, C. (2008). The social- cognitive mechanisms regulating adolescents' use of doping substances. *J. Sports Sci.* 26:456. doi: 10.1080/02640410701579370

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

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika* 58, 525–543.

Moore, C., Detert, J. R., Klebe Treviño, L., Baker, V. L., and Mayer, D. M. (2012). Why employees do bad things: moral disengagement and unethical organizational behavior. *Pers. Psychol.* 65, 1–48. doi: 10.1111/j.1744-6570.2011.01237.x

Muthén, L. K., and Muthén, B. O. (2007). *Mplus User's Guide (1998–2007)*. Los Angeles, CA: Muthén & Muthén.

Obermann, M.-L. (2011). Moral disengagement among bystanders to school bullying. *J. Sch. Violence* 10, 239–257. doi: 10.1080/15388220.2011.578276

Osofsky, M. J., Bandura, A., and Zimbardo, P. G. (2005). The role of moral disengagement in the execution process. *Law Hum. Behav.* 29, 371–393. doi: 10.1007/s10979-005-4930-1

Paciello, M., Fida, R., Tramontano, C., Lupinetti, C., and Caprara, G. V. (2008). Stability and change of moral disengagement and its impact on aggression and violence in late adolescence. *Child Dev.* 79, 1288–1309. doi: 10.1111/j.1467-8624.2008.01189.x

Paciello, M., Tramontano, C., Nocentini, A., Fida, R., and Menesini, E. (2020). The role of traditional and online moral disengagement on cyberbullying: do externalising problems make any difference? *Comput. Hum. Behav.* 103, 190–198. doi: 10.1016/j.chb.2019.09.024

Palladino, B. E., Nappa, M. R., Zambuto, V., and Menesini, E. (2020). Ethnic bullying victimization in italy: the role of acculturation orientation for ethnic minority adolescents with differing citizenship statuses. *Front. Psychol.* 11:499. doi: 10.3389/fpsyg.2020.00499

Palladino, B. E., Nocentini, A., and Menesini, E. (2016). Evidence-based intervention against bullying and cyberbullying: evaluation of the NoTrap! program in two independent trials. *Aggress. Behav.* 42, 194–206. doi: 10.1002/ab.21636

Pelton, J., Gound, M., Forehand, R., and Brody, G. (2004). The moral disengagement scale: extension with an American minority sample. *J. Psychopathol. Behav. Assess.* 26, 31–39. doi: 10.1023/B:JOBA.0000007454.34707.a5

R Studio Team (2020). *RStudio: Integrated Development for R*. Boston, MA: RStudio, PBC.

Rosseel, Y. (2012). Lavaan: an R package for structural equation modeling and more. Version 0.5–12 (BETA). *J. Stat. Softw.* 48, 1–36.

Smith, G. T., McCarthy, D. M., and Anderson, K. G. (2000). On the sins of short-form development. *Psychol. Assess.* 12, 102–111. doi: 10.1037/1040-3590.12.1.102

Thornberg, R., and Jungert, T. (2013). Bystander behavior in bullying situations: basic moral sensitivity, moral disengagement and defender self-efficacy. *J. Adolesc.* 36, 475–483. doi: 10.1016/j.adolescence.2013.02.003

Thornberg, R., and Jungert, T. (2014). School bullying and the mechanisms of moral disengagement. *Aggress. Behav.* 40, 99–108. doi: 10.1002/ab.21509

Thornberg, R., Daremark, E., Gottfridsson, J., and Gini, G. (2020). Situationally selective activation of moral disengagement mechanisms in school bullying: a repeated within-subjects experimental study. *Front. Psychol.* 11:1101. doi: 10.3389/fpsyg.2020.01101

Vandenberg, R. J., and Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organ. Res. Methods* 3, 4–70.

Vrieze, S. I. (2012). Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychol. Methods* 17:228. doi: 10.1037/a0027127

Webster, R. J., and Saucier, D. A. (2015). Demons are everywhere: the effects of belief in pure evil, demonization, and retribution on punishing criminal perpetrators. *Pers. Individ. Dif.* 74, 72–77. doi: 10.1016/j.paid.2014.09.041

Widaman, K. F., and Reise, S. P. (1997). "Exploring the measurement invariance of psychological instruments: applications in the substance use domain," in *The Science Of Prevention: Methodological Advances From Alcohol And Substance Abuse Research*, eds K. J. Bryant, M. Windle, and S. G. West (Washington, DC: American Psychological Association), 281–324.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

BRIEF RESEARCH REPORT
published: 21 January 2022
doi: 10.3389/fpsyg.2021.767596

Assessing and Optimizing Socio-Moral Reasoning Skills: Findings From the MorALERT Serious Video Game

*Hamza Zarglayoun[1,2], Juliette Laurendeau-Martin[1], Ange Tato[3], Evelyn Vera-Estay[4], Aurélie Blondin[1,2], Arnaud Lamy-Brunelle[2], Sameh Chaieb[5], Frédérick Morasse[1], Aude Dufresne[5], Roger Nkambou[3] and Miriam H. Beauchamp[1,2]**

[1] Department of Psychology, University of Montreal, Montreal, QC, Canada, [2] Sainte-Justine Hospital Research Center, Montreal, QC, Canada, [3] Department of Computer Science, Université du Québec à Montréal, Montreal, QC, Canada, [4] School of Psychology, Pontificia Universidad Católica de Chile, Santiago, Chile, [5] Department of Communication, University of Montreal, Montreal, QC, Canada

Background: Social cognition and competence are a key part of daily interactions and essential for satisfying relationships and well-being. Pediatric neurological and psychological conditions can affect social cognition and require assessment and remediation of social skills. To adequately approximate the complex and dynamic nature of real-world social interactions, innovative tools are needed. The aim of this study was to document the performance of adolescents on two versions of a serious video game presenting realistic, everyday, socio-moral conflicts, and to explore whether their performance is associated with empathy or sense of presence, factors known to influence social cognition.

Methods: Participants (12–17 years, M = 14.39; SD = 1.35) first completed a pre-test measure of socio-moral reasoning based on three dilemmas from a previously validated computer task. Then, they either played an evaluative version (n = 24) or an adaptive (n = 33) version of a video game presenting nine social situations in which they made socio-moral decisions and provided justifications. In the evaluative version, participants' audio justifications were recorded verbatim and coded manually to obtain a socio-moral reasoning maturity score. In the adaptive version (AV), tailored feedback and social reinforcements were provided based on participant responses. An automatic coding algorithm developed using artificial intelligence was used to determine socio-moral maturity level in real-time and to provide a basis for the feedback and reinforcements in the game. All participants then completed a three-dilemma post-test assessment.

Results: Those who played the adaptive version showed improved SMR across the pre-test, in-game and post-test moral maturity scores, $F(1.97,63.00) = 9.81$, $p_{HF} < 0.001$, $\epsilon^2 = 0.21$, but those who played the Evaluative version did not. Socio-moral reasoning scores from both versions combined did not correlate with empathy or sense of presence during the game, though results neared significance. The study findings support preliminary validation of the game as a promising method for assessing and remediating social skills during adolescence.

Keywords: moral reasoning, serious video games, adolescence, empathy, presence, neuropsychology, assessment, intervention

OPEN ACCESS

Edited by:
Yoshifumi Ikeda,
Joetsu University of Education, Japan

Reviewed by:
Hideyuki Takahashi,
Osaka University, Japan
Talena Day,
Stony Brook University, United States

****Correspondence:***
Miriam H. Beauchamp
miriam.beauchamp@umontreal.ca

Specialty section:
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

Received: *31 August 2021*
Accepted: *29 December 2021*
Published: *21 January 2022*

Citation:
Zarglayoun H, Laurendeau-Martin J, Tato A, Vera-Estay E, Blondin A, Lamy-Brunelle A, Chaieb S, Morasse F, Dufresne A, Nkambou R and Beauchamp MH (2022) Assessing and Optimizing Socio-Moral Reasoning Skills: Findings From the MorALERT Serious Video Game. Front. Psychol. 12:767596. doi: 10.3389/fpsyg.2021.767596

INTRODUCTION

Daily socio-emotional interactions play an important role in shaping the social brain, especially during childhood and adolescence (Blakemore, 2012; Immordino-Yang et al., 2019). In parallel, the emergence and maturation of socio-cognitive skills supports the ability to create bonds, be aware and understand social situations, and make decisions according to context and societal norms (Beauchamp and Anderson, 2010). Socio-moral reasoning (SMR) is an important socio-cognitive building block defined as the ability to analyze social situations according to moral criteria in order to distinguish right from wrong and regulate behavior in everyday life (Haidt, 2001; Gibbs, 2014). Sound SMR is associated with prosocial behavior, altruistic personality traits and overall better social competence (Eisenberg et al., 2002; Malti et al., 2009; Malti and Latzko, 2010). Conversely, SMR impairments have been linked to maladaptive behaviors, including aggression, rule-breaking and criminality (Arsenio and Lemerise, 2004; Stams et al., 2006; Van Vugt et al., 2011).

Socio-Cognitive Development During Adolescence

Adolescence is an important period for social maturation given the increased autonomy, social network complexity, as well as environmental, biological, neural and cognitive changes that characterize this period (Blakemore, 2012). Environmentally, adolescents tend to reduce their reliance on parents, focus on peer relationships and start modeling them to fit in (Clausen, 1991; Harris, 1995). Peer opinion plays a major role in social decision-making as they become more sensitive to the approval of others and experiment new social roles (Jones et al., 2014). Biologically, pubertal hormones bring about changes that influence how adolescents interact with their peers and surroundings, as well as a strong drive for reward seeking (Sato et al., 2008; Blakemore et al., 2010). Neurally, synaptic pruning and myelination in prefrontal regions continues into late adolescence, leading to more efficient cognitive processing and behavioral changes as well as an eventual reduction in risky behavior tendencies and expanded inhibitory capacities for refraining from inappropriate social behavior (Casey et al., 2005; Blakemore, 2008; Blakemore and Mills, 2014; Qu et al., 2015). Cognitively, executive and socio-cognitive functions such as affect recognition, theory of mind, and empathy develop in parallel allowing flexible processing of complex social stimuli (Tousignant et al., 2017; Beaudoin and Beauchamp, 2020). Thus, both experience and biology underpin the socio-cognitive foundations that promote SMR maturation (Vera-Estay et al., 2015, 2016; Beaudoin and Beauchamp, 2020). Viewed from a cognitive-developmental perspective, SMR development is depicted as a progression from egocentric viewpoints to internalization of societal values throughout childhood and adolescence (Gibbs, 2014). Children and adolescents also learn to distinguish moral, social and psychological knowledge related to moral issues such as fairness, justice, welfare and rights (Turiel, 2002; Killen et al., 2011).

Empathy and Socio-Moral Reasoning

Alongside SMR, empathy also undergoes protracted development during adolescence (Tousignant et al., 2017). Empathy consists of two primary components: an affective response to another person (sharing another's emotional state) and a cognitive component enabling perspective taking while maintaining a self-other distinction (Jackson et al., 2005). Since SMR in part depends on emotional state (Miller et al., 1996; Zarinpoush et al., 2000), it is posited that empathic tendencies contribute to SMR and decision-making. However, studies to date report mixed findings regarding such an association. Some report a positive correlation between SMR and empathy functions (Hoffman, 2001; Dooley et al., 2010; Vera-Estay et al., 2016; Morasse et al., 2021), while others suggest that empathy can, in some cases, lead to amoral behaviors due to partiality that can cloud moral judgment (Batson et al., 1995; Decety and Cowell, 2014). For example, it may be more difficult to maintain a moral stance when a family member or someone we identify with is in a dire situation. Thus, associations between empathy and SMR need to be clarified.

Socio-Moral Reasoning Difficulties in Adolescence

A number of risk factors impede optimal SMR maturation. Identifying and remediating putative SMR difficulties is thus essential during adolescence. Neurodevelopmental and acquired brain conditions such as Traumatic Brain Injury (Beauchamp et al., 2019), Autism Spectrum Disorders (Moran et al., 2011; Senland and Higgins-D'Alessandro, 2013) and Schizophrenia (Abdi and Sharma, 2004) have been associated with altered SMR. Environmental factors and psychological conditions can also constitute risk factors for altered social development and poor SMR, such as socioeconomic status (Bradley and Corwyn, 2002) or behavioral problems (Nelson et al., 1990; Stams et al., 2006) and may require remediation or rehabilitation. Conventional methods of intervention often take the form of cognitive development programs (e.g., moral dilemma discussion sessions) for high-risk adolescents with behavior disorders (Arbuthnot and Gordon, 1986) and moral reasoning promotion programs in high school education settings [e.g., reciprocal teaching style (Johnson and Ward, 2001; Mouratidou et al., 2007)]. However, these approaches can be limited by the use of hypothetical moral dilemmas and methodological constraints affecting engagement, motivation and involvement.

