

PREPRINT VERSION

Tolmatcheff, C., Galand, B., & Roskam, I (2021). Validation of the French Version of the Moral Disengagement in Bullying Scale: Testing Bandura's Conceptual Model. Journal of School Psychology.

Abstract

Although Bandura conceptualized moral disengagement as a global dimension encompassing a set of mechanisms, the validated factorial structures of the existing scales have not fully matched the conceptual model so far. The present study attempts to narrow the gap between the moral disengagement conceptual model and its measurement tools by testing five factorial structures based both on theory and previous empirical work. A sample of 1,107 pupils from grades 4 to 6 completed the French version of the Moral Disengagement in Bullying Scale (Fr-MDBS). Confirmatory Factorial Analyses revealed that a second-order factor structure, congruent with the conceptual model, was better than a single-factor and as good as a first-order multi-factorial structure. Results also supported both construct validity (gender comparison) and predictive validity of the scale. Moral disengagement mechanisms were associated with self- and peer-reported bullying and victimization, defender, and outsider behaviors. Euphemistic labeling, diffusion of responsibility, and distorting consequences were also related to bullying while moral justification and victim attribution were related to both bullying and victimization. Euphemistic labeling, diffusion of responsibility, distorting consequences, and victim attribution were negatively associated with defender behavior and positively related to outsider behavior. Advantageous comparison was not related to any of the examined variables. The Fr-MDBS seems to be a useful assessment tool to validly examine either the mechanisms or the global dimension of moral disengagement.

Keywords: moral disengagement, bullying, victimization, defenders, outsiders, peer-reported.

Validation of the French Version of the Moral Disengagement in Bullying Scale: Testing Bandura's Conceptual Model

How can we explain that otherwise kind and thoughtful children and adolescents sometimes behave in despicable and hurtful ways toward their peers and yet do not feel guilty about it? How can we explain that most students who witness these aggressions do not intervene to support the victims? These questions are particularly relevant to the context of school bullying, a situation of systematic, unprovoked, peer-to-peer abuse in a setting of power imbalance (Gini et al., 2011). Bullying is a prevalent and pervasive problem in schools worldwide, with potential negative consequences that may persist into adulthood, such as depression, anxiety, or suicidality (e.g., Wolke & Lereya, 2015). Whereas most school-age children and adolescents report to disapprove bullying and to be aware of its immoral nature (e.g., Gini et al., 2011; Salmivalli, 2001), a substantial portion of them do bully others, and reactions of defense or support toward victims from bystanders are rare (Hymel et al., 2010; Salmivalli, 2001). This discrepancy between students' attitudes and their actual behavior can interestingly be analyzed through the lenses of Bandura's model of moral disengagement (Hymel & Bonnano, 2014).

Mechanisms of Moral Disengagement

In his social cognitive theory of thought and action, Bandura (1986) offers an explanatory framework based on the central role of selective disengagement from one's moral principles. When growing up, children integrate moral standards that guide their conduct. Behaving in accordance with these moral principles leads to self-satisfaction and a sense of self-worth whereas transgressing them leads to self-sanctions like guilt and shame. However, some cognitive processes can selectively disengage internal controls, allowing one to act in an immoral way without undergoing self-condemnation. Bandura identified eight of these moral disengagement mechanisms, split into four domains of moral disengagement. The full

conceptual model is organized as follows (see Figure 1). The cognitive restructuring domain includes the mechanisms of *moral justification* (the immoral behavior serves a moral and worthy purpose), *euphemistic labeling* (the immoral behavior is negated through a sanitized and innocuous language), and *advantageous comparison* (the immoral behavior is made benign by comparing it with an even more harmful one). The second domain of minimizing one's agentic role operates through the mechanisms of *responsibility displacement* (responsibility belongs to someone/something else), and *responsibility diffusion* (responsibility is diffused among several people and no one bears it). In the third domain of distorting the consequences, one can minimize, deny, or even pretend that the consequences of the detrimental behavior are positive (*disregarding or distorting the consequences*). The last domain of victim attribution consists of *dehumanization* (victims are divested of their human being quality), and *blaming the victim* (victims are blamed for the situation).

[Please insert Figure 1 about here]

The use of these mechanisms of moral disengagement is not limited to exceptional contexts in which large-scale atrocities take place. Instead, these mechanisms are encountered in everyday life and belong to normal functioning (Bandura, 1986). Bandura's conceptual model of moral disengagement has been thought to be especially relevant in the context of school bullying, as cross-sectional studies have highlighted associations between moral disengagement on the one hand, and bullying perpetration, victimization, as well as bystanders' roles, on the other.

Moral Disengagement and Bullying

With regard to the perpetration of bullying (i.e., engaging in bullying behavior towards other students), a recent meta-analysis examining 49 studies found a mean correlation of $r = .31$ with moral disengagement (Killer et al., 2019). In addition, the relation between moral disengagement and bullying perpetration holds even when the influence of other factors like

aggression efficacy (Barchia & Bussey, 2011) or parental monitoring and mother-child warmth/support (Pelton et al., 2004) are controlled.

Results on the association between moral disengagement and victimization (i.e., being a recipient of bullying) are inconsistent. While some studies report a negative relation (Pornari & Wood, 2010), others point out a positive relation (Allison & Bussey, 2017), or no relation (Pozzoli et al., 2016). On the one hand, victims might have an enhanced sensitivity to moral responsibility due to their experience of being victimized (Perren et al., 2012). On the other hand, as victims often blame themselves for their plight (Graham & Juvonen, 1998; Tolmatcheff et al., 2019), they might display higher levels of moral disengagement, notably at the victim level (i.e., *dehumanization* and *blaming the victim*) (Perren et al., 2012). A meta-analytic review of 21 studies found a significant, yet weak, positive association between moral disengagement and victimization, with a mean effect size of $r = .08$ (Killer et al., 2019).

It should be noted that whether bullying perpetration and victimization are self- or peer-reported has been shown to influence their link with moral disengagement (Gini et al., 2014; Killer et al., 2019). For instance, the association with moral disengagement was equal to $r = .34$ for self-reported bullying but equal to $r = .17$ for peer-reported bullying in a recent meta-analysis (Killer et al., 2019). As moral disengagement is usually self-reported, this might be the result of shared method variance (Killer et al., 2019). There has been a longstanding debate to determine which of self- and peer-reported data are the most suitable method to assess bullying perpetration and victimization. Recent recommendations support the use of both forms of reports to compensate for their respective biases as well as to capture a more accurate picture of bullying (e.g., Hymel & Swearer, 2015; Volk et al., 2017).

Bystanders' key role in school bullying by either enhancing or dissuading it has long been recognized (Jenkins & Troop-Gordon, 2020; Salmivalli et al., 1996). A negative association was found between moral disengagement and the defender role (i.e., bystanders

providing support to the victim or even standing up to the bully), with a mean correlation of $r = -.17$ computed from 23 studies (Killer et al., 2019). As defenders' behavior is in line with their moral standards, there is no use for them to rely on moral disengagement mechanisms (Doramajian & Bukowski, 2015). For outsiders (i.e., bystanders not directly involved in bullying who passively observe it or walk away), the few existing results are mixed. Some studies reported a negative (e.g., Gini, 2006; Thornberg & Jungert, 2013) or null association (e.g., Gini et al., 2015). Most of recent findings, however, indicated a positive association (e.g., Doramajian & Bukowski, 2015; Sjögren et al., 2020; Thornberg et al., 2017), which is consistent with the theoretical expectations. As they refrain from helping, outsiders are likely to rely on moral disengagement mechanisms to distance themselves from the victim's distress and avoid self-sanctions (Doramajian & Bukowski, 2015). The assistant (i.e., bystander who joins in the bullying but does not initiate it) and reinforcer (i.e., bystander who displays encouraging behaviors, like laughing or cheering) roles have been traditionally considered in the literature as well (Salmivalli et al., 1996; Troop-Gordon et al., 2019). However, this categorization has met criticism because of the strong covariances between the bully, assistant, and reinforcer roles that question the conceptual distinction between them (e.g., Crapanzano, et al., 2011; Troop-Gordon et al., 2019). For this reason, we focused on defenders and outsiders in the present study.

Measurement of Moral Disengagement

Based on this set of results, moral disengagement has been suggested both as a relevant theoretical framework to understand bullying, and a promising avenue for prevention (e.g., Almeida et al., 2009; Hymel & Bonanno, 2014; Obermann, 2011; Wang & Goldberg, 2017). As such, a valid and reliable measurement tool is required to operationalize the concept of moral disengagement as well as to conduct intervention studies targeting moral disengagement. Although three scales inspired by Bandura's framework have been developed

to measure moral disengagement, their validated factorial structures have not fully matched the theoretical model so far. The most commonly used scale was developed by Bandura and colleagues (Bandura et al., 1996; Caprara et al., 1995) in three versions, namely for elementary children (14 items), adolescents (24 items) and adults (32 items). Another 13-item scale was proposed by Hymel et al. (2005) to assess moral disengagement in the specific context of school bullying. Whereas Bandura originally presented moral disengagement as a construct encompassing four domains and eight mechanisms, these two scales failed to identify distinct dimensions and were validated as unidimensional structures, thus calling into question either the theoretical model itself or its operationalization. Subsequent studies using either scale repeatedly displayed a single-factor solution (e.g., Almeida et al., 2009; Bandura et al., 2001; Gini, 2006; Obermann, 2011; Paciello et al., 2008; Pelton et al., 2004). To our knowledge, only one study using the shortened version of Bandura's scale identified the four broader domains -but not the eight mechanisms- of moral disengagement in the context of school bullying (Pozzoli et al., 2012). In that respect, Thornberg and Jungert (2014) made a valuable step forward in developing and validating a third scale, entitled Moral Disengagement Bullying Scale (MDBS). They initially created 55 items designed to capture the eight mechanisms of moral disengagement, and tested it in a sample of 372 Swedish children aged 10 to 14 years old. Through a systematic and iterative process of repeated Explorative Factor Analysis (EFA), they found a seven-factor solution based on 18 items that fit the theoretical subdivision into mechanisms, except for *dehumanization* and *blaming the victim*, which formed a common factor. Although Thornberg and Jungert's scale thus encompasses only seven of the eight theoretical mechanisms, it is to date the validated scale closest to Bandura's conceptual subdivision into mechanisms.