Socio-Moral Reasoning Assessment and Intervention

There is a rich history of social cognition and competence assessment, however, recent recommendations aimed at enhancing the validity and ecology of such assessments highlight obstacles associated with traditional tasks such as the use of written, static and hypothetical scenarios (Beauchamp, 2017). Such approaches often introduce perceptual and cognitive confounds, do not adequately mimic the complexity and dynamism of real-life social scenarios, and limit user engagement,

motivation, familiarity, presence, and immersion (Beauchamp, 2017; Morasse et al., 2021). A range of SMR assessment tools exist, including paper-and-pencil questionnaires, interviews and static cartoon presentation (Dooley et al., 2010). More recently, efforts have been deployed to increase the visual and dynamic nature of SMR assessment using pictures of real people in developmentally appropriate and realistic scenarios (Chiasson et al., 2017) and using virtual reality (Morasse et al., 2021).

Serious Video Games

Serious video games, defined as "video games that use computer-based entertainment technology to teach, train, or change behavior" (Baranowski et al., 2008) are another potential medium for the assessment and remediation of socio-cognitive skills. They have already proven useful in the socio-cognitive domain, such as to improve emotion recognition (Silver and Oakes, 2001; Lacava et al., 2007) and social competence (Beaumont and Sofronoff, 2008) in youth with Autism Spectrum Disorders, or to improve executive functions in adolescents with Attention Deficit Hyperactivity Disorder (Dovis et al., 2015). The appeal of this method is a result of the popularity of recreational video games, their interactive nature, and their dynamic and adaptive qualities, which can promote youth engagement, stimulate affective reactions and boost motivation (Baranowski et al., 2008). This technology tends to elicit a sense of presence from players, defined as "the extent to which one feels present in the mediated environment, rather than in the immediate physical environment" (Steuer, 1992). It has been shown to correlate positively with empathy (Bachen et al., 2016; Morasse et al., 2021), and may contribute to participants feeling more immersed and therefore more engaged. Serious video games could therefore provide an innovative and engaging modality for assessing and optimizing SMR in adolescents.

Objectives

The overarching objective of this study was to provide preliminary information on the development of a novel serious video game (MorALERT) for assessing and optimizing SMR in adolescence. Preliminary data are presented documenting SMR progression during an evaluative (EV) and an adaptive version (AV) of the game. The main difference between the versions is that the latter includes real-time assessment of SMR, automated scoring, and feedback and reinforcements that were directly tailored to responses provided. We hypothesized that players who completed the adaptive version, incorporating feedback and reinforcements in real-time, would improve their SMR throughout the game, but that players who completed the evaluative version, more comparable to a previously validated computerized task (SoMoral), would not. A secondary aim was to document the association between SMR, empathy, and sense of presence in adolescents who completed the game. It was expected that participants with higher empathic tendencies would have greater socio-moral maturity (i.e., higher SMR scores), and that those who felt most immersed in the game would show greater SMR and empathy.

METHODS

Participants

Fifty-seven participants (27 females) between the ages of 12 to 17 years (M = 14.4, SD = 1.4 years) were recruited *via* community web sites and youth organizations (e.g., sports groups, clubs). For inclusion, participants had to be fluent in French and be enrolled in a regular school curriculum without having repeated a grade. Participants were excluded if they had a diagnosis of any neurodevelopmental, genetic, psychiatric or metabolic disorder or history of acquired brain injury.

Procedure

Written consent was obtained from participants or their legal guardian. Participants completed either the EV (n = 24) or AV (n = 33) SMR serious video game MorALERT. Recruitment was conducted in two phases in parallel with developments in the game design itself. The first iteration of the game that was developed was the EV. A second development phase was subsequently initiated to develop the AV, thus recruitment occurred in sequence. For both groups, the assessment session included first a pre-test evaluation of SMR using three socio-moral dilemmas from a validated task (SoMoral, described below) to document participants' initial SMR level. They then played the video game (MorALERT) and finally performed a post-test SMR assessment again using three dilemmas from the SoMoral. To complete the SoMoral and MorALERT, participants were seated at a desk in front of a desktop or laptop computer and were provided headphones to hear the audio stimuli. Standard instructions for both tests were provided and participants completed them on their own with no further input from the examiner. Responses were thus documented using audio recordings. Questionnaires documenting socio-demographic characteristics, empathy, and presence, as well as a brief intellectual functioning assessment were performed after the SMR assessment and game.

Measures

Demographic Questionnaire

An in-house questionnaire was used to document age, sex, ethnicity, academic level, and parental education.

Intellectual Functioning

The two-subtest version (Vocabulary and Matrix Reasoning) of the Wechsler Abbreviated Scale Intelligence (WASI; Wechsler, 1999) was used to estimate general intellectual functioning (IQ, M = 100, SD = 15) for descriptive purposes.

Socio-Moral Aptitude Level Task

Six dilemmas from this previously validated task were equally divided into a pre- and a post-test to assess SMR progression before and after playing the video game. Detailed information on the SoMoral task, cognitive and affective factors associated with performance on the task, and performance in typically developing and clinical samples are presented elsewhere (Dooley et al., 2010; Beauchamp et al., 2013; Vera-Estay et al., 2016; Chiasson et al., 2017). Briefly, the SoMoral is a computerized

task composed of everyday, visual socio-moral dilemmas each depicted by three static pictures. An initial screen represents the name of the dilemma, the next three screens correspond to pictures of real actors playing out a social situation (i.e., a problem associated with justice, welfare, harm or human rights) in the first-person perspective. Participants are asked whether or not they would engage in the action portrayed (moral decision-making) and asked to provide a justification for their response. For example, one of the dilemmas presents a scene during which the participant is losing at a game and must reflect on whether or not they would cheat to get ahead in the game. After viewing the three pictures, the participant is shown a screen that asks them what they would do in this situation (decision-making) and why (justification). In this study, participants' justifications to the second question were recorded using a microphone function on the computer. The answers were then transcribed verbatim and the justifications were coded to obtain a SMR score using a cognitive-developmental approach with a score from 0 (no justification provided) to 5 points (highest level of socio-moral maturity). Lower scores are qualified by responses defined by egocentrism and fear of authority (e.g., I would not steal in a store because I could go to jail) and higher scores embody fundamental societal values such as people's rights to property and integration of diverse points of view (e.g., the shopkeeper depends on selling his things and if people take things from him, he won't have any money). For further details on the different stages and scoring system, please see **Table 1** in Chiasson et al. (2017).

TABLE 1 | Participants' sociodemographic characteristics.

Characteristics	Valid N	Mean or N	SD or%	Statistical comparison between groups
Age	57	14.4	1.4	$t(56) = 0.34$, $p = 0.73$
Sex	57			Fisher's Exact Test odds ratio = 1.59, $p = 0.43$
Female		27	47.4	
Male		30	52.6	
Ethnic background	48			$\chi^2(5) = 6.35$, $p = 0.27$
North America		14	24.6	
Europe		5	8.8	
Asia		15	26.3	
North Africa and Middle East		3	5.3	
Latin America		6	10.5	
Subsaharan Africa		5	8.8	
Education level (father)	39			$\chi^2(4) = 2.91$, $p = 0.57$
Elementary		0	0.0	
High school		4	7.0	
Cegep (college)		9	15.8	
Bachelor's		16	28.1	
Master's		6	10.5	
Doctorate		4	7.0	
Education level (mother)	43			$\chi^2(5) = 7.15$, p = 0.21
Elementary		1	1.8	
High school		4	7.0	
Cegep (college)		8	14.0	
Bachelor's		17	29.8	
Master's		11	19.3	
Doctorate		2	3.5	

MorALERT Serious Video Game

Two versions of the MorALERT video game were developed with Unity software and programmed in C# and Python languages. The game is played on a standard desktop computer and is in the third person perspective: the player incarnates a character whose avatar is visible on the screen. The player can choose the gender of the character in order to facilitate self-identification. Navigation in the virtual environment is performed by using the arrows of the keypad and all of the other actions are performed with the left click on the computer mouse.

Evaluative Version

This version of the game is composed of nine everyday socio-moral dilemmas presented in a predetermined, continuous sequence (see **Figure 1A**). The scenes are dynamic as opposed to those in the SoMoral task which present static pictures and are presented in the third-person perspective. In each dilemma, players encounter the same five avatars (non-player characters, NPC) representing people they know (e.g., friends, classmates, and family) playing out realistic social situations with a moral component (see **Figure 1C**). The scene is also narrated using an audio track. For example, in one scene, the player avatar walks down the street behind someone who drops their wallet. After the wallet has fallen, a voice over integrated in the game asks the player to choose what they would do in this situation (decision-making) by clicking on one of two options (e.g., whether to keep the wallet or not). They are then asked to verbally justify their decision and to record their justification using the microphone function. Then, the five NPC appear and the player can approach them and interact with them. As the player nears a character, they provide their perspective on the socio-moral dilemma presented. Each NPC perspective represents one of the five stages of socio-moral maturity from the SoMoral coding system. The player indicates whether they agree or not with the NPC's point of view by choosing a thumbs up or thumbs down. After each dilemma, a random score consisting of "likes" (a thumbs up symbol similar to that used on social media) and a tally of the player's number of friends is shown to the player in the bottom left of the screen to encourage them to continue the game; these are not, however, graded according to their individual responses. After playing the game, the player's justifications are transcribed verbatim and coded using the So-Moral coding system to obtain a SMR score corresponding to their level of socio-moral maturity.

Adaptive Version

This version of the game is similar to the EV in that they both have the same visual presentation, nine socio-moral dilemmas, and interactions with NPC avatars. The main difference is that the AV relies on an automated coding algorithm based on natural language processing, deep learning and expert knowledge

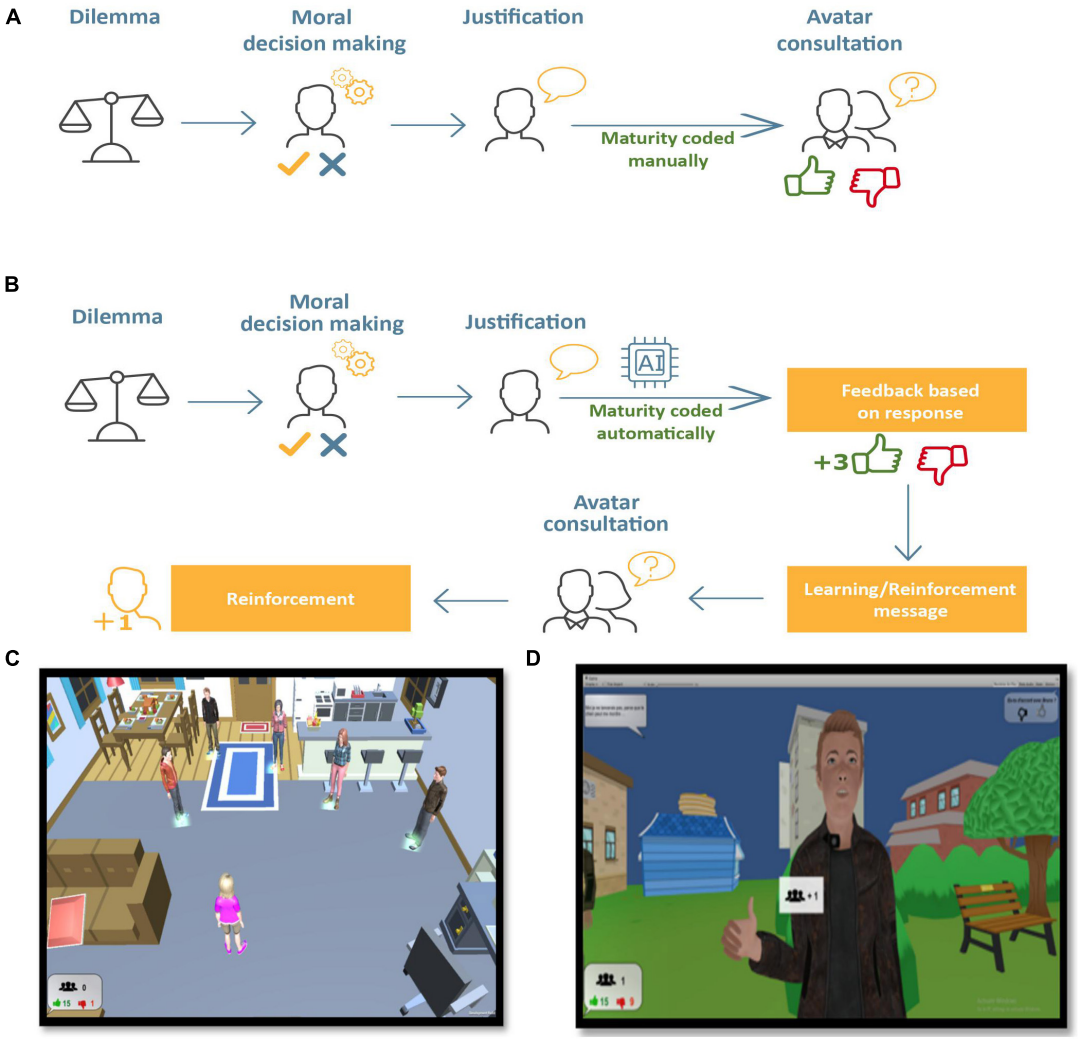


FIGURE 1 | Schematic representation of the sequence for one dilemma in the MorALERT game. **(A)** Evaluative version: the structure and coded are comparable to the original SoMoral task. Coding of justifications is performed manually. Players can consult with avatar friends to hear how they reason and indicate whether they agree (thumbs up) or disagree (thumbs down) with their reasoning, however, no feedback or reinforcement is provided. **(B)** Adaptive version: The overall structure of the dilemmas is comparable to the original SoMoral task; however, justifications are coded automatically in order to provide tailored feedback in the form of "likes" and audio learning/reinforcement messages based on moral maturity stage. After consulting with avatar friends to hear their reasoning about the dilemma, players and indicate whether they agree (thumbs up) or disagree (thumbs down) with their reasoning and obtain "friends" when they agree with a moral reasoning stage that is comparable or higher than their own. Screen captures from the MorALERT game. **(C)** The player, seen in the third-person perspective, faces five avatar friends. In both the evaluative and AVs, using the keyboard arrows, the player can consult each of the friends who will provide their own reasoning to the dilemma in the form of audio sound files. **(D)** In the AV, when the player agrees with friend justifications that are equal or higher maturity than their own response, they gain a "friend." The left bottom screen shows the number of friends and number of likes and dislikes accumulated in the game.

from which an immediate SMR maturity score is produced as players express their moral justifications *via* microphone to a given dilemma (see **Figure 1B**). The algorithm was developed based on manually coded justifications provided in previous work using the SoMoral task; technical information on the design and reliability of the algorithm are available elsewhere (Tato et al., 2017, 2019).

In the AV of the game, this real-time scoring is used to trigger feedback and reinforcements to players throughout the game that are adapted to their decisions and justifications. Feedback and reinforcements are provided in three ways:

(i) Social feedback: the player receives "likes" from the NPC corresponding to the level of SMR maturity provided in their justifications. These are attributed according to the principal of rewarding more mature reasoning, but not penalizing lower stage responses (stage 1 = 5 likes, stage 2 = 8 likes, stage 3 = 11 likes, stage 4 = 14 likes, and stage 5 = 17 likes).

(ii) Learning messages: When players reach a new stage of socio-moral maturity, a praise message is shown on the screen in a dialog window as well as a message reflecting the essential elements of the reasoning stage. For example,

the first time they reach Stage 3, they receive this message: "*Well done! You thought to consider others in making your decision*!" A learning message specific to the dilemma is also shown at the beginning of each dilemma in order to provide a clue for reaching the next stage of moral maturity. For instance, players whose response scored at Stage 2 in a dilemma about keeping a lost wallet would receive this message: "*Did you consider that the owner would like to have her wallet back*?"

(iii) Social reinforcements: as in the evaluative version, the adaptive design includes interactions with five avatar friends (NPC) in which players agree or disagree with their justifications using a thumbs up/thumbs down button. This process is also coupled with reinforcement mechanisms. If the player agrees (thumbs up) with a NPC's justification that is equal or superior to their own stage of SMR maturity, they receive a friend icon. If participants disagree (thumbs down) with a superior or equal stage of sociomoral maturity, the NPC makes a thumbs down as a form of deterrent and the player does not receive a friend icon. These reinforcements accumulate in a box in the bottom, left corner at the bottom of the computer screen (see **Figure 1D**). Similarly, if the player reacts positively toward an inferior sociomoral maturity stage, they do not receive a friend icon and there is no reaction from the NPC.

Interpersonal Reactivity Index

The French adaptation (Gilet et al., 2013) of the IRI (Davis, 1980) was used to document empathy using 28 items (e.g., I am moved by the events that I witness) for which participants rate their empathy on a 5-point scale from 1 (Does not describe me well) to 5 (Describes me very well). The scale contains four 7-item subscales: (1) perspective taking scale, the tendency to adopt the point of view of other people (2) fantasy scale, the tendency of the respondent to feel the actions and feelings of fictive characters in books, movies and plays (3) empathic concern scale, measuring the feelings and worries in reaction to the other's misfortune (4) personal distress scale, assessing personal feelings of anxiety and discomfort in interpersonal contexts. These items can be combined to obtain a global score (global IRI), an affective empathy subscore (affective IRI) and a cognitive empathy subcscore (cognitive IRI). The instrument's reliability (Cronbach's alpha between 0.67 and 0.87) and construct validity are adequate (Hawk et al., 2013).

Presence

The French version of the ITC-Sense of Presence Inventory (ITC-SOPI; Lessiter et al., 2001) was used to measure the individual experience of the participants while playing the game. The ITC-SOPI contains 44 items composed of a five-point Likert scale ranging from ″strongly disagree″ to ″strongly agree″. The scoring of the ITC-SOPI results in a total score encompassing four different factors of presence: Spatial presence, Engagement, Ecological Validity/Naturalness and Negative Effects.

Statistical Analyses

Statistical analyses were performed using the RStudio software (version 1.4.1103). Despite the sequential design, the participant groups (EV and AV) were nonetheless compared on socio-demographic characteristic to ensure general comparability of the samples for descriptive purposes. *T*-tests were used to compare age and IQ, Fisher's Exact Test was used to compare sex and Chi-squared tests were used to compare ethnic background and parental education. Given the non-randomized group attribution and preliminary nature of the study design, direct group comparisons on the main outcome (SMR) were not performed using mixed ANOVA. Instead, to document changes in socio-moral maturity, repeated measures ANOVAs were performed between the pre, in-game and post SMR maturity ratings for the EV and AV groups separately. To prevent sphericity issues, a Huynh-Feldt correction was applied to the obtained *p*-value (p_{HF}). Bonferroni *post hoc* tests were used to determine where differences occurred between the pre-, in game and post-test SMR. Finally, correlations were performed to test the associations between in-game SMR, empathy (global IRI, affective IRI, and cognitive IRI) and sense of presence (total ITC). For these comparisons, the results of participants from both game versions were combined given the absence of *a priori* hypothesis pertaining to differences in these associations between game versions. Of note, nine participants (all in the Evaluative group) had missing data for the IQ or IRI measure due to the later inclusion of these measures in the study design.

RESULTS

Sociodemographic Characteristics

Participants' sociodemographic characteristics are presented in **Table 1**. Their average age was 14.40 years and they were from diverse backgrounds (26% Asian, 25% North American, 11% Latin American). The two groups (EV and AV) did not differ significantly in terms of sex (Fisher's Exact Test odds ratio = 1.59, $p = 0.43$), age ($t(56) = 0.34$, $p = 0.73$), IQ ($t(42) = -1.51$, $p = 0.14$), ethnic background ($\chi^2(5) = 6.35$, $p = 0.27$), paternal ($\chi^2(4) = 2.91$, $p = 0.57$), and maternal ($\chi^2(5) = 7.15$, $p = 0.21$) education level.

Socio-Moral Reasoning Progression Between Pre-test, In-Game, and Post-test

Socio-moral reasoning scores for the pre, in-game and post-test are presented in **Table 2** along with the results of the secondary outcome measures (IRI, ITC). Repeated measures ANOVAs indicate that those who played the EV version did not improve their SMR ($F(2,46) = 0.88$, $p = 0.42$), while those who played the AV showed a significant improvement in SMR scores, with a strong effect size ($F(1.97,63.00) = 9.81$, $p_{HF} < 0.001$, $\epsilon^2 = 0.21$) (Cohen, 1988). A *post hoc* Bonferonni test showed that in the AV version, significant differences were found between the pre and in-game SMR ($t(64) = -4.39$, $p = 0.0001$) as well as between the pre and post SMR scores, ($t(64) = -2.68$, $p = 0.028$). No

TABLE 2 | Age and results on measures of interest according to group (evaluative and adaptive versions of MorALERT).

	Evaluative version				Adaptive version			
Scores	**M**	**SD**	**Range**	**n**	**M**	**SD**	**Range**	**n**
Age	14.46	1.35	12.00—17.00	24	14.33	1.36	12.00—17.00	33
IQ	104.67	10.07	89.00—128.00	15	110.0	10.39	89.00—123.00	28
ITC-SOPI	3.91	0.79	2.25—5.50	24	3.93	0.67	2.38—5.00	31
Pre-test SMR (So-Moral)	2.36	0.77	1.00—3.67	24	2.04	0.90	0.83—4.17	33
In-Game SMR (MorALERT)	2.57	0.58	1.61—3.89	24	2.67	0.56	1.83—3.83	33
Post-test SMR (So-Moral)	2.52	0.87	0.67—4.00	24	2.42	0.65	1.33—3.50	33
IRI-Global	3.42	0.39	2.68—3.82	15	3.33	0.49	2.00—4.07	32
IRI-Affective	3.26	0.31	2.71—3.71	15	3.10	0.47	1.79—4.00	32
IRI-Cognitive	3.57	0.60	2.43—4.36	15	3.57	0.63	2.00—4.71	32

IRI, Interpersonal Reactivity Index; ITC-SOPI, International Test Commission-Sense of Presence Inventory; SMR, Socio-moral reasoning.

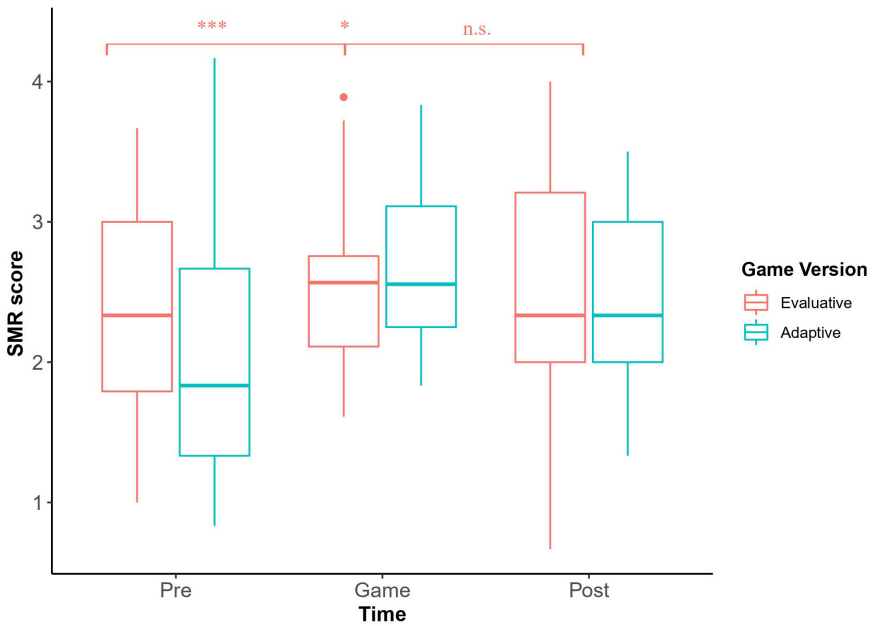


FIGURE 2 | SMR score averages before, during and after the gameplay. As shown above, the average score of the adaptive group significantly increased from pre to in-game and that improvement was maintained to post, whilst the evaluative group's average score stayed relatively stable throughout the experiment.

significant difference was found between the in-game and post-test scores, ($t(64) = 1.71$, $p = 0.21$). The visual representation of those results can be see in **Figure 2**.

Correlations Between Socio-Moral Reasoning, Empathy and Presence

Associations between the in-game SMR scores and empathy neared significance for the global IRI empathy score ($r = 0.26$, $p = 0.07$), and the cognitive IRI empathy subscore ($r = 0.28$, $p = 0.06$) with a medium effect size (Cohen, 1988). No significant correlations were found between SMR and sense of presence (ITC; $r = 0.17$, $p = 0.21$), between global empathy and sense of presence ($r = 0.16$, $p = 0.31$), between cognitive empathy and sense of presence ($r = 0.06$, $p = 0.69$) and between affective empathy and sense of presence ($r = 0.24$, $p = 0.11$).

DISCUSSION

The main objective of the study was to provide preliminary information on the development of a novel serious video game for assessing and optimizing SMR in adolescence and on the effect of evaluative and adaptive versions of the game. The results support the main study hypothesis in that adolescents who played an AV of the game, which incorporated feedback and reinforcement message in the form of learning and motivation cues, significantly improved their SMR by playing the game, while those who played an evaluative version (no feedback or other reinforcements) did not. However, the sequential and non-randomized study design precluded direct comparisons between the two versions of the game and a more controlled experimental design is necessary to draw clear conclusions as to the relative value of the two versions. Nonetheless, the study results provide initial information on this

novel approach to social cognition assessment and optimization. Contrary to expectations, no significant relations were found between SMR, empathy and sense of presence.

Efficacy of MorALERT

The main study findings provide preliminary support for the potential of a serious video game to assess and optimize SMR in adolescents. To our knowledge, this is the first video game that directly focusses on SMR skills. As such, there are no published studies to compare the findings with directly, however, game-like media have previously been used to target other sociocognitive skills. Our results align with studies reporting improvements in affect recognition (emotional regulation, recognition and expression) in children with ADHD using the *EmoGalaxy* serious video game (Hakimirad et al., 2019) and in neurotypical children with the *Socialdrome* serious video game (Tan et al., 2016).

On the surface, the two versions are similar since the same dilemmas are presented; however, the presence of real-time feedback and reinforcements that are tailored to the player's responses distinguishes the learning process experienced by participants who played the AV. Two well-established psychological principles could explain the SMR improvement observed in the AV: operant conditioning (Skinner, 1938) and the effect of personal relevance (Sorrentino et al., 1988). Operant conditioning effects are likely to be present in association with the reward "likes" offered to players when they provide answers of progressively higher moral maturity stages. The rewarding or distressing effect of social media-type likes and friend counts have been extensively reported in other contexts (Sherman et al., 2016, 2018; Lee et al., 2020) and, in the game, serve a similar reinforcement purpose. Notably, the game included only positive reinforcements. These were given when the target behavior was elicited, that is, providing reasoning at a higher moral stage than in the previous dilemma. No punishments or negative reinforcements were included.