At this point, it is not clear whether moral disengagement domains and mechanisms are all highly inter-related (unidimensional construct) or partially independent

(multidimensional construct). Moreover, since the domains and mechanisms of moral disengagement are conceptually distinct but supposed to operate together (Thornberg & Jungert, 2014), an even better suited factorial structure should include first-order factors corresponding either to the four domains or to the eight mechanisms, and a second-order factor corresponding to moral disengagement. In a recent study using the MDBS (Bjärehed et al., 2020), such a factorial structure encompassing the seven first-order factors and a higher order factor presented a better fit than a single-factor structure. However, whether this higher order factorial structure also prevails over the original seven first-order factor structure still needs to be clarified. In addition, bullying is recognized as a serious problem not only in high schools but also in primary schools (Hong & Espelage, 2012), with a “peak” suggested to occur between ages 9 and 15 (Álvarez-Garcia et al., 2015; Carney & Merrell, 2001). It is therefore of interest to have a measure of moral disengagement also validated in elementary school-age children.

The Current Study

Our main goal was to validate the French version of the Moral Disengagement in Bullying Scale (hereafter Fr-MDBS) (Thornberg & Jungert, 2014) in a sample of 8- to 12-year-old children. First, we tested the Fr-MDBS’ construct validity by examining its factorial structure. We tested five different models relying on both Bandura’s theoretical framework and previous empirical research. Second, since higher absolute levels of moral disengagement have consistently been found in boys compared to girls (Gini et al., 2014), higher levels of moral disengagement in boys would also support construct validity. We thereby tested measurement invariance across gender. When scalar invariance is reached, we can meaningfully compare boys and girls (Dimitrov, 2010). Third, we assessed the predictive validity of the scale by examining the associations between moral disengagement and bullying perpetration, victimization, and bystander roles (i.e., defender and outsider). Relying

on results from previous studies, we hypothesized that moral disengagement would be positively associated with both bullying perpetration and victimization (Killer et al., 2019). As we followed the recommendations to use both self- and peer-reports to evaluate bullying perpetration and victimization, we expected moral disengagement to be more strongly linked to bullying and victimization when the latter were self-reported (Gini et al., 2014; Killer et al., 2019). With regard to bystanders, we expected a negative association between moral disengagement and the defender role but a positive association between moral disengagement and the outsider role (Doramajian & Bukowski, 2015; Sjögren et al., 2020; Thornberg et al., 2017).

Method

Participants

An initial sample of 1,216 fourth, fifth and sixth graders participated in this study ($M = 10.15$ years; $SD = 1.02$; 52.4% girls). They came from 57 classrooms of nine French-speaking primary schools. These schools were diversified in terms of size, socioeconomic level, and geographical location. They were recruited on a voluntary basis in the context of a larger intervention study on school bullying. Passive parental consent was obtained prior to data collection through paper consent forms sent by the schools. Sixty-four students (5.2%) were absent during the data collection stage, and the parents of 37 pupils (3%) did not want their child to participate. These pupils, as well as eight recently arrived immigrant pupils (0.6%) who had difficulties reading French, were dropped from the analyses. The final sample was comprised of 1,107 pupils.

Procedure

The Fr-MDBS was obtained with a double-check procedure in collaboration with a certified English teacher from the Institute for Modern Languages of our university. Both the first author of this paper and the English teacher translated the scale in parallel. Discrepancies

between the two translations were then discussed to reach agreement. Minor modifications were made to simplify and adapt the items to French. These adjustments partly relied on a pilot study conducted with 38 fourth graders from a school outside the sample and are detailed in supplementary material (see Table S.1). The goal of this pilot study was to check that the youngest pupils of the present study (i.e., the fourth graders) would understand the items correctly and easily. In a first step, the 38 pupils were asked to help the researchers by testing a questionnaire that was intended for some other pupils of their age. They were specifically requested to report any unclear or difficult words or items. In a second step, a group of five pupils volunteered to examine the whole questionnaire together with the researchers to discuss the listed issues in depth and find the best alternatives. The final decisions were endorsed by the first author of this paper and the English teacher to ensure the accuracy and meaning of the items. The French version of the items is provided in supplementary material (see Table S.1).

Afterwards, pupils in the sample of the present study were asked to fill in confidential pen and paper questionnaires during regular class hours under the supervision of two researchers – a PhD student (first author of this paper) and a Master's student. Pupils were informed that their participation was not mandatory and that their answers were confidential. In the fourth graders' classrooms, the questionnaire was read aloud by a researcher during the completion to compensate for potential reading difficulties at this age. The fifth and sixth graders read and completed the questionnaire by themselves after receiving general instructions. The procedure was approved by the ethic committee of the Psychological Sciences Research Institute of our university.

Measures

Moral Disengagement

The 18 items of the Fr-MDBS represent seven of the eight mechanisms theorized by Bandura and colleagues (1996), as dehumanization (one item) and blaming the victim (two items) had been merged into one factor (victim attribution) in the original validation study (Thornberg and Jungert, 2014). Pupils were asked to indicate to which degree they agreed or disagreed with each statement (from 0 = *totally disagree* to 4 = *totally agree*). Factorial structure and internal consistency of the Fr-MDBS are presented in the results section. Reliability indices for the global index and seven subscales, as well as CFA results from Thornberg and Jungert's original validation study (2014) can be found in Appendix A.

Bullying Perpetration

Self-Report. Eleven items listed verbal (teasing, insulting), physical (hitting/kicking/slapping, biting/scratching/pulling hair), social (excluding, spreading rumors), material (racketing, hiding/damaging personal things) and cyber (insulting/scaring through the Internet, calling/texting, releasing pictures/videos) forms of bullying behaviors. Pupils had to indicate the number of times (from 0 = *zero times* to 4 = *four times or more*) they had engaged in each behavior toward another pupil over the last three months (ordinal alpha $\alpha = .87$). This scale has previously been used among elementary school students (Grades 4-6; Tolmatcheff et al., 2020) and has a high internal consistency ranging from $\alpha = .84$ to $\alpha = .86$ (e.g., Galand et al., 2014; Tolmatcheff et al., 2020).

Peer-Nomination. Following the self-reported measure, each participant received a list with the names of all the children in the class. The instruction was inspired by Salmivalli et al. (1996), but unlike the original procedure, a single item captured the bullying role by specifying that the previous list of items (from the self-reported measure of bullying) were examples of bullying behaviors. Pupils were asked to “tick an unlimited number of classmates

who are used to bullying¹ others, for instance, by doing some of the things described above”. The number of nominations received by each pupil was divided by the total number of pupils in the classroom minus one (as students were not allowed to nominate themselves), resulting in scores ranging from 0 to 1. The use of a single item is not unusual in peer-nomination research and its reliability arises from the multiple reporters instead of multiple items (Cillessen & Marks, 2017).²

Victimization

Self-Report. Eleven items listed the same common forms of bullying behaviors as in the self-reported bullying scale described above (Galand et al., 2014; Tolmatcheff et al., 2020). This time, pupils had to indicate the number of times (from 0 = *zero times* to 4 = *four times or more*) they had undergone the unpleasant experiences over the past three months, e.g., “another pupil has hidden or damaged my things on purpose.” (ordinal alpha $\alpha = .89$). This scale has previously been used among elementary school students (Grades 4-6; Tolmatcheff et al., 2020), with a high internal consistency ranging from $\alpha = .83$ to $\alpha = .89$ (e.g., Galand et al., 2014; Tolmatcheff et al., 2020).

Peer-Nomination. Pupils were asked to tick an unlimited number of classmates who used to be bullied by others. This single item (inspired by Salmivalli et al., 1996) captured the victim role by specifying that the previous list of items (from the self-reported measure of victimization) were examples of ways of being victimized by others. The number of nominations received by each pupil was divided by the total number of pupils in the classroom minus one, resulting in scores ranging from 0 to 1.²

Defender Behavior

¹ We did not use the French verb commonly used for bullying (‘*harceler*’) but instead paraphrased it. The wording “[...] les élèves qui embêtent souvent les autres” can be literally translated by “[...] the students who frequently harass others”.

² For a recent example of the use of a single item for bully/victim peer-nomination, see Garandeau et al. (2019).

We used the subset of items selected by Ruggieri et al. (2013) from the shortened version of the participant roles scale adapted by Sutton and Smith (1999) and originally developed by Salmivalli (1996). The three items investigated pupils' potential defending behavior when witnessing a bullying episode: "I ask an adult or peer for help", "I comfort the victim", "I try to help the victim". Pupils had to indicate how often they behaved as described in each item (0 = *never*, 1 = *seldom*, 2 = *sometimes*, 3 = *usually*, 4 = *always*). Ordinal alpha was $\alpha = .67$.

Outsider Behavior

Pupils were asked about behaving as an outsider when witnessing a bullying episode. We adapted three items based on the aforementioned shortened version of the participant roles scale (Sutton & Smith, 1999) to be formulated as self-reported (instead of peer-reported) statements. For instance, the description "Pretends not to notice what is happening" was changed into "I pretend not to notice what is happening". The two other items were: "I don't look in that direction or I stay away", and "I don't do anything". Children were asked to indicate how often they behaved as described in each item (0 = *never*, 1 = *seldom*, 2 = *sometimes*, 3 = *usually*, 4 = *always*). Ordinal alpha was $\alpha = .70$.