In terms of personal relevance, a significant procedural and design difference exists between the EV and AV of the game. The latter integrated an automated coding algorithm that allowed us to provide personally tailored advice to players throughout the game, and this may have made them feel more involved and heightened the relevance and salience of the dilemmas. Empirical work indicates that the more people feel personally involved in situations presented to them, the more they exhibit strong emotional reactions rather than remaining in a theoretical mindset (Darley and Lim, 1992). Inclusion of effective real-world social reinforcements throughout the game to elicit authentic emotional reactions, as well as the presentation of social scenarios typically experienced or encountered by youth bodes well for establishing both internal and ecological validity in future work. The findings on the AV also offer some support and promise in terms of the feasibility of using artificial intelligence within a social cognition tool, as a way to rapidly and accurately code SMR maturity online, in real-time, and bodes well for further applications in future versions of the game or other interactive or digital technologies.

Associations Between Socio-Moral Reasoning, Empathy and Presence

Associations between SMR, empathy and presence were not supported, though the correlations between SMR and empathy showed a trend toward a positive relation and may have been limited by the modest sample size. Some studies have found significant links between these two constructs (Hoffman, 2001; Nicovich et al., 2005; Barriga et al., 2009; Dooley et al., 2010; Bachen et al., 2016; Morasse et al., 2021). With regard specifically to the SoMoral, previous work by our group has reported equivocal findings in this regard. As in the current study, Vera-Estay et al. (2016) found only a near significant relation between SMR and cognitive empathy; however, Morasse et al. (2021) did find a significant relation when a virtual reality version of the SoMoral was tested and interpreted as being due to the more immersive nature of the task. Notably, MorALERT, is interactive, but not immersive. However, it is possible that players completing the AV may feel a greater sense of presence because of the heightened interactive component (feedback) compared to the EV. In this study, effects of presence and personal relevance may overlap or be confounded. Presence is defined as "the extent to which one feels present in the mediated environment, rather than in the immediate physical environment" (Steuer, 1992), while personal relevance is related to how much someone recognizes themselves in a situation and how relevant the situation is to their goals and values (Celsi and Olson, 1988). However, it is possible that those for whom the situation is most relevant also feel more present. Future work using a larger sample and simultaneous randomized to the two versions would allow for direct comparisons between the versions on presence, personal relevance, and other variables of interest.

Another reason for the lack of SMR-empathy link in the current study could be methodological differences in the presentation of the social scenarios. The original SoMoral consists of first-person perspective pictures, whereas the serious video game is played from a third-person perspective, which could make it more difficult for some participants to feel present and engaged when playing the game. Perspective-taking manipulation studies show that the affective processes underlying empathy are more exploited from a first-person perspective than from a third-person perspective (Jackson et al., 2006; Lamm et al., 2007). This increased distance between the player and the characters could dampen their ability to empathize with the avatars in the game.

Strengths, Limitations, and Future Directions

To our knowledge, this is the first serious video game designed to assess and optimize SMR in adolescents and is grounded in previous empirical work and validation studies in both neurotypical children and adolescents and those with acquired or neurodevelopmental disorders. The results, however, need to be considered in light of a number of limitations. First, this is an initial step in the development and study of a serious video game and the sample size is modest and may have

limited the detection of some associations between variables. Second, the study was not designed as a randomized control intervention trial and though the samples were comparable in terms of age and IQ, they were recruited sequentially limiting the possibility of direct comparisons between the two. Third, empathy data was only available for a subset of the sample and most of the participants who completed the empathy measure were from the adaptive group. Fourth, we did not collect information on socio-moral decision-making although this variable does exist in the original SoMoral task. However, in previous studies in typically developing children and adolescents, very few individuals made maladaptive decisions and it is likely that this score is subject to social desirability, thus this variable was not useful in characterizing performance. Significant changes in moral decision-making are, however, observable in some clinical populations, such as for example youth with Traumatic Brain Injury (Beauchamp et al., 2019). Thus, it is not clear in the current study how the participants' reasoning relates to their social decisions and behavior.

Future avenues of research should establish the psychometric properties of MorALERT and test its effects in a large sample, as well as verify which affective, cognitive, social, individual (e.g., learning styles, temperament, traits) and behavioral factors contribute to success in the game, or conversely, impede performance. It is possible that other methods of social learning may be comparable in effect to this video game. Using a longitudinal intervention design with children randomized to each of the video game conditions in addition to a more traditional, low-tech learning control condition could inform on the potential added value of the gamified approach and would allow for direct comparisons between experimental and control conditions, as well as verification of knowledge retention effects in the longer term. Finally, further methodological developments could explore differences in user perspective and test more immersive formats such as augmented or virtual reality for heightening engagement.

CONCLUSION

Serious video games offer an interesting avenue for quantifying and remediating social competence in typically developing youth at-risk. Gamifying knowledge acquisition can heighten learning (Vogel et al., 2006; Tüzün et al., 2009; Miller and Robertson, 2010) and constitutes a motivating medium for youth. The findings of this study using a serious SMR video game show promise in terms of its potential for assessing and possibly improving moral maturity, though further work both in terms of game design and empirical validation are necessary.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because Consent was not obtained from the parents to share the data. Requests to access the datasets should be directed to MB, miriam.beauchamp@umontreal.ca.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Montreal Research Ethics Commitee. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

JL-M, AT, EV-E, SC, AD, RN, and MB contributed to the study conception and design. MB, AD, and RN contributed to the development of the study hypotheses. JL-M, AT, AL-B, and AB performed the testing and data collection. HZ performed the data analysis and interpretation under the supervision of MB. HZ drafted the manuscript with input from MB. All authors read, revised and approved the final version of the manuscript for submission.

FUNDING

This study was funded by operating grants from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2020-04966) and Fonds de la Recherche du Québec-Audace (2022-AUDC-300932) and a team grant (AD, RN, and MB) from the Fonds de la recherche du Québec-Société et Culture.

ACKNOWLEDGMENTS

The authors thank the participants for their time and the ABCs lab for comments on the game development. The authors also would like to thank Floris T. van Vugt for help with the statistical analyses.

REFERENCES

Abdi, Z., and Sharma, T. (2004). Social cognition and its neural correlates in schizophrenia and autism. *CNS Spectr.* 9, 335–343. doi: 10.1017/S1092852900009317

Arbuthnot, J., and Gordon, D. A. (1986). Behavioral and cognitive effects of a moral reasoning development intervention for high-risk behavior-disordered adolescents. *J. Consul. Clin. Psychol.* 54:208. doi: 10.1037/0022-006X.54.2.208

Arsenio, W. F., and Lemerise, E. A. (2004). Aggression and moral development: integrating social information processing and moral domain models. *Child Dev.* 75, 987–1002. doi: 10.1111/j.1467-8624.2004.00720.x

Bachen, C. M., Hernández-Ramos, P., Raphael, C., and Waldron, A. (2016). How do presence, flow, and character identification affect players' empathy and interest in learning from a serious computer game? *Comput. Hum. Behav.* 64, 77–87. doi: 10.1016/j.chb.2016.06.043

Baranowski, T., Buday, R., Thompson, D. I., and Baranowski, J. (2008). Playing for real: video games and stories for health-related behavior

change. *Am. J. Prevent. Med.* 34, 74–82. doi: 10.1016/j.amepre.2007.09.027

Barriga, A. Q., Sullivan-Cosetti, M., and Gibbs, J. C. (2009). Moral cognitive correlates of empathy in juvenile delinquents. *Crim. Behav. Mental Health* 19, 253–264. doi: 10.1002/cbm.740

Batson, C. D., Klein, T. R., Highberger, L., and Shaw, L. L. (1995). Immorality from empathy-induced altruism: when compassion and justice conflict. *J. Pers. Soc. Psychol.* 68, 1042–1054. doi: 10.1037/0022-3514.68.6.1042

Beauchamp, M. H. (2017). Neuropsychology's social landscape: common ground with social neuroscience. *Neuropsychology* 31, 981–1002. doi: 10.1037/neu0000395

Beauchamp, M. H., and Anderson, V. (2010). SOCIAL: an integrative framework for the development of social skills. *Psychol. Bull.* 136, 39–64. doi: 10.1037/a0017768

Beauchamp, M. H., Dooley, J. J., and Anderson, V. (2013). A preliminary investigation of moral reasoning and empathy after traumatic brain injury in adolescents. *Brain Injury* 27, 896–902. doi: 10.3109/02699052.2013.775486

Beauchamp, M. H., Vera-Estay, E., Morasse, F., Anderson, V., and Dooley, J. (2019). Moral reasoning and decision-making in adolescents who sustain traumatic brain injury. *Brain Injury* 33, 32–39. doi: 10.1080/02699052.2018.1531307

Beaudoin, C., and Beauchamp, M. H. (2020). Social cognition. *Handb. Clin. Neurol.* 173, 255–264. doi: 10.1111/j.1469-7610.2008.01920.x

Beaumont, R., and Sofronoff, K. (2008). A multi-component social skills intervention for children with asperger syndrome: the junior detective training program. *J. Child Psychol. Psychiatry All. Discipl.* 49, 743–753. doi: 10.1111/j.1469-7610.2008.01920.x

Blakemore, S. J., Burnett, S., and Dahl, R. E. (2010). The role of puberty in the developing adolescent brain. *Hum.Brain Mapp.* 31, 926–933. doi: 10.1002/hbm.21052

Blakemore, S.-J. (2008). The social brain in adolescence. *Nat. Rev. Neurosci.* 9, 267–277. doi: 10.1038/nrn2353

Blakemore, S.-J. (2012). Development of the social brain in adolescence. *J. R. Soc. Med.* 105, 111–116. doi: 10.1258/jrsm.2011.110221

Blakemore, S.-J., and Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing? *Annu. Rev. Psychol.* 65, 187–207. doi: 10.1146/annurev-psych-010213-115202

Bradley, R. H., and Corwyn, R. F. (2002). Socioeconomic status and child development. *Annu. Rev. Psychol.* 53, 371–399. doi: 10.1146/annurev.psych.53.100901.135233

Casey, B., Galvan, A., and Hare, T. A. (2005). Changes in cerebral functional organization during cognitive development. *Curr. Opin. Neurobiol.* 15, 239–244. doi: 10.1016/j.conb.2005.03.012

Celsi, R. L., and Olson, J. C. (1988). The role of involvement in attention and comprehension processes. *J. Consum. Res.* 15, 210–224. doi: 10.1086/209158

Chiasson, V., Vera-Estay, E., Lalonde, G., Dooley, J., and Beauchamp, M. (2017). Assessing social cognition: age-related changes in moral reasoning in childhood and adolescence. *Clin. Neuropsychol.* 31, 515–530. doi: 10.1080/13854046.2016.1268650

Clausen, J. S. (1991). Adolescent competence and the shaping of the life course. *Am. J. Sociol.* 96, 805–842. doi: 10.1086/229609

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Darley, W. K., and Lim, J. S. (1992). The effect of consumers' emotional reactions on behavioral intention: the moderating role of personal relevance and self-monitoring. *Psychol. Market.* 9, 329–346. doi: 10.1002/mar.4220090406

Davis, M. (1980). A multidimensional approach to individual differences in empathy. *J. Pers. Soc. Psychol.* 10, 1–17.

Decety, J., and Cowell, J. M. (2014). Friends or foes: Is empathy necessary for moral behavior? *Perspect. Psychol. Sci.* 9, 525–537. doi: 10.1177/1745691614545130

Dooley, J. J., Beauchamp, M., and Anderson, V. A. (2010). The measurement of sociomoral reasoning in adolescents with traumatic brain injury: a pilot investigation. *Brain Impair.* 11, 152–161. doi: 10.1375/brim.11.2.152

Dovis, S., Van der Oord, S., Wiers, R. W., and Prins, P. J. (2015). Improving executive functioning in children with ADHD: training multiple executive functions within the context of a computer game a randomized double-blind placebo controlled trial. *PLoS One* 10:4. doi: 10.1371/journal.pone.0121651

Eisenberg, N., Guthrie, I. K., Cumberland, A., Murphy, B. C., Shepard, S. A., Zhou, Q., et al. (2002). Prosocial development in early adulthood: a longitudinal study. *J. Pers. Soc. Psychol.* 82, 993–1006. doi: 10.1037/0022-3514.82.6.993

Gibbs, J. C. (2014). *Moral Development and Reality: Beyond the Theories of Kohlberg, Hoffman, and Haidt*. New York, NY: Oxford University Press. doi: 10.1093/acprof:osobl/9780199976171.001.0001

Gilet, A.-L., Mella, N., Studer, J., Grühn, D., and Labouvie-Vief, G. (2013). Assessing dispositional empathy in adults: a french validation of the interpersonal reactivity index (IRI). *Canad. J. Behav. Sci. Rev. Canad. Sci. Comport.* 45, 42–48. doi: 10.1037/a0030425

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychol. Rev.* 108, 814–834. doi: 10.1037/0033-295X.108.4.814

Hakimirad, E., Kashani-Vahid, L., Hosseini, M. S., Irani, A., and Moradi, H. (2019). "Effectiveness of EmoGalaxy video game on social skills of children with ADHD," in *Proceedings of the 2019 International Serious Games Symposium (ISGS), Online*, (Piscataway, NJ: IEEE). doi: 10.1109/ISGS49501.2019.9046992

Harris, J. R. (1995). Where is the child's environment? A group socialization theory of development. *Psychol. Rev.* 102, 458–489. doi: 10.1037/0033-295X.102.3.458

Hawk, S. T., Keijsers, L., Branje, S. J., Graaff, J. V. d., Wied, M. d., and Meeus, W. (2013). Examining the interpersonal reactivity index (IRI) among early and late adolescents and their mothers. *J. Pers. Assess.* 95, 96–106. doi: 10.1080/00223891.2012.696080

Hoffman, M. L. (2001). *Empathy and Moral Development: Implications for Caring and Justice*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511805851

Immordino-Yang, M. H., Darling-Hammond, L., and Krone, C. R. (2019). Nurturing nature: how brain development is inherently social and emotional, and what this means for education. *Educ. Psychol.* 54, 185–204. doi: 10.1080/00461520.2019.1633924

Jackson, P. L., Brunet, E., Meltzoff, A. N., and Decety, J. (2006). Empathy examined through the neural mechanisms involved in imagining how I feel versus how you feel pain. *Neuropsychologia* 44, 752–761. doi: 10.1016/j.neuropsychologia.2005.07.015

Jackson, P. L., Meltzoff, A. N., and Decety, J. (2005). How do we perceive the pain of others? A window into the neural processes involved in empathy. *Neuroimage* 24, 771–779. doi: 10.1016/j.neuroimage.2004.09.006

Johnson, M., and Ward, P. (2001). Effects of classwide peer tutoring on correct performance of striking skills in 3rd grade physical education. *J. Teach. Phys. Educ.* 20, 247–263. doi: 10.1123/jtpe.20.3.247

Jones, R. M., Somerville, L. H., Li, J., Ruberry, E. J., Powers, A., Mehta, N., et al. (2014). Adolescent-specific patterns of behavior and neural activity during social reinforcement learning. *Cogn. Affect. Behav. Neurosci.* 14, 683–697. doi: 10.3758/s13415-014-0257-z

Killen, M., Rutland, A., and Ruck, M. D. (2011). Promoting equity, tolerance, and justice in childhood. *Soc. Policy Rep.* 25, 1–33. doi: 10.1002/j.2379-3988.2011.tb00069.x

Lacava, P. G., Golan, O., Baron-Cohen, S., and Smith Myles, B. (2007). Using assistive technology to teach emotion recognition to students with Asperger syndrome: a pilot study. *Remed. Spec. Educ.* 28, 174–181. doi: 10.1177/07419325070280030601

Lamm, C., Batson, C. D., and Decety, J. (2007). The neural substrate of human empathy: effects of perspective-taking and cognitive appraisal. *J. Cogn. Neurosci.* 19, 42–58. doi: 10.1162/jocn.2007.19.1.42

Lee, H. Y., Jamieson, J. P., Reis, H. T., Beevers, C. G., Josephs, R. A., Mullarkey, M. C., et al. (2020). Getting fewer "Likes" than others on social media elicits emotional distress among victimized adolescents. *Child Dev.* 91, 2141–2159. doi: 10.1111/cdev.13422

Lessiter, J., Freeman, J., Keogh, E., and Davidoff, J. (2001). A cross-media presence questionnaire: the ITC-sense of presence inventory. *Pres. Teleoperat. Virt. Environ.* 10, 282–297. doi: 10.1162/105474601300343612

Malti, T., and Latzko, B. (2010). Children's moral emotions and moral cognition: towards an integrative perspective. *New Direct. Child Adoles. Dev.* 2010, 1–10. doi: 10.1002/cd.272

Malti, T., Gummerum, M., Keller, M., and Buchmann, M. (2009). Children's moral motivation, sympathy, and prosocial behavior. *Child Dev.* 80, 442–460. doi: 10.1111/j.1467-8624.2009.01271.x

Miller, D. J., and Robertson, D. P. (2010). Using a games console in the primary classroom: effects of 'Brain Training'programme on computation and self-esteem. *Br. J. Educ. Technol.* 41, 242–255. doi: 10.1111/j.1467-8535.2008.00918.x

Miller, P. A., Eisenberg, N., Fabes, R. A., and Shell, R. (1996). Relations of moral reasoning and vicarious emotion to young children's prosocial behavior toward peers and adults. *Dev. Psychol.* 32, 210–219. doi: 10.1037/0012-1649.32.2.210

Moran, J. M., Young, L. L., Saxe, R., Lee, S. M., O'Young, D., Mavros, P. L., et al. (2011). Impaired theory of mind for moral judgment in high-functioning autism. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2688–2692. doi: 10.1073/pnas.1011734108

Morasse, F., Vera-Estay, E., and Beauchamp, M. H. (2021). Using virtual reality to optimize assessment of sociomoral skills. *Virt. Real.* 25, 123–132. doi: 10.1007/s10055-020-00443-9

Mouratidou, K., Goutza, S., and Chatzopoulos, D. (2007). Physical education and moral development: an intervention programme to promote moral reasoning through physical education in high school students. *Eur. Phys. Educ. Rev.* 13, 41–56. doi: 10.1177/1356336X07072675

Nelson, J. R., Smith, D. J., and Dodd, J. (1990). The moral reasoning of juvenile delinquents: a meta-analysis. *J. Abnorm. Child Psychol.* 18, 231–239. doi: 10.1007/BF00916562

Nicovich, S. G., Boller, G. W., and Cornwell, T. B. (2005). Experienced presence within computer-mediated communications: initial explorations on the effects of gender with respect to empathy and immersion. *J. Comput. Med. Commun.* 10:2. doi: 10.1111/j.1083-6101.2005.tb00243.x

Qu, Y., Galvan, A., Fuligni, A. J., Lieberman, M. D., and Telzer, E. H. (2015). Longitudinal changes in prefrontal cortex activation underlie declines in adolescent risk taking. *J. Neurosci.* 35, 11308–11314. doi: 10.1523/JNEUROSCI.1553-15.2015

Sato, S. M., Schulz, K. M., Sisk, C. L., and Wood, R. I. (2008). Adolescents and androgens, receptors and rewards. *Horm. Behav.* 53, 647–658. doi: 10.1016/j.yhbeh.2008.01.010

Senland, A. K., and Higgins-D'Alessandro, A. (2013). Moral reasoning and empathy in adolescents with autism spectrum disorder: implications for moral education. *J. Moral Educ.* 42, 209–223. doi: 10.1080/03057240.2012.752721

Sherman, L. E., Hernandez, L. M., Greenfield, P. M., and Dapretto, M. (2018). What the brain 'Likes': neural correlates of providing feedback on social media. *Soc. Cogn. Affect. Neurosci.* 13, 699–707. doi: 10.1093/scan/nsy051

Sherman, L. E., Payton, A. A., Hernandez, L. M., Greenfield, P. M., and Dapretto, M. (2016). The power of the like in adolescence: effects of peer influence on neural and behavioral responses to social media. *Psychol. Sci.* 27, 1027–1035. doi: 10.1177/0956797616645673

Silver, M., and Oakes, P. (2001). Evaluation of a new computer intervention to teach people with autism or Asperger syndrome to recognize and predict emotions in others. *Autism* 5, 299–316. doi: 10.1177/1362361301005003007

Skinner, B. F. (1938). *The Behavior of Organisms: An Experimental Analysis*. Oxford: Appleton-Century.

Sorrentino, R. M., Bobocel, D. R., Gitta, M. Z., Olson, J. M., and Hewitt, E. C. (1988). Uncertainty orientation and persuasion: individual differences in the effects of personal relevance on social judgments. *J. Pers. Soc. Psychol.* 55, 357–371. doi: 10.1037/0022-3514.55.3.357

Stams, G. J., Brugman, D., Deković, M., Van Rosmalen, L., Van Der Laan, P., and Gibbs, J. C. (2006). The moral judgment of juvenile delinquents: a meta-analysis. *J. Abnorm. Child Psychol.* 34, 692–708. doi: 10.1007/s10802-006-9056-5

Steuer, J. (1992). Defining virtual reality: dimensions determining telepresence. *J. Commun.* 42, 73–93. doi: 10.1111/j.1460-2466.1992.tb00812.x

Tan, J. L., Goh, D. H.-L., Ang, R. P., and Huan, V. S. (2016). Learning efficacy and user acceptance of a game-based social skills learning environment. *Int. J. Child Comput. Interact.* 9, 1–19. doi: 10.1016/j.ijcci.2016.09.001

Tato, A. A. N., Nkambou, R., and Dufresne, A. (2019). "Hybrid deep neural networks to predict socio-moral reasoning skills," in *Proceedings of the 12th International Conference on Educational Data Mining*, eds M. Desmarais, C. F. Lynch, A. Merceron, and R. Nkambou (Montreal, QC: International Educational Data Mining Society (IEDMS)), 623–626.

Tato, A., Nkambou, R., Dufresne, A., and Beauchamp, M. H. (2017). Convolutional neural network for automatic detection of sociomoral reasoning level. *Paper presented at the International Conference on Educational Data Mining (EDM). International Educational Data Mining Society, (10th, Wuhan, China, Jun 25-28, 2017), Online*, Wuhan.

Tousignant, B., Eugène, F., and Jackson, P. L. (2017). A developmental perspective on the neural bases of human empathy. *Infant Behav. Dev.* 48, 5–12. doi: 10.1016/j.infbeh.2015.11.006

Turiel, E. (2002). *The Culture of Morality: Social Development, Context, and Conflict*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511613500

Tüzün, H., Yılmaz-Soylu, M., Karakuş, T., Inal, Y., and Kızılkaya, G. (2009). The effects of computer games on primary school students' achievement and motivation in geography learning. *Comput. Educ.* 52, 68–77. doi: 10.1016/j.compedu.2008.06.008

Van Vugt, E., Gibbs, J., Stams, G. J., Bijleveld, C., Hendriks, J., and van der Laan, P. (2011). Moral development and recidivism: a meta-analysis. *Int. J. Off. Ther. Comparat. Criminol.* 55, 1234–1250. doi: 10.1177/0306624X11396441

Vera-Estay, E., Dooley, J., and Beauchamp, M. (2015). Cognitive underpinnings of moral reasoning in adolescence: the contribution of executive functions. *J. Moral Educ.* 44, 17–33. doi: 10.1080/03057240.2014.986077

Vera-Estay, E., Seni, A. G., Champagne, C., and Beauchamp, M. H. (2016). All for one: contributions of age, socioeconomic factors, executive functioning, and social cognition to moral reasoning in childhood. *Front. Psychol.* 7:227. doi: 10.3389/fpsyg.2016.00227

Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., and Wright, M. (2006). Computer gaming and interactive simulations for learning: a meta-analysis. *J. Educ. Comput. Res.* 34, 229–243. doi: 10.2190/FLHV-K4WA-WPVQ-H0YM

Wechsler, D. (1999). *WASI (Wechsler Adult Scale–Reduced)*. San Antonio, TX: The Psychological Corporation.

Zarinpoush, F., Cooper, M., and Moylan, S. (2000). The effects of happiness and sadness on moral reasoning. *J. Moral Educ.* 29, 397–412. doi: 10.1080/713679391

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Psychology

BRIEF RESEARCH REPORT
published: 18 February 2022
doi: 10.3389/fpsyg.2021.768503

Two Dimensions of Moral Cognition as Correlates of Different Forms of Participation in Bullying

Simona C. S. Caravita[1,2], Johannes N. Finne[1] and Hildegunn Fandrem[1]*

[1] Norwegian Centre for Learning Environment and Behavioural Research in Education, University of Stavanger, Stavanger, Norway, [2] Department of Psychology, Catholic University of the Sacred Heart, Milan, Italy

The present study investigated the extent to which moral disengagement and the tendency to consider moral rules as socio-conventional rules are distinct dimensions of morality, and their association with three different forms of participation in bullying (perpetrating bullying, defending the victim and passive bystander behavior). These two types of moral cognitions have been theorized in different models of morality and are usually studied independently, even if research on *moral shifts* (the interpretation of a moral rule transgression as a socio-conventional rule transgression) suggests some possible overlaps. A group of 276 Italian students from primary and middle school (aged 8–15) completed self-reports assessing moral disengagement, socio-conventional perception of moral rules, and participation in bullying as bully, defender of the victim and passive bystander. Results from structural equation modeling analysis confirmed that moral disengagement and socio-conventional comprehension of aggressions are separate and moderately connected morality dimensions. Controlling for age, gender and SES, only moral disengagement was positively associated with perpetrating bullying. These results point to moral disengagement as the critical component of moral cognitions to be addressed in interventions.