Data Analyses

We used Confirmatory Factor Analyses (CFAs) to compare the fit of five different models. **Model 1:** First, we tested a single factor solution. **Model 2:** Next, we tested a four-factor model with the four domains of moral disengagement as latent factors, as found by Pozzoli et al. (2012). **Model 3:** Then, we tested a seven-factor model, corresponding to the mechanisms, as found by Thornberg and Jungert (2014). **Model 4:** Congruent with Bandura's theoretical model, we tested a second-order factor model by adding a higher order factor to the four-factor model (i.e., Model 2). **Model 5:** Finally, we tested a second-order factor model by adding a higher order factor to the seven-factor model (i.e., Model 3). All CFAs were

performed with MPlus version 8.4 and took the nested structure of the data into account by using the TYPE=COMPLEX option of the ANALYSIS command in conjunction with the CLUSTER option of the VARIABLE command. In this approach, the Maximum Likelihood Robust (MLR) method of estimation is used by default when all observed dependent variables are continuous. In MLR, missing data are handled using full-information maximum likelihood (FIML). Data were missing on all observed variables in seven cases (0.6%), which were not included in the analyses. The covariance coverage of data was above 98% for all the indicators. The Chi-square (χ^2) test has long been used as a measure of fit, although it is highly sensitive to sample size and can thus lead to model rejection even when discrepancies from expected values are trivial (e.g., Dimitrov, 2010). Because of the unsatisfactory nature of the χ^2 statistic for model fit assessment, diverse alternative indices have been developed (Schermelleh-Engel et al., 2003). We reported the root mean square error of approximation (RMSEA), the standardized root mean square residual (SRMR), the comparative fit index (CFI), and the Tucker-Lewis index (TLI) as descriptive goodness-of-fit indices, along with the χ^2 test, as recommended by Hu and Bentler (1999). In addition, we reported the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), which can also be used to compare non-nested models (Schermelleh-Engel et al., 2003). For RMSEA and SRMR, values are considered as acceptable when they are equal to or lower than .08, and good when they are equal to or lower than .05. For CFI and TLI, values should be equal or higher than .90 to be acceptable, and equal or higher than .95 to be good (Little, 2013).

AIC and BIC values cannot be interpreted in isolation. Within a set of models for the same data, the lowest values indicate the best fitting model (Schermelleh-Engel et al., 2003). However, it is common for AIC and BIC to disagree as they handle the tradeoff between sensitivity and specificity differently. The BIC rewards model parsimony more than AIC, and is usually preferred in large samples because its probability of selecting the true model

increases with N (Dziak et al., 2020). As the χ^2 difference test applied to nested models is subject to the same shortcomings as the χ^2 test applied to any single model (i.e., trivial differences may become significant even in modest samples), it is not of use for comparing nested models in the present study (Schermelleh-Engel et al., 2003). It should be noted that since RMSEA, SRMR, CFI, and TLI are absolute (as opposed to relative) indices of fit (Dimitrov, 2010), their values should not be directly used to compare competing models to each other (i.e., comparing a CFI of, for example, .98 to a CFI of .96, which both indicate a good fit, is meaningless). If both competing nested models are well fitting according to the conventional thresholds listed above, the more restricted model should not be rejected, provided that there is theoretical justification supporting this model (Byrne, 2005). It should be pointed out that this applies to the comparison of competing models only. Different recommendations prevail in the context of measurement invariance and are described below.

We then measured the reliability of the items grouped according to the selected model. Following Gadermann et al.'s recommendations (2012) for estimating reliability coefficients with ordinal Likert-type scales, we used the polychoric correlation matrix to compute ordinal alphas³. Conventionally, the required threshold value is .70.

Regarding measurement invariance across gender, we used a stepwise multi-group CFA analysis (MGCFA). Again, the nested structure of the data was taken into account by using the TYPE=COMPLEX option of the ANALYSIS command in conjunction with the SUBPOPULATION option of the VARIABLE command. The recommended forward approach to test factorial invariance across groups relies upon a sequence of comparisons between pairs of nested models with gradually increasing parameters constraints (Dimitrov,

³ It should be noted that the empirical use of ordinal alpha is still subject to debate in the literature. Some authors advise to limit its use to a hypothetical tool. See, for example, Chalmers, R. P. (2018). On Misconceptions and the Limited Usefulness of Ordinal Alpha. *Educational and Psychological Measurement*, 78(6), 1056-1071. <https://doi.org/10.1177/0013164417727036>

2010). In the first step, baseline models were determined for both groups, then combined into a configural CFA model to test the model simultaneously across groups without equality restriction on any model parameters other than the number of factors and the patterns of factor loadings. (When applied to models with a second order-factor, each of the following steps must be tested for both the first-order and the second-order factors parameters.) In the second step, we tested *metric invariance*, that is to say invariance of the factor loadings. In the third step, we tested *scalar invariance*, which requires the intercepts to be set as equal across groups (Dimitrov, 2010). As mentioned above, the χ^2 difference test is known to reject the null hypothesis in large samples even when violations are minor in a pair of nested models. Specific recommendations have been developed for comparison of adjacent models in the context of measurement invariance. When the sample is large enough (>300), a CFI difference bigger than .01 supplemented by a difference in the RMSEA bigger than .015 or a difference in SRMR bigger than .03 indicates metric non-invariance (Chen, 2007). The same applies to the scalar invariance, except that the difference in SRMR should be larger than .01 (Chen, 2007). To our knowledge, these recommendations are applicable in the context of measurement invariance only, and there is to date no equivalent benchmarks for the comparison of competing models.

Next, using the IBM SPSS Statistics version 25 program, we compared boys and girls as an indicator of construct validity (available in Appendix B, see Table B.1). Then we computed inter-correlations between the factors emerging from the CFAs and the other variables of the study (available in Appendix B, see Table B.2). Finally, we carried out hierarchical regression analyses (Enter method, as recommended by Field, 2018) on bullying perpetration, victimization, defender behavior, and outsider behavior to explore the predictive validity of the scale. Preliminary analyses included collinearity tests. Variance inflation factors (VIF) ranged from 1 to 1.88, confirming the absence of collinearity. In the first step of

the regressions, we entered age and gender as control variables as both are well-known and frequently examined predictors of bullying perpetration, victimization, and bystander roles (Hong & Espelage, 2012). In the second step, we entered the factors of moral disengagement emerged from the CFAs. As bullying perpetration and victimization are known to be strongly correlated (for a recent meta-analysis, see Walters, 2020), the regression analyses for self-reported bullying and victimization were both run a second time with the other variable entered in the first step as control variable along with age and gender. By default, SPSS uses pairwise (e.g., in correlation analysis) or listwise (e.g., in regression analysis) exclusion of missing values. As a consequence, sample sizes indicated in the tables slightly vary from one analysis to the other.

Results

Confirmatory Factor Analyses

Table 1 displays goodness-of-fit indices of the five tested models.

[Please insert Table 1 about here]

Both Model 3 (the seven-factor model, corresponding to the mechanisms) and Model 5 (adding a higher order factor to Model 3) indicated a good fit of the observed data. The small decrease in the fit of Model 5 is not surprising given that this model is less flexible (more constrained) than Model 3. Both models presented a good fit according to the conventional thresholds, hence adding a higher order factor did not significantly worsen the fit of the model. We contend that the higher order factor structure best represents the data for the following reasons: (a) both models demonstrated a good fit; (b) there is a strong theoretical justification to consider moral disengagement as a higher order construct encompassing lower order subdimensions; (c) the higher order model is more parsimonious (as supported by the lowest BIC); and (d) correlations among the seven first-order factors – provided in Appendix B (see Table B.2) are satisfying (Byrne, 2005). Standardized factor

loadings for this model are displayed in Appendix C (Fig. C.1). All the estimated factor loadings were significant at $p < .001$, and ranged from .42 to .80.

Reliability

Ordinal Cronbach's alphas for the subscales were .64 (moral justification), .75 (euphemistic labelling), .75 (advantageous comparison), .68 (displacement of responsibility), .60 (diffusion of responsibility), .70 (distorting consequences), and .63 (victim attribution). For the global score of moral disengagement, ordinal Cronbach's alpha was .82. For a comparison with the reliability indices of the original scale (Thornberg & Jungert, 2014), see Table S.1 in supplementary material.

Measurement Invariance across Gender

A stepwise multi-group CFA analysis (MGCFA) was performed to test the invariance of the second-order factor model between girls and boys. As can be seen in Table 2, goodness-of-fit indices revealed an acceptable (CFI, TLI) to good (RMSEA, SRMR) fit for both genders when tested separately. The acceptable to good fit of the configural model indicated the same number and pattern of factor loadings across genders. CFI differences smaller than .01 provided support for first-order and second-order factor loadings invariance (M1 and M2, respectively), thus achieving metric invariance. CFI differences provided support for scalar invariance as well, since setting first-order and second-order factor intercepts equal across groups (M3 and M4, respectively) did not worsen the model fit ($\Delta\text{CFI} > -.01$). This allowed us to meaningfully compare the means of the underlying factors across genders (Dimitrov, 2010).

[Please insert Table 2 about here]

Gender Differences

The homogeneity of the variances' test indicated heterogeneity of variance for the global score of moral disengagement and all the mechanisms, the displacement of

responsibility excepted. Therefore, we used an independent sample student *t*-test for the displacement of responsibility and Welch's *t*-test for the six remaining mechanisms and the global index. The results of the analysis (see Table B.1 in Appendix B) showed that boys reported to rely more on all the mechanisms, the displacement of responsibility excepted, and had a higher global score of moral disengagement than girls. Effect sizes were small to moderate, ranging from .11 to .38.