Keywords: moral disengagement, moral domains, bullying, defending the victim, passive bystanding

OPEN ACCESS

Edited by:
Ann Dowker,
University of Oxford, United Kingdom

Reviewed by:
Patrice Rusconi,
University of Messina, Italy
Francesco Margoni,
University of Oslo, Norway
Gina L. Peyton,
Nova Southeastern University, United States

****Correspondence:***
Simona C. S. Caravita
simona.c.caravita@uis.no

Specialty section:
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

Received: *31 August 2021*
Accepted: *07 December 2021*
Published: *18 February 2022*

Citation:
Caravita SCS, Finne JN and Fandrem H (2022) Two Dimensions of Moral Cognition as Correlates of Different Forms of Participation in Bullying. Front. Psychol. 12:768503. doi: 10.3389/fpsyg.2021.768503

INTRODUCTION

Research has devoted considerable attention to processes explaining the associations between moral cognitions and bullying behavior, and two theoretical perspectives on morality have been mainly used: (i) using mechanisms of self-justification that allow the person to act in an aggressive way without feeling guilty by cognitively restructuring the situation, namely *moral disengagement* (Bandura, 1991); (ii) judging moral rules forbidding to harm others as breakable because of a wrong conception of them as dependent on the authorities' statements, and, by consequence as non-worthy by themselves, so that their transgression can be accepted (*social domain theory*; Turiel, 1983).

There is evidence that also bullying perpetrators can evaluate bullying as wrong (Gasser and Keller, 2009); nevertheless they bully peers. Both the social domain and the moral disengagement theories provide some explanation of this gap between the moral evaluation and the actual perpetration of bullying. Extensive literature (e.g., Killer et al., 2019) has provided evidence that higher moral disengagement is associated with higher bullying perpetration and lower defending, and in some studies with higher passive bystanding (Gini, 2006). In the few studies on the perception of bullying as a socio-conventional rule transgression, understanding bullying as violation of socio-conventional rules has been found to be associated with

increased bullying perpetration and lower defending (Caravita et al., 2012).

The theoretical frameworks supporting these two types of moral knowledge are different, and they may be conceived as distinct moral mechanisms. Nevertheless, scarce literature has examined these two constructs in the same framework (Caravita et al., 2012; Thornberg and Jungert, 2013). The phenomenon of the *moral domain shift* (Leenders and Brugman, 2005) suggests the possibility of their inter-connection. If they may relate to each other, however, they may relate differently to forms of participation in bullying when they are considered in the same framework. In this study we aim to contribute to fill in this gap in the literature, by investigating to what extent these two types of moral knowledge are related to each other and to forms of participation in bullying, in order to further light on the organization of moral mechanisms in relation to social behaviors.

Moral Disengagement

Moral disengagement (Bandura, 1991) refers to social cognitive processes through which the person can commit actions that they evaluate to be wrong, by cognitively restructuring the events and selectively avoiding moral censure.

Moral disengagement activates four clusters of cognitive processes aimed at: (1) redefining one's own behavior according to personal purposes; (2) displacing personal responsibility for one's own conduct to other persons or within the group; (3) minimizing the behavior consequences; and (4) considering the victim responsible for the situation or denying the victim's human characteristics. The motivation of activating moral disengagement comes from the need to solve the cognitive dissonance (Festinger, 1957), as the uncomfortable inner state stemming from inconsistencies between one's own actions, beliefs, attitudes or feelings. Moral disengagement mechanisms are used to reduce this uncomfortable inner state due to evaluating one's own behavior as wrong. Moral disengagement has also been conceptualized as learnt socially. It first acts as an *a posteriori* mechanism, after the perpetration of the transgression, as the child learns from other social agents to (self-) justify what they did in order to avoid the subsequent guilt or shame feelings (Bandura, 1986). Then, with the use, moral disengagement starts to be used in the *while* or *before* perpetrating the transgressive action. Accordingly, early adolescents can learn to use moral disengagement from their peers (Caravita et al., 2014). With reference to school bullying, higher levels of moral disengagement are associated with increased bullying (Killer et al., 2019), and lower defending of the victims (Jiang et al., 2020). In some studies, moral disengagement was associated also with higher passive bystanding (Jiang et al., 2020), but the research on this behavior in bullying is still scarce and inconsistent (Mazzone et al., 2016).

Social Domains

From a different theoretical perspective, Turiel's (1983) social domain theory proposes that the social cognition is organized in separate domains. Persons' social experiences influence the development and organization of their social knowledge in varying domains referred to the rules that allow or forbid social behaviors. The basic domains have been defined as *moral* (about concepts on fairness, rights, harm and welfare), *socio-conventional* (about concepts on social organization, social systems, and social conventions), and *personal* (about concepts on persons, self, identity and internal states) (Smetana, 1995). An important finding is that people consider transgressions in the moral domain as more serious and less acceptable than transgressions in the socio-conventional and personal domain (Nucci, 2001).

The moral domain is constructed and developed through experiences of actions that have negative or positive effects on the welfare of others or oneself. The socio-conventional domain, instead, refers to actions regulated by rules thought to depend on authorities' statement (context authority's dependence) and not on superior moral values, they are considered non-universally valid and their transgressions are evaluated less serious than breaking moral rules. The organization of moral knowledge in latent domain structures, informing and influencing the social information processing, has been hypothesized to be shaped by the repeated social interactions (Arsenio and Lemerise, 2004). After established, these mental structures may act *a priori*, leading the action.

Moral Disengagement and Social Domains in Relation to Bullying

Compared to peers, bullying perpetrators also evaluate moral rule transgressions as wrong (Gasser and Keller, 2009). Nevertheless, in early adolescence they are more prone to consider both moral and socio-conventional rules as dependent of the context, thus to attribute also moral rule transgressions to the domain of socio-conventional rules (Caravita et al., 2009). Furthermore, considering breaking moral rules as acceptable (thus as socio-conventional rule transgression) has been found to be associated with higher bullying (Caravita et al., 2012). To our knowledge, no studies have examined the *understanding* of bullying as breaking of a socio-conventional rule in relation to passive bystanding, even if this behavior was associated with higher recognition of bullying as harming the victim and empathizing with the victim (Thornberg and Jungert, 2013).

In general, very few studies have investigated moral disengagement and social domains in the same framework in relation to youth's behavior in bullying (Caravita et al., 2012; Thornberg and Jungert, 2013). Both these two moral dimensions may be relevant in explaining behaviors in bullying, and they are possibly related to each other, even if they were conceptualized within separate theoretical frameworks. In a community sample of adolescents Leenders and Brugman (2005) found that a *domain shift* from the moral toward non-moral (socio-conventional and personal) domains appeared when the adolescents evaluated hypothetical situations about delinquent behavior. That is, they attributed the delinquent behaviors that they perpetrated, but not other delinquent behaviors, to non-moral rule domains. This domain shift, only emerging in relation to one's own transgressions, may serve a similar function as moral disengagement, by legitimizing one's own transgressive behaviors, reducing the cognitive dissonance and preserving the self-esteem. This theorization about domain shifts suggests a possible relation, if not partial overlap, between

the organization of moral knowledge in domains and the moral disengagement mechanisms. Nevertheless, the scarcity of research on the association between the two constructs does not allow to establish to what degree they are inter-related and their relative weights in explaining social behaviors.

The Current Study

In this study we investigated the intertwin of moral disengagement and the perception of aggressions as socio-conventional rule transgressions in association with perpetrating bullying, defending the victim and passive bystander behavior in bullying, focusing on late-childhood and early adolescence as critical age levels (Caravita et al., 2009, 2012).

Our first purpose was to better investigate whether and to what extent moral disengagement and the understanding of aggressions as socio-conventional rule transgressions are related expressions of the moral knowledge organization. We hypothesized that these two morality components are mainly distinct, as the social domains should work more as a static *a priori* (before the action) organization of the socio-moral knowledge, while moral disengagement should be a more dynamic process, at the beginning working *a posteriori*, after the perpetration of the action. Nevertheless, as both these constructs should emerge from social interactions (Bandura, 1986; Arsenio and Lemerise, 2004), and may show some possible overlap in their functioning, (moral shifts; Leenders & Brugman), we hypothesize that they can be inter-correlated. We examined these hypotheses by running a structural equation model (SEM) in which the two moral dimensions were indicators of two separate latent factors, tested against a model in which they were indicators of only one latent factor of morality.

Our second purpose was to investigate whether these two components of the morality differently predict three main forms of participation in bullying, as bully, defender of the victim, and passive bystander. Based on the literature, we hypothesized that both these moral dimensions are associated with higher bullying perpetration and lower defending, while the literature on passive bystanding is too scarce to formulate clear hypotheses. As only few studies have considered the two dimensions of morality in the same framework, we cannot formulate clear hypotheses as well on which of the two dimensions is the most relevant in explaining participations in bullying. Nevertheless, the consistency of the literature on moral disengagement and bullying suggests that this mechanism may be the most important.

METHOD

Participants

Participants were 276 fourth to eight-graders (8–15 years; *Mage* = 11.21; *SD* = 1.52; 50% girls), attending one primary school (44.2%, *Mage* = 9.80, SD = 0.68) and one middle school (55.8%, *Mage* = 12.31, SD = 1.01) in Northern Italy. Participants' majority (85.9%) had an Italian background. To assess the socioeconomic/cultural status (SES) participants reported their parents' jobs and qualifications: 32.9% had families of low-average SES. 32.2% of average SES and 31.8% of average-high SES; 3.1% were not able to provide this information.

Measures

Moral Disengagement Scale

We administered a self-report questionnaire specifically devised to assess moral disengagement in bullying situations (Caravita et al., 2011; Oliveira et al., 2019). The measure was an adaptation of the scale for children by Caprara et al. (1995). It consisted of 17 items assessing seven of the eight mechanisms of moral disengagement (the mechanisms assessed in the version for children of the original adapted measure): moral justification, euphemistic language, advantageous comparison, displacement of responsibility, distorting consequences, dehumanization of the victim, attribution of blame to the victims. Each measure item presented a statement of moral exoneration of bullying conduct (e.g., "Hitting annoying classmates is just like giving them a lesson"; 5 point response scale: 1 strongly disagree to 5 strongly agree). Higher scores indicated higher moral disengagement for bullying. We performed a second-order Confirmatory Factor Analysis (estimator MLR; Mplus 8.4, Muthén and Muthén, 1998–2017) to test the structure of the scale according to Bandura's (1986) theoretical model, with parcel scores of the seven mechanisms loading the four moral disengagement clusters, which in turn loaded a unique latent factor of moral disengagement. The model fitted the data well: $\chi^2(12) = 8.422$ $p = 0.751$, *CFI* = 1.000, *RMSEA* = 0.000 (90% CI.000.044), moral disengagement factor Chronbach's alpha = 0.78.

Socio-Conventional Perception of Moral Rules on Aggressions

A measure adapted by the scale developed by Caravita et al. (2012) was administered. The measure included 16 items, consisting in scenarios where a school socio-conventional rule (4 stories) or a moral rule (12 stories) was broken. The 12 scenarios describing the break of a moral rule (preserving other's wellbeing) regarded situations of aggressive behaviors of three types: physical (4 scenarios), verbal (4 scenarios) and relational aggressions (4 scenarios). For the present study only the 12 scenarios on aggressions were considered. In all the items the main character was a student breaking a school rule. In half of the scenarios characters were girls and in half boys. For each scenario the respondent evaluated on a 4 point likert scale (1 = totally wrong to 4 = totally right) whether the character's behavior is acceptable under three socio-conventionality conditions: if allowed by the principal (main school authority dependency), if allowed by the class teacher (class authority dependency), if behaved out of school (context dependency). Higher scores indicated higher perceived socio-conventionality. We confirmed the structure of the scale by means of a second-order CFA, with parcel scores of the three conditions as manifest indicators of the latent factors of the three types of aggressions, which were indicators of the overall socio-conventionality attributed to the moral rules: model fit $\chi^2(15) = 22.956$ $p = 0.085$, *CFI* = 0.995, *RMSEA* = 0.044 (90% CI.000.078), socio-conventionality factor Chronbach's alpha 0.96.

Forms of Participation in Bullying

We assessed forms of participation in bullying as perpetrator, defender, and passive bystander by administering a peer report measure adapted from Pozzoli et al.'s (2012) scale.

Each of the three behaviors was assessed by three items describing the participation form in situations of verbal, physical and relational bullying. The respondent had to nominate the classmates who more often behaved the way described (unlimited peer nominations). A definition of bullying was provided at the beginning of the questionnaire. Per each item the sum of the received peer nominations was standardized among classmates. The participation behavior score was the item average. Chronbach's alphas: bully 0.78, defender 0.84, passive bystander 0.68.

Procedure

School principals and teachers' committees authorized the participation of some classrooms in grades four and five (last two grades of primary school), and six to eight (middle school). Participants' parents/legal guardians authorized students' participation in the study by providing active consent in response to a letter describing the study and its aims. The 72.3% of families authorized their children participation. Measures were group-administered in classroom situations, during the regular school hours, by a researcher assistant who discussed with the participants the definition of bullying and answered their questions. Participants were informed that they could withdraw from the study at any time without providing any explanation.

Strategy of Analysis

We used Structural Equation Modeling (Robust Maximum Likelihood MLR estimator, MPlus 8.0, Muthén and Muthén, 1998–2017) to test our hypotheses. As first step, we run a set of Exploratory Factor Analyses (EFAs; one to seven factors, factor axes extraction method, Geomin obliquation rotation) to examine the dimensionality emerging from the four clusters of moral disengagement and the nine parcel scores of the aggression conventionality. Then, we run Confirmatory Factor Analyses (CFAs) to test two competitive models: the one-factor model of morality, in which the latent factors of the four clusters of moral disengagement and the three latent factors of the socio-conventional perception of physical, verbal and relational aggressions were specified as indicators of one second-order latent factor of morality; the two factor model in which the two latent factors of moral disengagement and the socio-conventional perception of aggression were separate morality latent factors (**Figure 1**).

As third step, we tested a SEM model in which the three forms of behaviors in bullying were regressed on the final latent structure of morality emerging from step two. Gender (*one* boy, *two* girl), SES and age were included among predictors to control for their effects.

The model goodness of fit was examined against these indices: Chi-square (χ^2), which needs to be non-significant for good fitting models, but which is also sensitive to the sample size and tends to become significant for large samples; the *CFI* index, with a value $\geq$0.90 for acceptable fit (Bollen, 1989) and $\geq$0.95 for good fit (Hu and Bentler, 1995); the *RMSEA* index with a value $\leq$0.08 for acceptable fit and $\leq$0.05 for good fit. Comparison between competitive models was also based on the Chi-square difference test for nested models, retaining the model with the best fit and if the decrease of the Chi-square was significant. The appropriateness of the sample size was established by running a power analysis, which showed that for testing a model with 180 degrees of freedom (our most complex model) and expecting a value of RMSEA equal to 0.08, a sample of 200 participants was enough to provide a power of 1.00 (MacCallum et al., 1996). Furthermore, the criteria recommended for performing CFAs (with p variables, N/p should be $\geq$10; Marsh et al., 1998), indicated in 230 the minimum participants' number to test a CFA model including 23 variables (our two-factor second-order CFA).

RESULTS

Moral Knowledge Organization

The EFAs indicated a six factor model as the best model of parcel scores (**Supplementary Material**): $\chi^2(39) = 33.937$ $p = 0.670$, *CFI* = 1.000, TLI = 1.000, *RMSEA* = 0.000 (90% CI.000.033), 5 factor model χ^2 Difference test (11) = 88.975 $p = 0.000$. One factor (the fourth extracted factor), was loaded only by the seven parcel scores of moral disengagement. The remaining five factors expressed the dimensionality of the measure assessing the socio-conventionality of aggressions. The first extracted factor expressed the authority dependence and was positively loaded by the six indicators of teacher and principal dependency. The three scores of context dependency characterized the second extracted factor. The third extracted factor was loaded positively by teacher dependency of verbal aggression and negatively by the principal dependency of the physical aggression, thus mainly expressing the different perception of seriousness of these two forms of aggression. The fifth extracted factor was mainly loaded by the teacher dependency (0.661) and the principal dependency (0.295) of relational aggression, thus expressing the specificity of this form of aggression. The sixth extracted factor was mainly characterized by principal dependency of verbal (0.247) and relational violence (0.221), expressing some residual variance explained by this type of authority. The moral disengagement factor correlated moderately and significantly with first (0.301, $p < 0.05$), second (0.463, $p > 0.05$), and fifth extracted factors (0.220, $p < 0.05$) and non-significantly with third factor (0.014). On summary, EFAs provided a first indication that moral disengagement and conventionality of aggression are distinct morality dimensions.

In the CFAs, the one-factor model of morality had a non-adequate fit: $\chi^2(90) = 263.730$ $p = 0.000$, *CFI* = 0.934, *RMSEA* = 0.084 (90% CI.072.095). The two-factor model fitted the data adequately, with significantly lower Chi-square than the one-factor model: $\chi^2(89) = 136.920$ $p = 0.000$, *CFI* = 0.982, *RMSEA* = 0.044 (90% CI.029.058), χ^2 Difference test with Satorra–Benter correction (1) = 94.307 $p = 0.000$. Then the two-factor model was retained as the final morality model (**Figure 1**). The two dimensions of morality were moderately associated (0.381).

Moral Dimensions and Participant Roles Behavior in Bullying

The model (**Figure 2**) in which the two morality dimensions predicted the variance of behaviors in bullying had an adequate fit: $\chi^2(180) = 282.201$ $p = 0.000$, *CFI* = 0.965,

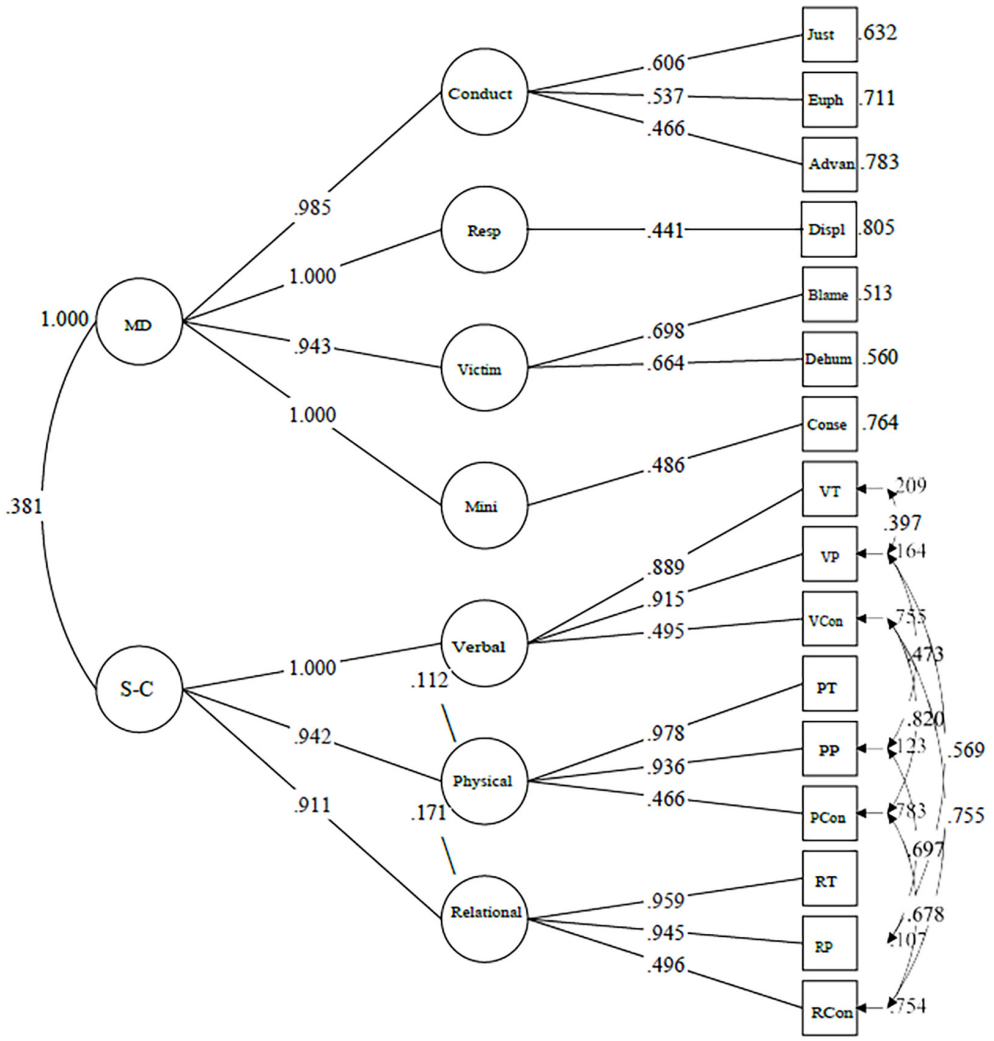


FIGURE 1 | Moral dimensions, two latent factors model. Only significant paths ($p < 0.05$) are displayed in the figure. Standardized indices. MD, Moral Disengagement, S-C, Socio-conventionality of the moral rules for aggression; Conduct, conduct restrictuding mechanisms, Resp, responsibility restructuring mechanisms, Victim, victim restructuring mechanism, Mini, consequence minimization moral disengagement mechanism, Just, moral justification mechanism, Euph, euphemistic labeling mechanism,Advan, advantagious comparison mechanism, Displ, responsibility displacement mechanism, Blame, victim blame mechanism, Dehum, dehumanization mechanism; VT, verbal violence teacher's permission, VP, verbal violence principal's permission, VCon, verbal violence context dependence; PT, physical violence teacher's permission, PP, physical violence principal's permission, PCon, physical violence context dependence; RT, relational violence teacher's permission, RP, relational violence principal's permission, RCon, relational violence context dependence.

RMSEA = 0.048 (90% CI.037.058). Bullying was higher for boys and defending for girls. Older youths showed higher defending. Moral disengagement was positively associated with bullying perpetration and marginally ($p = 0.067$) negatively with defending. The socio-conventional perception of aggressions was not associated with any of the three behaviors.

DISCUSSION

Goal of this study was to investigate whether and to what extent moral disengagement and the socio-conventional perception of aggressions (moral rule transgressions) are possibly related dimensions of morality, and their associations with participation in bullying as bullying perpetrator, defender of the victim, and passive bystander. Our findings confirmed that moral disengagement mechanisms and the socio-conventional perception of aggressions are distinct dimensions of morality, only moderately associated. Only moral disengagement was associated with higher perpetration of bullying and (marginally) with lower defending.

The distinction between the two moral dimensions may relate to the fact that social domains may be a more static organization of moral knowledge, mainly working *a priori*, before the perpetration of the action, while moral disengagement is a mechanism working (originally) *a posteriori*, after the transgression perpetration, to avoid negative emotions. This

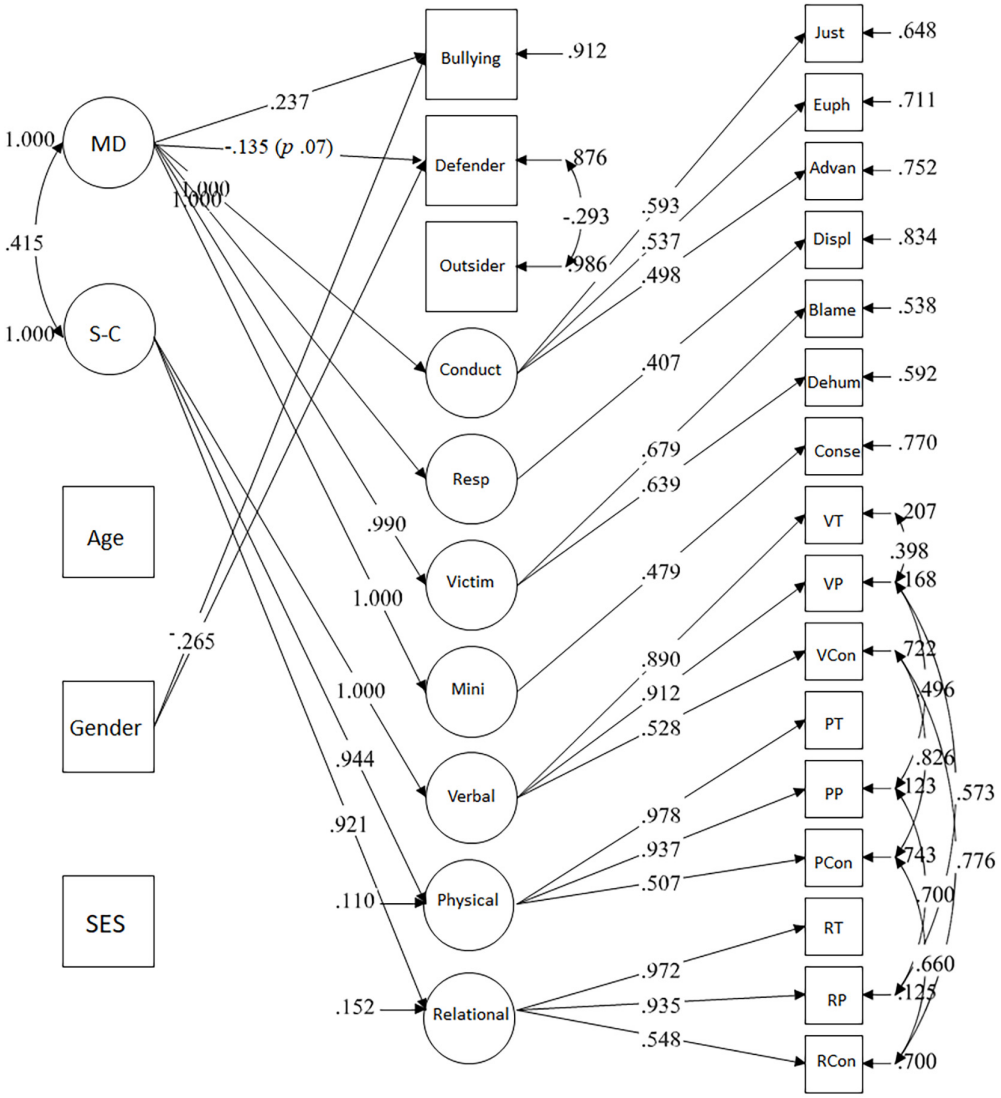


FIGURE 2 | Moral dimensions and forms of participation in bullying. Only significant paths ($p < 0.05$) are displayed in the figure. Standardized indices. Md, Moral Disengagement, S-C, Socio-conventionality of the moral rules for aggression; Conduct, conduct restrictuding mechanisms, Resp, responsibility restructuring mechanisms, victim, victim restructuring mechanism, Mini, consequence minimization moral disengagement mechanism, mJust, moral justification mechanism, Euph, euphemistic labeling mechanism, Advan, advantagious comparison mechanism, Displ, responsibility displacement mechanism, Blame, victim blame mechanism, Dehum, dehumanization mechanism; VT, verbal violence teacher's permission, VP, verbal violence principal's permission, VCon, verbal violence context dependence; PT, physical violence teacher's permission, PV, physical violence principal's permission, PCon, physical violence context dependence; RT, relational violence teacher's permission, RP, relational violence principal's permission, RCon, relational violence context dependence.