Moral Disengagement and Bullying Perpetration

Table 3 presents the results of the two-step hierarchical regression analyses with self-reported bullying (left) and peer-reported bullying (right) as the dependent variable.

[Please insert Table 3 about here]

Five of the seven moral disengagement mechanisms were significantly and positively correlated to self-reported bullying. Fourteen percent of variance in self-reported bullying is related to these five mechanisms, as indicated by the change in R^2 . When running the analysis a second time with victimization included in step 1, the change in R^2 dropped to 10.4% (which is not represented in the table for the sake of clarity). The two mechanisms that were not related to self-reported bullying were advantageous comparison and displacement of responsibility. Only two mechanisms had very small positive correlations with peer-reported bullying: moral justification and victim attribution. Two percent of variance in peer-reported bullying was related to these two mechanisms, as indicated by the change in R^2 .

Moral disengagement and Victimization

Two-step hierarchical regression analyses were performed on self-reported victimization (left) and peer-reported victimization (right) (Table 4). Two moral disengagement mechanisms were significantly and positively associated with self-reported victimization: moral justification and victim attribution. Three percent of variance in self-reported victimization is related to these two mechanisms, as shown by the change in R^2 . This

proportion dropped to 1% when bullying is included in step 1 (which is not represented in the table for clarity purposes). Two mechanisms of moral disengagement presented very small significant correlations with peer-reported victimization; euphemistic labeling (negative), and victim attribution (positive). These two mechanisms explained 1.7% of variance in peer-reported victimization, as indicated by the R^2 change.

[Please insert Table 4 about here]

Moral disengagement and Bystanders' Behavior

To examine moral disengagement mechanisms associated with bystanders' behavior, two-step hierarchical regression analyses were carried out on defender behavior (left) and outsider behavior (right) (Table 5). Euphemistic labeling, distorting consequences, diffusion of responsibility, and victim attribution were negatively associated with defender behavior. Displacement of responsibility, on the other hand, presented a positive association, while moral justification and advantageous comparison were not related to defender behavior. Almost five percent of variance in defender behavior is related to these five mechanisms, as shown by the R^2 change. Euphemistic labeling, distorting consequences, diffusion of responsibility, and victim attribution were positively associated to outsider behavior, but moral justification, advantageous comparison, and displacement of responsibility were not related to it. The change in R^2 indicates that almost seven percent of variance in outsider behavior is related to these four mechanisms.

[Please insert Table 5 about here]

Discussion

This study aimed to validate the French version of the Moral Disengagement in Bullying Scale (MDBS; Thornberg & Jungert, 2014) in a sample of 8- to 12-year-old children. First, we examined the Fr-MDBS's construct validity by testing its factorial structure. Both the seven-factor model and its nested higher order-factor model had a good fit

to the data. This supports the existence of an overarching dimension (moral disengagement) that encompassed the seven first-order factors, corresponding to the mechanisms. This result is important for both theory and research. From a theoretical perspective, the discrepancy between Bandura's conceptual model and the factorial structures of the developed scales was questioning either the theoretical model or its operationalization. The present study has addressed this issue by validating a factorial structure that is more congruent with the conceptual model. To our knowledge, only one study using the MDBS has demonstrated such a factorial solution to be better than a single dimension structure (Bjärehed et al., 2020), and none has demonstrated it to be as good as a first-order multi-dimension structure so far. From a research perspective, being limited to a single dimension of moral disengagement can lead to missing insight into the potential mechanisms involved in both bullies' and bystanders' behavior (Pozzoli et al., 2012). Yet, examining moral disengagement as a single dimension remains relevant given that moral disengagement mechanisms are conceptually believed to operate together and given the strong inter-correlations between them (Thornberg & Jungert, 2014). The validation of a second-order factor structure entitles researchers to examine either the mechanisms or the global dimension of moral disengagement, according to their research goals.

Second, we tested the Fr-MDBS's measurement invariance across gender. As scalar invariance was reached, we could meaningfully compare boys and girls (Dimitrov, 2010). Gender differences followed the expected direction, which supports construct validity: Positive associations between being a boy and the use of all moral disengagement mechanisms were found, except for the displacement of responsibility. A specific subsection will be devoted to the findings related to this mechanism below.

Predictive validity of the Fr-MDBS was tested by examining associations between bullying perpetration, victimization, bystander roles (i.e., defender and outsider), and the

factors of moral disengagement emerging from the CFAs. We expected bullying to be positively linked to moral disengagement factors (Gini et al., 2014; Killer et al., 2019). This study corroborated previous research, as most of the moral disengagement mechanisms were significantly and positively related to self-reported bullying, even when taking into account potential concurrent victimization. As in previous studies (Bjärehed et al., 2020; Thornberg & Jungert, 2014), both moral justification and victim attribution were significantly related to self-reported bullying. In addition, the use of euphemistic labeling, diffusion of responsibility, and distorting consequences was also related to self-reported bullying in our results. One explanation might be that the larger size of our sample enabled us to detect more significant associations. When considering peer-reported bullying, only moral justification and victim attribution were significantly related. This may lend support to a greater role of these two specific mechanisms in bullying behavior, which is in line with previous findings (Bjärehed et al., 2020; Thornberg & Jungert, 2014).

We hypothesized that victimization would be positively associated with moral disengagement (Killer, 2019). Our results confirmed this assumption, as children who tended to use moral justification and victim attribution as moral disengagement mechanisms were more likely to report being victimized, even when self-reported bullying perpetration was controlled for. Besides, victim attribution was also associated with peer-reported victimization. On the other hand, children who reported using more euphemistic labeling (i.e., using innocuous words to qualify bullying) were less likely to be designated by their peers as victims. This might be explained by a greater awareness among victims of the harmful effects bullying has on them. The association of victim attribution with both self- and peer-reported victimization is consistent with the tendency toward self-blame displayed by students experiencing chronic victimization (Graham & Juvonen, 1998; Tolmatcheff et al., 2019). In contrast with these results, Thornberg and Jungert (2014) did not find any mechanism

associated with victimization when concurrent bullying perpetration was controlled. This is likely to be caused by insufficient statistical power due to the sample size, as suggested by the very small effect size of the associations found in our study as well as in meta-analyses (Killer, 2019).

Following recent recommendations (e.g., Hymel & Swearer, 2015; Volk et al., 2017), the present study used a combination of self-reported and peer-reported data for bullying perpetration and victimization. Because of shared method variance, we expected stronger associations between moral disengagement and both bullying perpetration and victimization with self-reported data than with peer-reported data (Killer et al., 2019). Our results proved conclusive: moral disengagement mechanisms explained 12% more of the variance in self-reported bullying compared to peer-reported bullying, and 1.3% more of the variance in self-reported victimization compared to peer-reported victimization. It is noteworthy that expected associations, although weaker and fewer, can also be found with peer-reported measures.

Regarding bystanders' behavior, we assumed a negative association between moral disengagement and the defender role (Killer et al., 2019) but a positive association between moral disengagement and the outsider's role (e.g., Doramajian & Bukowski, 2015; Sjögren et al., 2020; Thornberg et al., 2017). These assumptions were mostly confirmed. As in Thornberg and Jungert's study (2014), defending was negatively predicted by the use of diffusion of responsibility and victim attribution. Furthermore, in our results, defending was also negatively predicted by using euphemistic labeling and distorting consequences. On the other hand, we found that children who reported using displacement of responsibility were more likely to report acting as defenders. Regarding the outsider role, we found a positive association with euphemistic labeling, distorting consequences, diffusion of responsibility, and victim attribution. These four mechanisms appear to play a particular role in bystander behavior as children who reported using them were both more likely to passively observe or

walk away when bullying takes place and less likely to comfort the victim or stand up against the bully.

Displacement of responsibility presented an intriguing pattern throughout our results. This mechanism did not predict bullying perpetration (whether self- or peer-reported) or outsider behavior and was the only mechanism that was not used more by boys than by girls. In contrast, children who reported this mechanism were more likely to act as defenders. This result may be explained by children's unexpected reaction to the three corresponding items as observed by the researchers during the data collection phase. Many children reacted in an empathetic way, expressing compassion for the fictional children in the items that were having "family problems", had "parents who are mean or violent toward them", or had had "a poor upbringing", and agreed that these unfortunate children could not be blamed for their consequential bullying behavior. We can imagine that differently phrased items, measuring displacement of responsibility without eliciting the respondents' empathy for the fictional bullies, might yield different results. No significant link with defending was found in the original validation study (Thornberg & Jungert, 2014), but this could again be related to the small sample size. In contrast to these authors' and our results, Bjärehed et al. (2020) found that boys used displacement of responsibility more than girls. However, no significant associations with bullying perpetration were found in the regression analyses of any of the three studies.

Limitations and Future Directions

This study has some limitations. First, some of the subscales presented moderate ordinal alphas (e.g., diffusion of responsibility = .60), which weakens their reliability. Note that the original scale (Thornberg & Jungert, 2014) presented several low reliability indices (even below .60), as indicated in Appendix A. Yet, these rather low alphas can be explained by the small numbers of items in the subscales (Thornberg & Jungert, 2014). Second, as

mentioned before, displacement of responsibility was unexpectedly interpreted by the children, which may have concealed this mechanism's actual contribution to bullying perpetration and bystanders' behavior. Future studies might test differently-worded items to better capture this moral disengagement mechanism. Some suggestions of items could be "If no adult is present to supervise the playground during recess, children cannot be blamed for bullying others" (inspired from Hymel et al., 2005) or "If someone got up on the wrong side of the bed, it is not his or her fault if he or she is being mean with other students". Third, we opted for behavior-based (as opposed to definition-based) self-reported measures of bullying perpetration and victimization (e.g., Brighi et al., 2012; Espelage et al., 2003), which do not specify the three definitional characteristics of bullying (repetition, intentionality, power imbalance). While this behavioral approach enables researchers to avoid some shortcomings of providing a definition or even using the word "bullying" in surveys (see Kert et al., 2010), it may also capture more generalized peer aggression and peer victimization. Fourth, as *dehumanization* and *blaming the victim* were originally merged into one factor (Thornberg & Jungert, 2014), the scale only taps into seven out of the eight theoretical mechanisms. This could imply either that the conceptual distinction between these two subdimensions does not empirically exist or that the items contained in the original pool could not capture them. Although the seven-factor structure already provides a rich picture of moral disengagement mechanisms, future studies should re-examine the existence of these two distinct factors. Finally, the directionality of the relationship between moral disengagement and bullying perpetration is not yet clearly established (Gini et al., 2014). While some studies suggest that moral disengagement predicts bullying behavior (e.g., Barchia & Bussey, 2011; Sticca & Perren, 2015; Troop-Gordon et al., 2019), others indicate the opposite (e.g., Obermann, 2013). The association might also be bidirectional, with moral disengagement enabling bullying perpetration and, in turn, repeated bullying perpetration making the use of the

mechanisms easier (Gini et al., 2014; Mazzone et al., 2019). Longitudinal studies, and in particular longitudinal intervention studies that aim to decrease bullying perpetration by challenging students' use of moral disengagement mechanisms, are needed to clarify this issue (Gini et al., 2014).

Implications

As far as we know, only one study has assessed an intervention that explicitly targeted moral disengagement to help reduce bullying in primary schools, providing mixed but promising results (Wang & Goldberg, 2017). Several researchers have called for further research to determine the potential of this approach (Almeida et al., 2009; Hymel & Bonanno, 2014), notably given that such interventions could be a way of targeting both bystanders and bullies (Obermann, 2011; Wang & Goldberg, 2017). The rationale for targeting moral disengagement in anti-bullying interventions is that outlining and invalidating the moral disengagement mechanisms should prevent children from using them. If they can no longer rely on these justifications, they will expose themselves to self-sanctions (guilt, shame), which, in turn, should discourage immoral behavior. In other words, bullying perpetration or refraining from defending the victim is expected to be less likely if children can no longer justify their behavior and therefore undergo unpleasant internal sanctions. Sound knowledge of the moral disengagement mechanisms could enable school psychologists and school staff in general to detect and invalidate them when dealing with a bullying situation.

Practically speaking, using storybooks displaying bullying situations to illustrate the use of moral disengagement mechanisms (e.g., Wang & Goldberg, 2017) seems to be a promising avenue to target moral disengagement. Bibliotherapy is often used by school psychologists and teachers with regard to conduct problems and can easily be integrated into the regular instruction curriculum. This makes it a particularly suitable approach for schools given their well-known time constraints (Wang & Goldberg, 2017). Moreover, integration of

bullying prevention into daily instruction has been suggested as facilitating skill generalization (Leadbeater et al., 2016; Swearer et al., 2009).

Extending Thornberg and Jungert's work, the present study provides greater insight into the moral disengagement mechanisms related to bullying perpetration, victimization, and bystanders' roles, which can help to design more effective anti-bullying interventions. In that respect, interventions should address victim attribution in priority, given that this mechanism appeared to be related to all the investigated variables. By targeting victim attribution, interventions could prevent bullies from dehumanizing and blaming their victims, but also remind the victims that they are not responsible for their plight, exhort defenders to keep supporting the victims, and finally encourage outsiders to react when bullying occurs. Euphemistic labeling, diffusion of responsibility, and distorting the consequences were involved in both bullying perpetration and bystander (defender and outsider) roles, and are therefore worth targeting for anti-bullying interventions too. In particular, the role of diffusion of responsibility in dissuading bystanders from intervening is well already known (Salmivalli, 2010). Since moral justification was associated with both self-reported and peer-reported bullying, it could be targeted in selective interventions designed for bullies specifically. As highlighted in previous studies (Pornari & Wood, 2010; Thornberg & Jungert, 2014), when bullies use moral justification, they are likely to experience positive feelings for serving a moral purpose, which, in turn, will ease their bullying behavior. This mechanism, in particular, thus seems to perniciously reinforce bullying perpetration. Finally, advantageous comparison was not related to any of the examined variables. More studies are needed to demonstrate that this mechanism does not play any role in bullying dynamics, but this result suggests that anti-bullying interventions should not focus on it.

In conclusion, the validation of the Fr-MBDS empirically tested a second-order factor structure, thus narrowing the gap between Bandura's conceptual model of moral

disengagement and its operationalization. This study contributes to enable researchers to validly examine either the mechanisms or the global dimension of moral disengagement, depending on their research interests. Adding to the current literature, this study showed that more moral disengagement mechanisms are involved in bullying perpetration, victimization, defender behavior, and outsider behavior than was previously thought. This research also provided more detailed information about the respective contribution of these mechanisms to the bullying dynamics and brought useful insight into the mechanisms which can be prioritized for the design of anti-bullying interventions. Future intervention studies targeting moral disengagement in children should move beyond the single dimension of moral disengagement and investigate their effects on the distinct mechanisms in both bullies and bystanders.

Acknowledgements

We would like to thank Aurélie Bertrand, Michaël Parmentier, and Thibault Coppe for their advice on the statistical analyses. We are grateful to René Veenstra and his WALM research group at the University of Groningen, the Netherlands, for their advice and valuable suggestions for improving the paper.

Funding: This research was funded by a PhD grant (FC23715) from the National Fund for Scientific Research (F.R.S.-F.N.R.S.).

References

- Allison, K. R., & Bussey, K. (2017). Individual and collective moral influences on intervention in cyberbullying. *Computers in Human Behavior*, 74, 7-15.
<https://doi.org/10.1016/j.chb.2017.04.019>
- Almeida, A., Correia, I., & Marinho, S. (2009). Moral disengagement, normative beliefs of peer group, and attitudes regarding roles in bullying. *Journal of School Violence*, 9(1), 23-36. <https://doi.org/10.1080/15388220903185639>

- Álvarez-García, D., García, T., & Núñez, J. C. (2015). Predictors of school bullying perpetration in adolescence: A systematic review. *Aggression and Violent Behavior, 23*, 126-136. <http://dx.doi.org/10.1016/j.avb.2015.05.007>
- Bandura, A. (1986). *Social foundations of thought and action*. Englewood Cliffs, NJ, 23-28.
- Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology, 71*(2), 364-374. <https://doi.org/10.1037/0022-3514.71.2.364>
- Bandura, A., Caprara, G. V., Barbaranelli, C., Pastorelli, C., & Regalia, C. (2001). Sociocognitive self-regulatory mechanisms governing transgressive behavior. *Journal of Personality and Social Psychology, 80*(1), 125-135. <https://doi.org/10.1037/0022-3514.80.1.125>
- Barchia, K., & Bussey, K. (2011). Individual and collective social cognitive influences on peer aggression: Exploring the contribution of aggression efficacy, moral disengagement, and collective efficacy. *Aggressive Behavior, 37*(2), 107-120. <https://doi.org/10.1002/ab.20375>
- Bjärehed, M., Thornberg, R., Wänström, L., & Gini, G. (2020). Mechanisms of moral disengagement and their associations with indirect bullying, direct bullying, and pro-aggressive bystander behavior. *The Journal of Early Adolescence, 40*(1), 28-55. <https://doi.org/10.1177/0272431618824745>
- Brighi, A., Ortega, R., Scheitauer, H., Smith, P. K., Tsormpatzoudis, C., Barkoukis, V., Del Rey, R., & Thompson, J. (2012). European bullying intervention project questionnaire (EBIPQ). University of bologna.
- Byrne, B. M. (2005). Factor analytic models: Viewing the structure of an assessment instrument from three perspectives. *Journal of Personality Assessment, 85*(1), 17-32. https://doi.org/10.1207/s15327752jpa8501_02

- Carney, A. G., & Merrell, K. W. (2001). Bullying in schools: Perspectives on understanding and preventing an international problem. *School Psychology International*, 22(3), 364-382. <https://doi.org/10.1177/0143034301223011>
- Caprara, G. V., Pastorelli, C., & Bandura, A. (1995). La misura del disimpegno morale in età evolutiva. *Età Evolutiva*, 51, 18-29.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233-255. https://doi.org/10.1207/S15328007SEM0902_5
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14(3), 464-504. <https://doi.org/10.1080/10705510701301834>
- Cillessen, A. H., & Marks, P. E. (2017). Methodological choices in peer nomination research. *New Directions for Child and Adolescent Development*, 157, 21-44. <https://doi.org/10.1002/cad.20206>
- Crapanzano, A. M., Frick, P. J., Childs, K., & Terranova, A. M. (2011). Gender differences in the assessment, stability, and correlates to bullying roles in middle school children. *Behavioral Sciences & the Law*, 29(5), 677-694. <https://doi.org/10.1002/bsl.1000>
- Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, 43(2), 121-149. <https://doi.org/10.1177/0748175610373459>
- Doramajian, C., & Bukowski, W. M. (2015). A longitudinal study of the associations between moral disengagement and active defending versus passive bystanding during bullying situations. *Merrill-Palmer Quarterly*, 61(1), 144-172. <https://doi.org/10.13110/merrpalmquar1982.61.1.0144>

- Dziak, J. J., Coffman, D. L., Lanza, S. T., Li, R., & Jermiin, L. S. (2020). Sensitivity and specificity of information criteria. *Briefings in Bioinformatics*, 21(2), 553-565.
<https://doi.org/10.1093/bib/bbz016>
- Espelage, D. L., Holt, M. K., & Henkel, R. R. (2003). Examination of peer-group contextual effects on aggression during early adolescence. *Child Development*, 74(1), 205-220.
<https://doi.org/10.1111/1467-8624.00531>
- Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics* 5th ed.
- Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research, and Evaluation*, 17(1), 3.
<https://doi.org/10.7275/n560-j767>
- Galand, B., Hospel, V. & Baudoin, N. (2014). Prévenir le harcèlement via les pratiques de classe ? Une étude multiniveaux. [Preventing bullying through classroom practices? A multilevel study.] *Revue Québécoise de Psychologie*, 35(3), p. 137-157.
<http://hdl.handle.net/2078.1/152752>
- Garandeau, C. F., Yanagida, T., Vermande, M. M., Strohmeier, D., & Salmivalli, C. (2019). Classroom size and the prevalence of bullying and victimization: testing three explanations for the negative association. *Frontiers in Psychology*, 10, 21-25.
<https://doi.org/10.3389/fpsyg.2019.02125>
- Gini, G. (2006). Social cognition and moral cognition in bullying: What's wrong?. *Aggressive Behavior*, 32(6), 528-539. <https://doi.org/10.1002/ab.20153>
- Gini, G., Pozzoli, T., & Bussey, K. (2015). The role of individual and collective moral disengagement in peer aggression and bystanding: A multilevel analysis. *Journal of Abnormal Child Psychology*, 43(3), 441-452. <https://doi.org/10.1007/s10802-014-9920-7>

- Gini, G., Pozzoli, T., & Hauser, M. (2011). Bullies have enhanced moral competence to judge relative to victims, but lack moral compassion. *Personality and Individual Differences*, 50(5), 603-608. <https://doi.org/10.1016/j.paid.2010.12.002>
- Gini, G., Pozzoli, T., & Hymel, S. (2014). Moral disengagement among children and youth: A meta-analytic review of links to aggressive behavior. *Aggressive Behavior*, 40(1), 56-68. <https://doi.org/10.1002/ab.21502>
- Graham, S., & Juvonen, J. (1998). Self-blame and peer victimization in middle school: an attributional analysis. *Developmental Psychology*, 34(3), 587. <https://doi.org/10.1037/0012-1649.34.3.587>
- Hong, J. S., & Espelage, D. L. (2012). A review of research on bullying and peer victimization in school: An ecological system analysis. *Aggression and Violent Behavior*, 17(4), 311-322. <https://doi.org/10.1016/j.avb.2012.03.003>
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- Hymel, S., & Bonanno, R. A. (2014). Moral disengagement processes in bullying. *Theory into Practice*, 53(4), 278-285. <https://doi.org/10.1080/00405841.2014.947219>
- Hymel, S., Schonert-Reichl, K. A., Bonanno, R. A., Vaillancourt, T., & Rocke-Henderson, N. (2010). Bullying and morality: Understanding how good kids can behave badly. In S. R. Jimerson, S. M. Swearer, & D. L. Espelage (Eds.), *Handbook of bullying in schools: An international perspective* (pp. 101–118). Routledge/Taylor & Francis Group.
- Hymel, S., & Swearer, S. M. (2015). Four decades of research on school bullying: An introduction. *American Psychologist*, 70(4), 293. <https://doi.org/10.1037/a0038928>

- Hymel, S., Rocke-Henderson, N., & Bonanno, R. A. (2005). Moral disengagement: A framework for understanding bullying among adolescents. *Journal of Social Sciences*, 8(1), 1-11.
- Jenkins, L. N., & Troop-Gordon, W. (2020). Bystanders in bullying : Introduction to the special issue. *Journal of School Psychology*, 81, 47-50.
<https://doi.org/10.1016/j.jsp.2020.07.002>
- Kert, A. S., Coddington, R. S., Tryon, G. S., & Shiyko, M. (2010). Impact of the word “bully” on the reported rate of bullying behavior. *Psychology in the Schools*, 47(2), 193-204.
<https://doi.org/10.1002/pits.20464>
- Killer, B., Bussey, K., Hawes, D. J., & Hunt, C. (2019). A meta-analysis of the relationship between moral disengagement and bullying roles in youth. *Aggressive Behavior*, 45(4), 450-462. <https://doi.org/10.1002/ab.21833>
- Leadbeater, B. J., Thompson, K., & Sukhawathanakul, P. (2016). Enhancing social responsibility and prosocial leadership to prevent aggression, peer victimization, and emotional problems in elementary school children. *American Journal of Community Psychology*, 58(3-4), 365-376. <https://doi.org/10.1002/ajcp.12092>
- Little, T. D. (2013). *Longitudinal structural equation modeling*. Guilford press.
- Mazzone, A., Yanagida, T., CS Caravita, S., & Strohmeier, D. (2019). Moral emotions and moral disengagement: Concurrent and longitudinal associations with aggressive behavior among early adolescents. *Journal of Early Adolescence*, 39(6), 839-863.
<https://doi.org/10.1177/0272431618791276>
- Obermann, M. L. (2011). Moral disengagement in self-reported and peer-nominated school bullying. *Aggressive Behavior*, 37(2), 133-144. <https://doi.org/10.1002/ab.20378>

- Obermann, M. L. (2013). Temporal aspects of moral disengagement in school bullying: Crystallization or escalation?. *Journal of School Violence*, 12(2), 193-210.
<https://doi.org/10.1080/15388220.2013.766133>
- Paciello, M., Fida, R., Tramontano, C., Lupinetti, C., & Caprara, G. V. (2008). Stability and change of moral disengagement and its impact on aggression and violence in late adolescence. *Child Development*, 79(5), 1288-1309. <https://doi.org/10.1111/j.1467-8624.2008.01189.x>
- Pelton, J., Gound, M., Forehand, R., & Brody, G. (2004). The moral disengagement scale: Extension with an American minority sample. *Journal of Psychopathology and Behavioral Assessment*, 26(1), 31-39.
<https://doi.org/10.1023/B:JOBA.0000007454.34707.a5>
- Perren, S., Gutzwiller-Helfenfinger, E., Malti, T., & Hymel, S. (2012). Moral reasoning and emotion attributions of adolescent bullies, victims, and bully-victims. *British Journal of Developmental Psychology*, 30(4), 511-530. <https://doi.org/10.1111/j.2044-835X.2011.02059.x>
- Pornari, C. D., & Wood, J. (2010). Peer and cyber aggression in secondary school students: The role of moral disengagement, hostile attribution bias, and outcome expectancies. *Aggressive Behavior*, 36(2), 81-94. <https://doi.org/10.1002/ab.20336>
- Pozzoli, T., Gini, G., & Vieno, A. (2012). Individual and class moral disengagement in bullying among elementary school children. *Aggressive Behavior*, 38(5), 378-388.
<https://doi.org/10.1002/ab.21442>
- Pozzoli, T., Gini, G., & Thornberg, R. (2016). Bullying and defending behavior: The role of explicit and implicit moral cognition. *Journal of School Psychology*, 59, 67-81.
<https://doi.org/10.1016/j.jsp.2016.09.005>

- Ruggieri, S., Friemel, T., Sticca, F., Perren, S., & Alsaker, F. (2013). Selection and influence effects in defending a victim of bullying: The moderating effects of school context. *Procedia-Social and Behavioral Sciences*, 79, 117-126.
<https://dx.doi.org/10.1016/j.sbspro.2013.05.060>
- Salmivalli, C. (2001). Group view on victimization. *Peer Harassment in School*, 398-417.
- Salmivalli, C. (2010). Bullying and the peer group: A review. *Aggression and Violent Behavior*, 15(2), 112-120. <https://doi.org/10.1016/j.avb.2009.08.007>
- Salmivalli, C., Lagerspetz, K., Björkqvist, K., Österman, K., & Kaukiainen, A. (1996). Bullying as a group process: Participant roles and their relations to social status within the group. *Aggressive Behavior*, 22(1), 1-15. [https://doi.org/10.1002/\(SICI\)1098-2337\(1996\)22:1<1::AID-AB1>3.0.CO;2-T](https://doi.org/10.1002/(SICI)1098-2337(1996)22:1<1::AID-AB1>3.0.CO;2-T)
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8(2), 23-74. <http://www.mpr-online.de>
- Sjögren, B., Thornberg, R., Wänström, L., & Gini, G. (2020). Bystander behaviour in peer victimisation: moral disengagement, defender self-efficacy and student-teacher relationship quality. *Research Papers in Education*, 1-23.
<https://doi.org/10.1080/02671522.2020.1723679>
- Sticca, F., & Perren, S. (2015). The chicken and the egg: Longitudinal associations between moral deficiencies and bullying: A parallel process latent growth model. *Merrill-Palmer Quarterly*, 61(1), 85-100. <https://doi.org/10.13110/merrpalmquar1982.61.1.0085>
- Sutton, J., & Smith, P. K. (1999). Bullying as a group process: An adaptation of the participant role approach. *Aggressive Behavior*, 25(2), 97-111.
[https://doi.org/10.1002/\(SICI\)1098-2337\(1999\)25:2<97::AID-AB3>3.0.CO;2-7](https://doi.org/10.1002/(SICI)1098-2337(1999)25:2<97::AID-AB3>3.0.CO;2-7)

- Swearer, S. M., Espelage, D. L., & Napolitano, S. A. (2009). *Bullying prevention and intervention: Realistic strategies for schools*. Guilford press.
- Thornberg, R., & Jungert, T. (2013). Bystander behavior in bullying situations: Basic moral sensitivity, moral disengagement and defender self-efficacy. *Journal of Adolescence*, 36(3), 475-483. <https://doi.org/10.1016/j.adolescence.2013.02.003>
- Thornberg, R., & Jungert, T. (2014). School bullying and the mechanisms of moral disengagement. *Aggressive Behavior*, 40(2), 99-108. <https://doi.org/10.1002/ab.21509>
- Thornberg, R., Wänström, L., Hong, J. S., & Espelage, D. L. (2017). Classroom relationship qualities and social-cognitive correlates of defending and passive bystanding in school bullying in Sweden: A multilevel analysis. *Journal of School Psychology*, 63, 49-62. <https://doi.org/10.1016/j.jsp.2017.03.002>
- Tolmatcheff, C., Bebiroglu, N., & Galand, B. (2020) L'aménagement des cours de récréation et l'organisation d'espaces de parole permettent-ils de réduire le harcèlement scolaire ? [Can redesigning the school playground and setting up classroom forums help to reduce bullying?] *Revue française de Pédagogie*, 209, 129-147. <https://doi.org/10.4000/rfp.9846>
- Tolmatcheff, C., Hénoumont, F., Klée, E., & Galand, B. (2019). Stratégies et réactions des victimes et de leur entourage face au harcèlement scolaire : une étude rétrospective [Strategies and reactions of school bullying victims and their entourage : A retrospective study] *Psychologie française*, 64(4), 391-407. <https://doi.org/10.1016/j.psfr.2018.09.002>
- Troop-Gordon et al. (2019). Predicting the development of pro-bullying bystander behavior : A short-term longitudinal analysis. *Journal of School Psychology*, 77, 77-89. <https://doi.org/10.1016/j.jsp.2019.10.004>

Volk, A. A., Veenstra, R., & Espelage, D. L. (2017). So you want to study bullying?

Recommendations to enhance the validity, transparency, and compatibility of bullying research. *Aggression and Violent Behavior*, 36, 34-43.

<https://doi.org/10.1016/j.avb.2017.07.003>

Walters, G. D. (2020). School-age bullying victimization and perpetration: A meta-analysis of prospective studies and research. *Trauma, Violence, & Abuse*.

<https://doi.org/10.1177/1524838020906513>

Wang, C., & Goldberg, T. S. (2017). Using children's literature to decrease moral

disengagement and victimization among elementary school students. *Psychology in the Schools*, 54(9), 918-931. <https://doi.org/10.1002/pits.22042>

Wolke, D., & Lereya, S. T. (2015). Long-term effects of bullying. *Archives of Disease in*

Childhood, 100(9), 879-885. <http://dx.doi.org/10.1136/archdischild-2014-306667>

Table 1. Goodness-of-fit indices of Confirmatory Analyses of the five models

Models	S-B χ^2	df	RMSEA	90% CI for RMSEA		CFI	TLI	SRMR	AIC	BIC	
				LL	UL						
Model 1 (F1)	701.4*	135	.062	.057	.066	.78	.75	.056	62513.9	62784.1	Poor fit
Model 2 (F4)	575.6*	129	.056	.051	.061	.83	.79	.050	62354.5	62654.7	Poor fit
Model 3 (F7)	198.9*	114	.026	.020	.032	.97	.96	.029	61907.7	62282.9	Good fit
Model 4 (F4+1)	609.7*	132	.057	.053	.062	.81	.78	.057	62392.4	62677.6	Poor fit
Model 5 (F7+1)	255.5*	128	.030	.025	.035	.95	.94	.033	61950.4	62255.6	Good fit

$N = 1100$

* $p < .05$

F1 = single factor model; F4 = four first-order factor model (see Pozzoli et al., 2012); F7 = seven first-order factor model (see Thornberg & Jungert, 2014); F4+1 = second-order factor model with four first-order factors; F7+1 = second-order factor model with seven first-order factors

Table 2. Testing for Factorial Invariance Across Genders

Model	χ^2	df	RMSEA	90% CI for RMSEA		Δ RMSEA	CFI	Δ CFI	TLI	SRMR	Δ SRMR
				LL	UL						
Baseline girls ($N = 575$)	199.8*	128	.031	.023	.039	—	.942	—	.931	.040	—
Baseline boys ($N = 525$)	214.9*	128	.036	.027	.044	—	.933	—	.919	.044	—
M0 configural model	413.9*	256	.033	.027	.039	—	.938	—	.925	.042	—
M1 1st order- metric invariance	422.6*	267	.033	.027	.038	0	.939	.001	.930	.044	.002
M2 2nd order- metric invariance	433.6*	273	.033	.027	.038	0	.937	-.002	.929	.045	.001
M3 1st order- scalar invariance	467.5*	284	.034	.029	.040	.001	.928	-.009	.922	.047	.002
M4 2nd order- scalar invariance	473.2*	290	.034	.028	.039	0	.928	0	.924	.047	0

$N = 1100$

Table 3. Regression Analysis for Bullying Perpetration

	Self-reported bullying				Peer-reported bullying			
	<i>B</i>	<i>SE B</i>	β	R^2	<i>B</i>	<i>SE B</i>	β	R^2
Step 1				.025				.068
Age	-.021	.015	-.041		-.013	.004	-.092**	
Gender (girls = 0, boys = 1)	.159	.031	.153***		.067	.008	.243***	
Step 2				.169				.087
Age	-.010	.014	-.020		-.011	.004	-.080**	
Gender (girls = 0, boys = 1)	.079	.030	.075**		.060	.008	.218***	
Moral justification	.045	.014	.100**		.008	.004	.064*	
Euphemistic labeling	.062	.015	.134***		.000	.004	-.003	
Advantageous comparison	-.013	.019	-.026		-.002	.005	-.012	
Displacement of responsibility	.013	.014	.028		.000	.004	.000	
Diffusion of responsibility	.055	.017	.106**		.006	.005	.047	
Distorting consequences	.049	.024	.078*		.003	.007	.020	
Victim attribution	.105	.022	.154***		.013	.006	.072*	

Note. Self-reported bullying (left): $F(2, 1086) = 13.9, p < .001$ for the first model; $F(7, 1079) = 24.5, p < .001$ for the second model; $\Delta R^2 = .145$. Peer-reported bullying (right): $F(2, 1094) = 39.6, p < .001$ for the first model; $F(9, 1087) = 11.5, p < .001$ for the second model; $\Delta R^2 = .019$.

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 4. Regression Analysis for Victimization

	Self-reported victimization				Peer-reported victimization			
	<i>B</i>	<i>SE B</i>	β	R^2	<i>B</i>	<i>SE B</i>	β	R^2
Step 1				.029				.035
Age	-.130	.023	-.168***		-.022	.003	-.188***	
Gender (girls = 0, boys = 1)	.041	.047	.026		.004	.007	.015	
Step 2				.062				.052
Age	-.117	.023	-.151***		-.021	.004	-.180***	
Gender (girls = 0, boys = 1)	.000	.048	.000		.004	.007	.015	
Moral justification	.050	.022	.074*		.000	.003	-.004	
Euphemistic labeling	.002	.023	.003		-.011	.004	-.105**	
Advantageous comparison	.023	.030	.030		-.003	.005	-.030	
Displacement of responsibility	.016	.022	.023		-.005	.003	-.044	
Diffusion of responsibility	-.015	.028	-.019		.005	.004	.043	
Distorting consequences	-.026	.038	-.027		.005	.006	.035	
Victim attribution	.154	.036	.150***		.012	.005	.076*	

Note. Self-reported victimization (left): $F(2, 1090) = 16.3, p < .001$ for the first model; $F(9, 1083) = 7.9, p < .001$ for the second model; $\Delta R^2 = .033$. Peer-reported victimization (right): $F(2, 1094) = 20.1, p < .001$ for the first model; $F(9, 1087) = 6.7, p < .001$ for the second model; $\Delta R^2 = .017$.

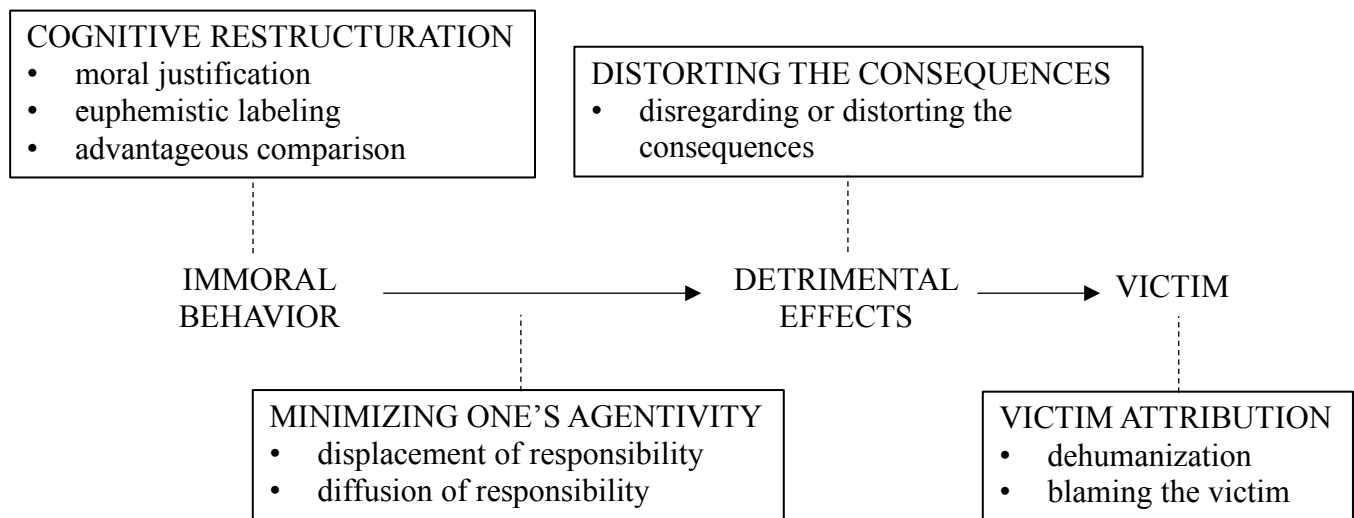
* $p < .05$, ** $p < .01$, *** $p < .001$

Table 5. Regression Analysis for Bystander's behavior

	Defender behavior				Outsider behavior			
	<i>B</i>	<i>SE B</i>	β	R^2	<i>B</i>	<i>SE B</i>	β	R^2
Step 1				.029				.019
Age	-.007	.028	-.008		.010	.025	.012	
Gender (girls = 0, boys = 1)	-.323	.056	-.171***		.236	.052	.137***	
Step 2				.076				.086
Age	-.011	.027	-.011		.016	.025	.019	
Gender (girls = 0, boys = 1)	-.261	.057	-.138***		.154	.051	.090**	
Moral justification	.030	.026	.037		-.006	.024	-.009	
Euphemistic labeling	-.064	.028	-.076*		.067	.025	.087**	
Advantageous comparison	.038	.036	.042		-.043	.032	-.052	
Displacement of responsibility	.067	.026	.079*		-.038	.024	-.050	
Diffusion of responsibility	-.077	.033	-.084*		.114	.030	.134***	
Distorting consequences	-.118	.045	-.104**		.120	.041	.116**	
Victim attribution	-.117	.042	-.095**		.093	.038	.083*	

Note. Defender behavior: $F(2, 1088) = 16.5, p < .001$ for the first model; $F(9, 1081) = 9.9, p < .001$ for the second model; $\Delta R^2 = .047$. Outsider behavior: $F(2, 1083) = 10.5, p < .001$ for the first model; $F(9, 1076) = 11.3, p < .001$ for the second model; $\Delta R^2 = .067$.

* $p < .05$, ** $p < .01$, *** $p < .001$

Figure 1*Bandura's Framework of Moral Disengagement*

Note. This figure was adapted from “Mechanisms of moral disengagement in the exercise of moral agency,” by A. Bandura, C. Barbaranelli, G. V. Caprara, and C. Pastorelli, 1996 *Journal of Personality and Social Psychology*, 71(2), p. 365. (<https://doi.org/10.1037/0022-3514.71.2.364>) Copyright 1996 by the American Psychological Association.

Appendix A.

Reliability indices of the global index and seven factors of moral disengagement in the original validation study of the MDBS (Thornberg & Jungert, 2014): Moral justification ($r = .59$); Euphemistic labeling (Cronbach's $\alpha = .65$); Advantageous comparison ($r = .49$); Displacement of responsibility (Cronbach's $\alpha = .70$); Diffusion of responsibility ($r = .26$); Distorting consequences ($r = .56$); Victim attribution (Cronbach's $\alpha = .64$); Global index of moral disengagement (Cronbach's $\alpha = .84$).

Results of the Confirmatory Factor Analysis of the original validation study of the MDBS (Thornberg & Jungert, 2014): Satorra-Bentler scaled $\chi^2(114, N=372) = 118.56, p < .366$, CFI = .99, RMSEA = .01, 90% CI .00 - .03

Appendix B.*Table B.1. Gender differences in Moral Disengagement mechanisms*

Moral Disengagement variables	Total		Girls		Boys		95% Confidence Interval				
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i> -value	Lower	Upper	Cohen's <i>d</i>
Moral justification	1.72	1.16	1.54	1.09	1.9	1.21	-5.19 ^a	<.001	0.50	0.23	.32
Euphemistic labeling	1.05	1.13	0.91	1.03	1.2	1.21	-4.21 ^a	<.001	0.42	0.15	.26
Advantageous comparison	1	1.05	0.85	0.97	1.17	1.11	-5.16 ^a	<.001	0.45	0.20	.31
Displacement of responsibility	1.75	1.13	1.7	1.11	1.82	1.14	-1.74	.082	0.25	0.01	.11
Diffusion of responsibility	0.88	1.02	0.72	0.90	1.04	1.11	-5.21 ^a	<.001	0.44	0.20	.32
Distorting consequences	0.82	0.84	0.69	0.79	0.96	0.86	-5.38 ^a	<.001	0.37	0.17	.33
Victim attribution	0.61	0.77	0.53	0.71	0.69	0.83	-3.54 ^a	<.001	0.26	0.07	.22
Moral disengagement (global)	0.93	0.66	0.81	0.60	1.05	0.69	-6.32 ^a	<.001	0.31	0.16	.38

Note. a = Welch's *t*

N = 1107

* *p* < .05

** *p* < .01

*** *p* < .001

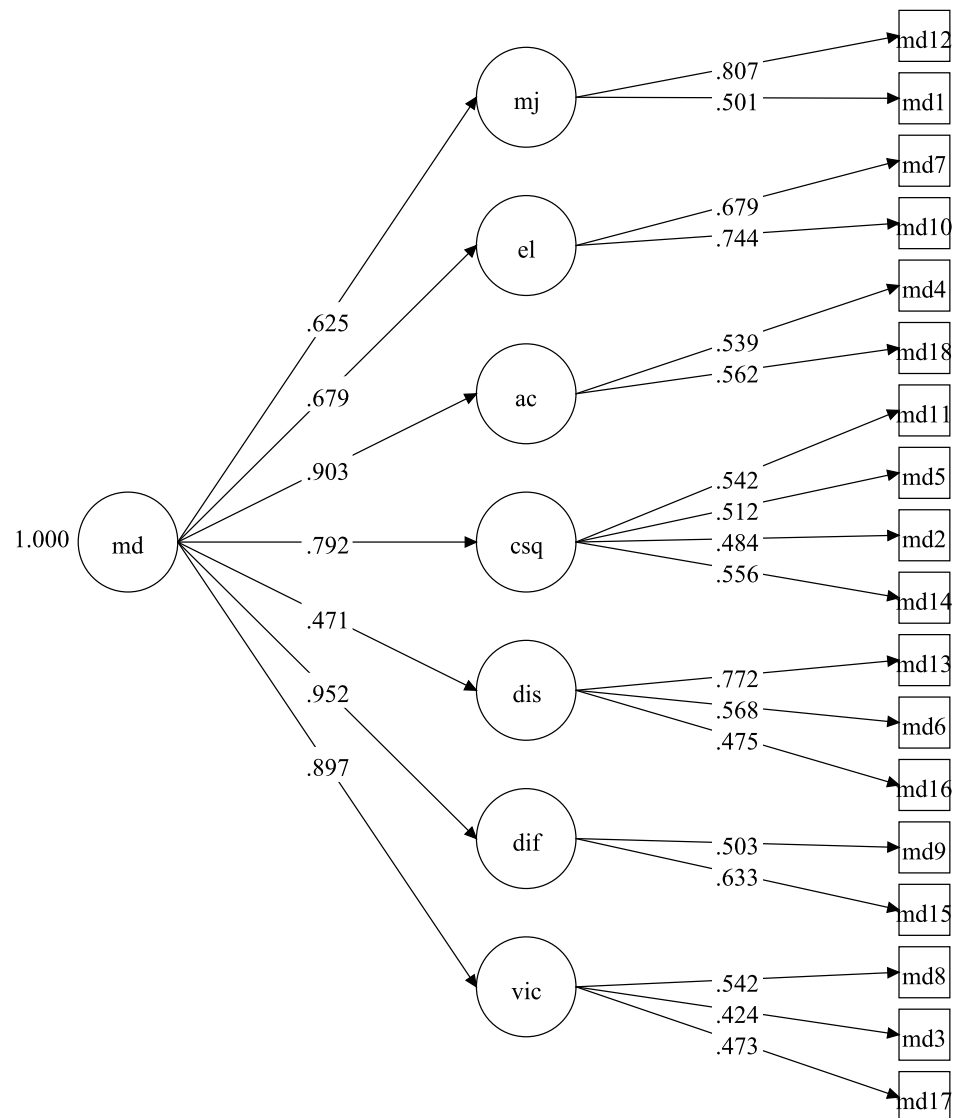
Table B.2. Inter-correlations between all variables

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Age	—														
2. Gender (girls = 0, boys = 1)	.01	—													
3. Self-reported bullying	-.04	.15***	—												
4. Peer-reported bullying	-.09**	.25***	.33***	—											
5. Self-reported victimization	-.17***	.02	.44***	.31***	—										
6. Peer-reported victimization	-.18***	.02	.15***	.35***	.41***	—									
7. Defender	-.01	-.17***	-.10***	-.03	.15***	.05	—								
8. Outsider	.01	.14***	.24***	.06*	.10**	.08**	-.30***	—							
9. Moral justification	-.02	.16***	.25***	.14***	.12***	-.00	-.06	.10**	—						
10. Euphemistic labeling	.02	.13***	.27***	.08**	.06*	-.08**	-.15***	.18***	.29***	—					
11. Advantageous comparison	-.07*	.16***	.21***	.10**	.09**	-.00	-.11***	.13***	.29***	.37***	—				
12. Displacement of responsibility	-.08**	.05	.16***	.06*	.08*	-.02	.01	.04	.25***	.20***	.23***	—			
13. Diffusion of responsibility	-.10**	.16***	.29***	.14***	.10**	.05	-.16***	.22***	.32***	.33***	.37***	.27***	—		
14. Distorting consequences	.00	.16***	.25***	.11***	.07*	.01	-.17***	.20***	.29***	.36***	.65***	.23***	.37***	—	
15. Victim attribution	-.09**	.11***	.30***	.15***	.18***	.08**	-.17***	.19***	.26***	.29***	.32***	.20***	.20***	.47***	—
16. Moral disengagement (global)	-.09**	.19***	.39***	.17***	.16***	.02	-.17***	.23***	.59***	.61***	.65***	.59***	.59***	.67***	.63***

N = 1107* *p* < .05** *p* < .01

*** $p < .001$

Appendix C.

Figure C.1.*Standardized Factor Loadings of Model 5.*

md = moral disengagement; mj = moral justification; el = euphemistic labeling; ac = advantageous comparison; csq = distorting the consequences; dis = displacement of responsibility; dif = diffusion of responsibility, vic = victim attribution