difference between the two dimensions can be particularly evident in the case of bullying. Students who bully others are aware of the importance of other's well-being and they know that bullying is wrong (Gasser and Keller, 2009); yet they perpetrate bullying and use moral reasons less frequently and are less prone to referring to the victim's harm in their judgments on bullying (Thornberg et al., 2017). Our findings suggest that the moral distortions of bullying perpetrators may lay more in the *a posteriori* self-justification of behaviors that they know to be wrong than in the *a priori* distortion in the organization of the knowledge about social behaviors in domains. The main difficulty about bullying perpetration, therefore, may lay in motives other than a static organization of moral knowledge, in which bullying is interpreted as a behavior that can be allowed under some circumstances (as a socio-conventional transgression), in that the distinction may be blurred in people's mind, or can vary significantly depending on the person (Kelly et al., 2007; Margoni and Surian, 2021). Bullying has been consistently found to be related to a high popularity peer status, bullies' motivation to achieve and keep dominance among peers (Caravita and Cillessen, 2012; Thornberg and Delby, 2019), peer norms (Salmivalli and Voeten, 2004) and perceived peer pressure as well (Juvonen and Galván, 2008). These factors may be the relevant motives that lead youth to behave bullying, but, as they

evaluate bullying wrong, they need to use moral disengagement to avoid the subsequent negative feelings.

Moral disengagement, however, was significantly related only with bullying perpetration. Consistent with previous studies (Gini et al., 2014: Thornberg et al., 2021), a negative, but marginal, association appeared between moral disengagement and defending. This outcome may indicate that moral factors other than moral reasoning may be more relevant to explain defending. Eisenberg et al. (2016) suggested that there are several motives for performing prosocial behaviors, including egoistic concerns (e.g., the expectation of reciprocity), practical concerns (e.g., preventing an unwanted situation or helping), and other-oriented concerns (e.g., sympathy). Defenders also show high levels of moral sensitivity to the distress of victims, which may lead to higher sympathetic emotions (Menesini et al., 2003). Therefore, for defending the victim, the most important dimensions of the moral functioning may be more related to emotions than cognitive mechanisms.

Passive bystanding was associated with neither moral disengagement nor social domain knowledge. Also for this behavior other factors may be more relevant. Thornberg and Jungert (2013) found that students who are low in defender self-efficacy are more inclined to act as passive bystanders, even if they display low moral disengagement. Furthermore, students may not intervene in bullying episodes because they have a low sense of safety at school (Gini et al., 2008). Hence, the non-intervention in bullying situations may stem not from moral evaluations of behaviors, but from other elements. Moreover, Obermann (2011) suggested a distinction between unconcerned passive bystanders (high in moral disengagement) and guilty passive bystanders (low in moral disengagement), which may equalize the effect of each other in a sample if this distinction is not addressed in the analysis. Finally, we used measures focused on the moral evaluation of aggressions and bullying. By developing measures able to assess how much *not intervening* in bullying situations is perceived as a moral transgression (due to withdrawing from prosocial behaviors) may detect some associations between the two moral dimensions we investigated and passive bystanding.

Limitations and Future Directions of Research

The main limitation of this study consists in the cross-sectional nature of the data that does not allow to draw strong conclusions on the direction of the associations. Notwithstanding, to our knowledge this is the first study investigating the possible relation between the organization of moral knowledge in domains and moral disengagement mechanisms, also in connection with three participant behaviors in bullying. In this perspective, our results contribute to the literature on morality clarifying that the two moral dimensions are actually distinct constructs, only moderately associated, and with different relations with youth's behaviors in bullying. We interpreted this distinction mainly with the *a priori* and *a posteriori* functioning of these constructs, but we need more studies investigating the possible intersections of different dimensions of morality, also in relation to moral emotions.

Lastly, our results can contribute to the literature on anti-bullying interventions, as we confirmed that moral disengagement is a moral cognitive dimension affecting the perpetration of bullying more than others. Therefore, we need more anti-bullying interventions addressing this aspect.

DATA AVAILABILITY STATEMENT

Due to limitations from the Privacy Law restrictions and the participants' ethical consents, the raw data cannot be made accessible out of the research team. Aggregated analyses results, however, can be requested from the corresponding author.

AUTHOR CONTRIBUTIONS

SC made substantial contributions to the conception of the work, the acquisition, analysis, and interpretation of data for the work, drafting the work, and revising it critically for important intellectual content, provided approval for publication of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work were appropriately investigated and resolved. JF made substantial contributions to the conception of the work, interpretation of data for the work, drafting the work and revising it critically for important intellectual content, provided approval for publication of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work were appropriately investigated and resolved. HF made substantial contributions to the conception of the work, drafting the work, and revising it critically for important intellectual content, provided approval for publication of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work were appropriately investigated and resolved.

FUNDING

This research was supported by grant PRIN 2017, nr. 20173E3Z7W_001 from the Ministero dell'Istruzione dell'Università e della Ricerca (Ministry of Education, University and Research—Italy).

ACKNOWLEDGMENTS

The authors are grateful to the students, teachers, and school administrators who participated in this study. Special thanks are due to Linda Andrico for her help with the data collection.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.768503/full#supplementary-material

REFERENCES

Arsenio, W. F., and Lemerise, E. A. (2004). Aggression and moral development: integrating social information processing and moral domain models. *Child Dev.* 75, 987–1002. doi: 10.1111/j.1467-8624.2004.00720.x

Bandura, A. (1986). *Social Foundations Of Thought And Action: A Social Cognitive Theory*. Hoboken: Prentice-Hall.

Bandura, A. (1991). "Social cognitive theory of moral thought and action," in *Handbook Of Moral Behavior And Development*, Vol. 1, eds W. M. Kurtines and J. L. Gewirtz (Mahwah: Erlbaum), 45–103.

Bollen, K. A. (1989). *Structural Equations With Latent Variables*. Hoboken: John Wiley and Sons, Inc.

Caprara, G. V., Pastorelli, C., and Bandura, A. (1995). La misura del disimpegno morale in età evolutiva. *Età Evol.* 46, 18–29.

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

Caravita, S. C., Gini, G., and Pozzoli, T. (2012). Main and moderated effects of moral cognition and status on bullying and defending. *Aggress. Behav.* 38, 456–468. doi: 10.1002/ab.21447

Caravita, S. C., Gini, G., and Pozzoli, T. (2011). "Moral disengagement for victimization in bullying: individual or contextual processes?" in *Paper presented at the SRCD Biennial Meeting*, Montreal.

Caravita, S. C. S., Miragoli, S., and Di Blasio, P. (2009). "Why should I behave in this way? Rule discrimination within the school context related to children's bullying," in *Social Development*, ed. L. R. Elling (Hauppauge: Nova Science Publishers), 269–290.

Caravita, S. C. S., Sijtsema, J. J., Rambaran, J. A., and Gini, G. (2014). Moral Disengagement development: peer influences on cognitive self-justification mechanisms of social transgression. *J. Youth Adolesc.* 43, 193–207. doi: 10.1007/s10964-013-9953-1

Eisenberg, N., VanSchyndel, S. K., and Spinrad, T. L. (2016). Prosocial motivation: inferences from an opaque body of work. *Child Dev.* 87, 1668–1678.

Festinger, L. (1957). *A Theory Of Cognitive Dissonance*. Evantson: Row, Peterson and Co.

Gasser, L., and Keller, M. (2009). Are the competent the morally good? Perspective taking and moral motivation of children involved in bullying. *Soc. Dev.* 18, 798–816. doi: 10.1111/j.1467-9507.2008.00516.x

Gini, G. (2006). Social cognition and moral cognition in bullying: what's wrong? *Aggress. Behav.* 32, 528–539. doi: 10.1002/ab.20153

Gini, G., Pozzoli, T., Borghi, F., and Franzoni, L. (2008). The role of bystanders in students' perception of bullying and sense of safety. *J. Sch. Psychol.* 46, 617–638. doi: 10.1016/j.jsp.2008.02.001

Gini, G., Pozzoli, T., and Hymel, S. (2014). Moral disengagement among children and youth: a meta-analytic review of links to aggressive behavior. *Aggress. Behav.* 40, 56–68. doi: 10.1002/ab.21502

Hu, L. T., and Bentler, P. M. (1995). "Evaluating model fit," in *Structural Equation Modeling: Concepts, Issues, And Applications*, ed. R. H. Hoyle (Thousand Oaks: Sage Publications), 76–99.

Jiang, S., Liu, R.-D., Ding, Y., Jiang, R., Fu, X., and Hong, W. (2020). Why the victims of bullying are more likely to avoid involvement when witnessing bullying situations: the role of bullying sensitivity and moral disengagement. *J. Interpers. Viol.* 7:886260520948142. doi: 10.1177/0886260520948142

Juvonen, J., and Galván, A. (2008). "Peer influence in involuntary social groups: lessons from research on bullying," in *Understanding Peer Influence in Children and Adolescents*, eds M. J. Prinstein and K. A. Dodge (The Guilford Press), 225–244.

Kelly, D., Stich, S., Haley, K. J., Eng, S. J., and Fessler, D. M. T. (2007). Harm, affect, and the moral/conventional distinction. *Mind Lang.* 22, 117–131. doi: 10.1111/j.1468-0017.2007.00302.x

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

Leenders, I., and Brugman, D. (2005). Moral/non-moral domain shift in young adolescents in relation to delinquent behaviour. *Dev. Psychol.* 23, 65–79. doi: 10.1348/026151004X20676

MacCallum, R. C., Browne, M. W., and Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychol. Methods* 1, 130–149. doi: 10.1037/1082-989X.1.2.130

Margoni, F., and Surian, L. (2021). Question framing effects and the processing of the moral-conventional distinction. *Philos. Psychol.* 34, 76–101. doi: 10.1080/09515089.2020.1845311

Marsh, H. W., Hau, K.-T., Balla, J. R., and Grayson, D. (1998). Is more ever too much? The number of in- dicators per factor in confirmatory factor analysis. *Multivar. Behav. Res.* 33, 181–220. doi: 10.1207/s15327906mbr3302_1

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

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

Muthén, L. K., and Muthén, B. O. (1998–2017). *Mplus User's Guide*, 8th Edn. Los Angeles: Muthén & Muthén.

Nucci, L. P. (2001). *Education In The Moral Domain*. Cambridge: Cambridge University Press.

Obermann, M. L. (2011). Moral disengagement among bystanders to school bullying. *J. Sch. Viol.* 10, 239–257. doi: 10.1080/15388220.2011.578276

Oliveira, W. A., Caravita, S. C. S., Colombo, B., Donghi, E., da Silva, J. L., and Iossi Silva, M. A. (2019). Self-justification processes related to bullying among Brazilian adolescents: a mixed methods study. *Front. Psychol.* 10:1086. doi: 10.3389/fpsyg.2019.01086

Pozzoli, T., Gini, G., and Vieno, A. (2012). The role of individual correlates and class norms in defending and passive bystanding behavior in bullying: a multilevel analysis. *Child Dev.* 83, 1917–1931.

Salmivalli, C., and Voeten, M. (2004). Connections between attitudes, group norms, and behaviour in bullying situations. *Int. J. Behav. Dev.* 28, 246–258. doi: 10.1080/01650250344000488

Smetana, J. G. (1995). "Morality in context: abstractions, ambiguities and applications," in *Annals of Child Development*, ed. R. Vasta (London: Jessica Kingsley), 83–130.

Thornberg, R., and Delby, H. (2019). How do secondary school students explain bullying? *Educ. Res.* 61, 142–160. doi: 10.1080/00131881.2019.1600376

Thornberg, R., and Jungert, T. (2013). Bystander behavior in bullying situations: basic moral sensitivity, moral disengagement and defender self-efficacy. *J. Adolesc.* 36, 475–483. doi: 10.1016/j.adolescence.2013.02.003

Thornberg, R., Pozzoli, T., and Gini, G. (2021). Defending or remaining passive as a bystander of school bullying in Sweden: the role of moral disengagement and antibullying class norms. *J. Interpers. Viol.* doi: 10.1177/08862605211037427 [Epub Online ahead of Print].

Thornberg, R., Pozzoli, T., Gini, G., and Hong, J. S. (2017). Bullying and repeated conventional transgressions in Swedish schools: how do gender and bullying roles affect students' conceptions? *Psychol. Sch.* 54, 1189–1201. doi: 10.1002/pits.22054

Turiel, E. (1983). *The Development Of Social Knowledge: Morality And Convention*. Cambridge: Cambridge University Press.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership