TITLE:
The Role of Traditional and Online Moral Disengagement on Cyberbullying: Do Externalising Problems Make Any Difference?

Marinella Paciello¹*, Carlo Tramontano², Annalaura Nocentini³, Roberta Fida⁴, Ersilia Menesini³

¹ Uninettuno Telematic International University, Rome, Italy
² Centre for Global Learning: Education and Attainment, Coventry University, Coventry, United Kingdom
³ Department of Educational Sciences and Psychology, University of Florence, Firenze, Italy;
⁴ Norwich Business School, University of East Anglia, Norwich, United Kingdom

* Corresponding Author:
Marinella Paciello
m.paciello@uninettunouniversity.net
Uninettuno Telematic International University
Corso Vittorio Emanuele II, 39
00186, Rome
Italy
TITLE:

The Role of Traditional and Online Moral Disengagement on Cyberbullying: Do Externalising Problems Make Any Difference?
Abstract

This study examines the differential role of traditional and online moral disengagement (MD) in relation to cyberbullying. Traditional MD is operationalised as a process operating across contexts, whereas online MD as a contextualised process related to online settings. We hypothesised that they are separate, although correlated, and have different roles depending on externalising tendencies. The sample comprised 856 high school students (mean age = 14.7, S.D. = 1.7; 45.6% females). Regression analyses highlighted that: a) for low externalising adolescents, only online MD was significantly related to cyberbullying; b) for medium externalising adolescents, both online and traditional MD are significant, with the former more strongly associated with cyberbullying; c) for high externalising adolescents, traditional MD is key. Cluster analyses identified five configurations: 1) the Externalising Traditionally Disengaged; 2) the Externalising Not-Disengaged; 3) the Online Disengaged; 4) the All Good; and 5) the Unsuspected. The Online Disengaged has the highest engagement in cyberbullying. The Unsuspected (showing the same low externalising behaviour but significantly higher level of online MD than the All Good) engage in cyberbullying as much as Externalising Traditionally Disengaged and Not-Disengaged.

These findings have implications for intervention programmes, underlining the relevance of considering the moral processes within the online environment.

Keywords: moral disengagement, cyberbullying, externalising problems, ICT, online
Introduction

Albert Bandura recently stated that ‘the advent of the Internet ushered in a ubiquitous vehicle for disengaging moral self-sanctions from transgressive conduct. The Internet was designed as a highly decentralized system that defies regulation. Anybody can get into the act, and nobody is in charge’ (2016, p.68). Rather than demonising the web, Bandura claims that we need to better understand the social-cognitive processes that deactivate moral agency, namely Moral Disengagement (MD), during online interactions. To this end, within Bandura’s theoretical framework (1991, 2016), this study examined the role of MD (traditional and online) in relation to one of the most discussed ‘millennials’ online deviant behaviours: cyberbullying. Indeed, data from surveys conducted in industrialised countries suggest a prevalence rate for cyberbullying ranging from 5% to 21% (UNESCO, 2017). Worriingly, cyberbulling is increasing also in Europe, with the exposure to cyberbullying increasing from 7 to 12% between 2010 and 2014 among adolescents (UNESCO, 2018). Moreover, recent meta-analyses reported prevalence rates for cyberbullying victimisation among school aged children as ranging from 10 to 40% across studies (Kowalski, Giumetti, Schroeder, & Lattanner, 2014; Kowalski, Limber, & McCord, 2019).

Cyberbullying has been identified as a serious public health problem, which can dramatically impact the lives of adolescents (Abouhaoude, Savage, Starcevic, & Salame, 2015). Cyberbullying may lead to undesirable behavioural and health-related outcomes, including having a negative impact on psychological well-being; cyberbullying involvement has been found to be related to depression, anxiety, stress, emotional problems, low self-esteem, and suicidal thoughts (Kowalski et al., 2019). Furthermore, adolescents who have been cybervictimised also report poor physical health (Sourander, 2010).
Given the rapid rise in accessing the internet and other Information and Communications Technologies, and the impact cyberbullying may have on young people’s life, the understanding of moral regulation processes in relation to cyberbullying in adolescents is paramount. As such, this study investigated, for the first time, the concurrent contribution of off and online MD.

MD refers to a set of cognitive manoeuvres aimed at selectively ‘switching off’ individuals’ moral agency, allowing the engagement in misconduct that they would generally avoid. MD has been recognised as one of the most important disinhibitory factors able to explain a range of deviant and aggressive behaviours in a variety of contexts (e.g. Bandura et al., 2001; Gini Pozzoli, & Hymel, 2014), including traditional bullying in school (Kowalski et al., 2014). When considering cyberbullying - defined as aggressive behaviour perpetrated via information and communication technologies (Smith et al., 2008) - the association with MD is still unclear and results are inconclusive. Some studies provide evidence for this association (e.g., Kowalski et al., 2014), underlining that the characteristics of the online environment may promote the recourse to MD mechanisms (Bauman, 2010). In particular, the cyberbully may not be aware of the consequences of their actions, and may misinterpret victims’ perception of hurtful messages (Dehue, Bolman & Vollink, 2008). Others studies have claimed that the contribution of MD is strong in relation to traditional bullying, but only weak for cyberbullying, because in the online setting, the perpetrators may not need to rely on justification processes as much as in ‘real life’ (e.g., Perren & Gutzwiller-Helfenfinger, 2012; Pornari & Wood, 2010).

As suggested by Bussey and colleagues (Bussey, Fitzpatrick, & Raman, 2014), one of the reasons for the inconclusive findings on the link between cyberbullying and MD, may be related to the way the latter has been measured. Indeed, the literature seems to suggest that, rather than adopting a ‘one-size-fits-all’ approach, it is important to operationalise MD specifically referring to the context and the behaviour under study (Bandura, 2016). For
instance, in the last decade, various scholars have developed and adopted domain-specific measures such as: civil (Caprara et al., 2009), work (Fida et al., 2015; Moore et al., 2012), underage drinking (Quinn & Bussay, 2015), and sport (Boardley & Kavussanu, 2007) MD. When investigating cyberbullying, however, the majority of researchers have generally used the traditional measure (Bandura, Barbaranelli, Caprara, & Pastorelli, 1996), which operationalises MD in relation to a broad range of offline deviant and aggressive behaviours. Only a few studies have used ad hoc measures (i.e., cyberbullying scenarios: Sticca et al., 2013; Cyber Bullying MD Scale: Bussey & Fitzpatrick, 2014; adapted MD Scale: Hymel, Rocke-Henderson, & Bonanno, 2005; DeSmet et al., 2016), and to the best of our knowledge, none have concurrently tested the contribution of both online and traditional MD. We believe that these represent two separate, although correlated, set of mechanisms and are likely to operate differently, particularly in relation to cyber misconduct (Bandura et al., 2006; Pelton, Gound, Forehand, & Brody, 2004). While traditional MD may represent a crossecutting process operating across different contexts to legitimise a variety of deviant behaviours, online MD may represent a contextualised process specifically related to online social setting and misbehaviours.

In line with a social-ecological framework, to understand bullying it is important to consider the individual within their context (Cross et al., 2015; National Academies of Science, Engineering, and Medicine, 2016). Hence, online MD is expected to be particularly informative in relation to cyber deviance, since it is a domain-specific construct referring to processes weakening moral control in technologically mediated contexts. These contexts represent ‘places’ where individuals navigate relationships and social situations without clear interpersonal borders and behavioural codes. In this regard, a number of authors (e.g., Suler, 2004) have suggested that there are some factors in online settings, such as invisibility, asynchronicity, and minimisation of authority that may foster a disinhibitory effect. This is in
line with the results from the experimental study conducted by Naquin and colleagues (Naquin, Kurtzberg, & Belkin, 2010), showing that people are more likely to feel released from strict moral principles in an online environment. In particular, this may be due to the greater psychological distance between the actor and the potential victim, and between deviant actions and their harmful consequences in online settings. More recently, Runions and Bak (2015) claimed that some technological affordances may facilitate the activation of MD mechanisms aimed at silencing the individual moral control. This is supported by studies conducted on different online deviant behaviours, such as online racism (D’Errico & Paciello, 2018; Faulkner & Bliuc, 2016), and software piracy (Garbharran & Thatcher, 2011; Lowry, Zhang & Wu, 2017).

So far, the relevance of the specific online context has been considered. However, following Bandura’s reciprocal triadic model (1997), it is pivotal to take into account how the environment, the behaviour and the individual interact. Hence, to gain a better understanding about how MD may work in the online context, we hypothesised that individual differences in externalising conduct may play a central role. In particular, we theorised that online and traditional MD measures are differently informative, depending on individuals’ tendencies to engage in aggressive and deviant conduct. The interconnections between externalising problems, MD, and bullying are well-established in the literature. For example, longitudinal studies have shown that a stable engagement to aggressive and deviant behaviours tends to be associated to a chronic use of MD (Paciello, Fida, Tramontano, Lupinetti & Caprara, 2008). In addition, the literature also suggests that traditional bullying is generally associated with externalising tendencies (Gradinger, Strohmeier, & Spiel, 2009), and similar evidence has been provided in relation to cyberbullies, who often also show a problematic behavioural
profile ‘outside the web’ (Juvonen & Gross, 2008; Kowalski et al., 2008; Raskauskas & Stoltz, 2007; Sticca et al., 2013; Ybarra & Mitchell, 2004a, 2004b).

The study of these aggressive-related variables, however, should not be limited to problematic adolescents. In line with the individual by context perspective (Cross et al., 2015) and social cognitive theory (Bandura, 1991; 2016), we believe that it is important to understand whether and how in the online environment, adolescents who do not typically display externalising problems might be at risk of engaging in cyberbullying. To this end, in the present study we adopted a twofold analytical approach. Firstly, by using a variable-oriented approach, we aimed to examine the specific contribution of online and traditional MD in understanding cyberbullying, considering the moderating role of externalising problems. This goal is particularly helpful to investigate whether MD processes operating in offline and online environments play different roles for adolescents with different levels of problematic behaviours. Specifically, we hypothesised that for adolescents with high levels of externalising problems, online MD may represent an ‘extension’ of offline MD, and as such would not provide any specific additional contribution to explain the engagement in cyberbullying. In contrast, for those adolescents who are not usually aggressive in face-to-face interactions, we expected that online MD could provide an additive contribution, above and beyond offline MD.

Secondly, by using a person-oriented approach, we aimed to investigate, in greater depth, the potential interplay of MD processes in offline and online environments, and externalising problems. We did not have precise expectations about the number of configurations, but we had some hypotheses in terms of the profile of clusters. In particular, we were expecting to most likely identify a cluster characterised by low levels in both online and traditional MD, as well as in externalising problems. This would represent the non-
problematic group of adolescents, not involved in cyberbullying. Furthermore, we considered that adolescents with high externalising problems might also have associated high levels of online and traditional MD, and might most likely be expected to also frequently engage in cyberbullying. However, we also considered that cyberbullying may not necessarily be a matter of externalising tendencies; specifically, for those adolescents with lower externalising problems, traditional MD could be equally low, but online MD could represent a set of specifically internet-related mechanisms activated during online social interactions. In other words, in online environments (Cross et al., 2015), the activation of specific self-exonerative processes might become easier or more accessible also to 'unsuspected' cyberbullies. Cyberbullying may represent for some adolescents a situational behaviour, related to ‘contextualised self-exonerative’ processes, while for others it may just be another expression of the same problematic functioning, related to individuals’ externalising tendencies. This has the potential to be particularly informative for designing interventions to prevent cyberbullying in adolescents.

**Method**

*Participants and procedure*

The present study is part of a research project on the problematic use of technologies during adolescence conducted on 856 students (45.6% females) in two junior (27.8%) and two senior (72.1%) high schools (with .1% of the participants not reporting this information in the questionnaire). The mean age of the participants was 14.7 years (*S.D.* = 1.7). The project was approved by the Ethics Committee of the University to which the first author is affiliated. It was then approved by the Head Teachers and the Board of Teachers in each school. All the teachers in each school received an information sheet detailing the research procedure, and the measurements included in the questionnaire for the children. All parents
received an information sheet, detailing information about privacy and confidentiality, and were required to provide informed consent for their children to take part in the project. Following this, administration sessions were scheduled with the schools according to their calendar. Between April and June 2017 children (divided by classes) were invited in the IT room of the schools to complete an online questionnaire. The administration process was supervised by a researcher and a teacher (trained by the research team) and generally took one-hour per class to complete. The preliminary findings were presented to teachers in a seminar and each school was given a report summarising the results.

Measures

*Online Moral Disengagement (OMD)* was assessed by eight items specifically created for the present study. The scale aimed to translate MD into the online setting. In developing the items, we referred to the *loci* (i.e., behaviour, agent, outcome, and victim) and the mechanisms theorised by Bandura (1991, 2016). In particular, at the behaviour locus, where the online action is ‘re-structured’, three items, each one corresponding to a specific mechanism (*euphemistic labelling, advantageous comparison, moral justification*), were generated. At agent locus, where the responsibility for an online action can be moved from oneself to someone else, two items were generated (*displacement of responsibility; diffusion of responsibility*). At outcome locus, where the consequences of the online action are distorted or disregarded, one item was generated (*disregarding/distorting consequences*). Finally, at victim locus, where victims themselves end up being considered responsible or are deprived of their human characteristics, two items were generated (*attribution of blame; de-humanisation*). Participants were required to rate their level of agreement with each statement on a scale ranging from 1 = strongly disagree to 5 = strongly agree. Factorial structure and reliability for this scale are presented in the result section.
Moral disengagement (MD) was assessed by the 14-item scale developed by Pozzoli and colleagues (Pozzoli, Gini, & Vieno, 2012). This measure comprised items referring to the eight mechanisms and the corresponding four loci: behaviour (cognitive restructuring, six items); agent (minimising one’s agentive role, three items), outcome (disregarding/distorting consequences, two items), victim (blaming/dehumanising the victim, three items). Participants were required to rate their level of agreement with each statement on a scale ranging from 1 = strongly disagree to 5 = strongly agree. It is important to note that there were no missing data. Preliminary Confirmatory Factor Analysis was implemented in Mplus using the WLSMV (Weighted Least Squares with Mean and Variance adjustment) estimator, resulting in a good fit: Chi Square = 235.796, df = 71, p < .001; RMSEA = .052 (95% CI = .045 - .060), p = .310; CFI = .930, WRMR = 1.025. The reliability was tested by computing the Composite Reliability (Raykov, 1997) that resulted equal to .88.

Cyber-bullying Perpetration was assessed by 13 items developed by Palladino, Nocentini and Menesini (2015). Participants were asked to report how frequently they engaged in a range of online bullying behaviours using a 5-point response scale (ranging from 1 = never to 5 = several times per week). Cyberbullying perpetration was defined as a second order factor, with four first-order factors, namely: written verbal bullying (three items); visual bullying (three items); impersonation (four items); and exclusion (three items). Again, there were no missing data. Preliminary Confirmatory Factor Analysis was implemented in Mplus using the WLSMV, resulting in a good fit: Chi Square = 151.631, df = 72, p < .001; RMSEA = .036 (95% CI = .028 - .044), p = .998; CFI = .988, WRMR = .780. The reliability was tested by computing the Composite Reliability (Raykov, 1997) that resulted equal to .96. Given the highly asymmetrical distribution of the mean score, a logarithmic transformation has been used for the analyses.
Externalising behaviour was assessed by 7 items included in the measure developed by Achenbach and colleagues (Achenbach et al., 2011). This scale assesses behaviour towards others in an external environment, characterised by limited or scarce compliance with requests from the social context. Participants were asked to rate the frequency of their engagement in these behaviours using a 3-step response scale (ranging from 0 = not true to 2 = very true). There were no missing data. Preliminary Confirmatory Factor Analysis was implemented in Mplus using the WLSMV, resulting in a satisfactory fit: Chi Square = 58.424, df = 12, p < .001; RMSEA = .067 (95% CI = .051 - .085), p = .045; CFI = .957, WRMR = 1.070. The reliability was tested by computing the Composite Reliability (Raykov, 1997) that resulted equal to .75.

Planned analyses

First, a Confirmatory Factor Analysis (CFA) was implemented in Mplus 7.2 on the newly developed OMD scale to test the adequacy of the predicted one-factor model. Following this, an additional CFA was implemented including both the OMD and the MD measures to test whether they were clearly differentiated in terms of factorial structure, and could be considered as two different constructs.

In order to test our main hypotheses, we performed a three steps multivariate regression model in IBM-SPSS 21. At step 1, gender and age were entered. At step 2, OMD, MD and externalising problems were entered to test their direct contribution. Finally, at step 3, in order to test whether MD operating in offline and online environments play different roles for adolescents with different levels of problematic behaviours, we entered the interactions between externalising behaviours and MD. Only significant interactions were included in the final model. All the variables were centred in order to avoid multicollinearity. Significant
interactions were further investigated through simple slope analysis to test the moderation effect (Robinson, Tomekm & Schumacker, 2013).

Finally, following Asendorpf, Borkenau, Ostendorf, and Van Aken (2001), a two-phased cluster analysis was implemented with SLEIPNER v. 2.1 (Bergman, Magnusson, & El-Khoury, 2003). The number of clusters to extract was established, taking into account the interpretability of the solution, the increase in error sum of squares, the Explained Sum of Squares (ESS), and the homogeneity cluster coefficients. An explained ESS approaching 2/3 of the total (i.e., about 66%) and homogeneity cluster coefficients lower than 1 are considered, suggesting a good final non-hierarchical classification. Gender and school-grade differences were tested using $\chi^2$ and, when significant, examining adjusted standardised residuals. Differences in cyberbullying perpetration among identified clusters were analysed by paired-comparison test, using the Bonferroni correction. Cluster membership was entered as independent variable, cyberbullying perpetration as dependent variable, and gender and school grade (junior vs senior high school) as covariates.

**Results**

Descriptive statistics for the OMD items are presented in Table 1. Skewness and kurtosis for all the OMD items were greater than the $|1|$ cut-off, hence the CFA was implemented in Mplus using the WLSMV estimator. Results confirmed the excellent fit of the one factor model to the data: Chi Square = 24.429, $df = 20$, $p = .224$; RMSEA = .016 (95% CI = .000 -.035), $p = .999$; CFI = .996, WRMR = .514. The reliability of the OMD scale was tested by computing the Composite Reliability (Raykov, 1997) that resulted equal to .81.

[Table 1]

An additional CFA including both the OMD and the MD scales was then implemented to test whether the two measures are differentiated in terms of factorial structure. The results of the CFA model, implemented in Mplus using the WLSMV estimator, suggested an adequate
fit of the model to the data: Chi Square = 475.101, df. = 204, p=.000; RMSEA = .039 (95% CI = .035 - .044), p = 1.000; CFI = .938, WRMR = 1.058. The correlation between OMD and the second order MD factors was .88.

Modification indices (MI) suggested cross-loadings for item on dehumanisation in the OMD scale on the blaming/dehumanising factor of MD (MI = 41.555), as well as in the cognitive restructuring factor (MI = 28.451) and in the second order MD factor (MI = 33.225). Given the interest in maximising the differentiation of the two constructs, OMD and MD, the CFA model was replicated after excluding that item. The resulting fit was satisfactory: Chi Square = 378.526, df = 184, p = .000; RMSEA = .035 (95% CI = .030 - .040), p = 1.000; CFI = .95, WRMR =.970. The correlation between OMD and the second order MD factors was .84. The item on dehumanisation of the OMD scale was therefore excluded from the following analyses.

Table 2 summarises the findings from the regression analyses. In the final model MD, OMD, externalising behaviours and the interaction between MD and externalising behaviours showed a significant association in relation to cyberbullying, controlling for age and gender. In particular, higher levels of both traditional and online MD were associated with higher levels of cyberbullying, confirming the additive role of these two forms of moral processes. In addition, higher levels of externalising behaviours were associated with a higher likelihood of engaging in cyberbullying.

In order to interpret the significant interaction term, we performed a simple slope analysis through three separate regressions on three groups defined on the basis of participants’ level of externalising behaviours: low externalising (below the 20 percentile), medium externalising (between 20 and 80 percentile), and high externalising (above 80 percentile). Table 3 shows the simple slope estimates, and Figure 1 shows the slopes plotted
in the three groups separately. In the low externalising group, only OMD was significantly associated with cyberbullying. In the medium externalising group, both MD and OMD were significantly associated with cyberbullying, with regression coefficients for OMD double those of the MD coefficient. Finally, in the high externalising group, MD was significantly associated with cyberbullying, while OMD was not significant.

[Table 3]

[Figure 1]

Cluster analysis was implemented in Sleipner 2.1, integrating hierarchical and non-hierarchical methods to optimise the quality of the final solution. Initially, the RESIDUE module was run to identify and exclude outliers, which would affect the reliability of the cluster solution. As a result, six participants were excluded from the following analyses. The CLUSTER module was then implemented using a hierarchical procedure (Ward’s method), minimising the variance within clusters. The increase in error score was 3.40564 (moving from a 10- to a 9-cluster solution), 3.46613 (from 9- to 8-cluster), 3.54148 (from 8- to 7-cluster), 6.61548 (from 7- to 6-cluster), 7.47870 (from 6- to 5-cluster), and 10.35895 (from 5- to 4-cluster solution). This suggested a 5-cluster solution, initially explaining 61.27% of variance, with the following fit indices: (1) point-biserial correlation, PBC=.339; (2) Gamma index=.662; (3) C-Index=.0729; (4) G (+) index=.709; and (5) W/B index=.2609. Homogeneity coefficient ranged from .106 to .404.

The module RELOCATE was finally run to reclassify participants of the 5-cluster solution, using a non-hierarchical methods (i.e., k-means). The integration of hierarchical and non-hierarchical methods aimed at improving the homogeneity of each cluster and, in turn, increased the explained variance of the solution. In this case, the final explained variance was 67.59%, which was considered satisfactory. Cluster internal validity was examined through a
set of Univariate ANOVAs. In particular, the five clusters significantly differ in online MD 
\[ F(4, 845) = 435.455, p < .001 \], traditional MD \[ F(4, 845) = 600.749, p < .001 \], and 
externalising behaviour MD \[ F(4, 845) = 243.950, p < .001 \]. Significant gender (\( \chi^2 = 51.567, 
df = 4, p < .001 \)), age \[ F(4, 845) = 6.466, p < .001 \], and school grade differences (\( \chi^2 = 
36.238, df = 4, p < .001 \)) were tested and detected.

The five clusters are depicted in Figure 2 and described further below:

[Figure 2]

Cluster 1, *Externalising Traditionally Disengaged*, comprises 121 participants (14.2% 
of the total sample; 30.6% females; \( \text{Age}_{\text{mean}} = 14.8, \text{s.d.} = 1.6 \)) characterised by average level 
of online MD, medium-high traditional MD and externalising tendencies. **Overall, members** 
of this cluster are mainly males (as supported by \( \chi^2 \) analyses) and stand out for their 
level of traditional MD and externalising tendencies – dimensions usually associated in 
the literature with aggressive phenomena.

Cluster 2, *Externalising Not Disengaged*, comprises 172 participants (20.2% of the 
total sample; 56.4% females; \( \text{Age}_{\text{mean}} = 15.0, \text{s.d.} = 1.4 \)) characterised by low online and 
traditional MD, but high externalising behaviour. **This was an unanticipated configuration,** 
that stand out for including mainly older (in senior high school) females, presenting 
externalising behaviour not associated to any form of moral disengagement.

Cluster 3, *Online Disengaged*, comprises 63 participants (7.4% of the total sample; 
28.6% females; \( \text{Age}_{\text{mean}} = 13.8, \text{s.d.} = 1.8 \)), characterised by extremely high online MD, high 
traditional MD, and medium-high externalising behaviour. **Overall this cluster, mainly** 
including younger participants (in junior high school) stands out for the level of MD, 
especially online MD (highest level across clusters).
Cluster 4, *All Good*, comprises 299 participants (35.2% of the total sample; 57.2% females; \( \text{Age}_{\text{mean}} = 14.7, \text{s.d.} = 1.7 \)) characterised by extremely low levels in all the three variable. **Overall this cluster, mainly including females, is the expected non problematic configuration.**

Cluster 5, *Unsuspected*, comprises 195 participants (22.9% of the total sample; 34.9% females; \( \text{Age}_{\text{mean}} = 14.6, \text{s.d.} = 1.8 \)), characterised by the same extremely low level of externalising behaviour as Cluster 4, along with average traditional MD and medium-high online MD. **This cluster, mainly including males, despite not presenting externalising tendencies is nevertheless characterised by the relevant level of online MD.**

Moreover, the five clusters show different levels of cyberbullying \( [F (4, 848) = 8.14, p < .001] \). Paired-comparison test attested that the Cluster 3 *Online Disengaged* shows the highest cyberbullying perpetration level. Cluster 4 *All Good* shows the lowest level. Cluster 1 *Externalising Traditionally Disengaged*, Cluster 2 *Externalising Not Disengaged* and Cluster 5 *Unsuspected* are characterised by the same level of intermediate cyberbullying perpetration level.

**Discussion**

This was the first study to concurrently examine the relationship of traditional and online MD with engagement in cyberbullying. The results demonstrate that online MD, focused on aggressive and deviant behaviour in online settings, is a separate, albeit correlated, construct compared to traditional MD, focused on behaviours in off-line settings. Moreover, our findings provide evidence that online MD is a key dimension to consider when examining engagement in hurtful and damaging behaviour in online contexts, such as engaging in cyberbullying. In addition, our findings support the need for considering
concurrently online and traditional MD in relation to adolescents’ engagement in cyberbullying, since they can be differently informative depending on individual differences in externalising behaviours. Indeed, although the traditional MD was associated with cyberbullying, online MD provided a stronger contribution to the understanding of this damaging conduct. Hence, contrary to a number of previous studies suggesting that the impact of MD on cyberbullying is limited (Pornari & Wood, 2010) or not significant (Perren & Gutzwiller-Helfenfinger, 2012), our results suggest that adolescents may need to morally disengage to some extent and in doing this they may use clues from the specific online setting. Hence, to fully appreciate the role of MD in online environments, it is pivotal to consider its peculiar ‘expressions’ in relation to problematic behaviour within that specific environment. This is particularly relevant given that, as Bandura (2016) stated, in the name of their freedom of speech and expression, individuals may engage in aggressive online behaviour whose consequences are far away in time and space, and as a result they may end up inflicting wounds to the self-esteem of individuals whose pain they cannot see. Overall, this is consistent with the social-cognitive approach that supports the need to operationalise the relevant theoretical constructs, such as MD, within specific domains of functioning. Indeed, as claimed by Bandura (1997, 2016), the closer the cognitions are to the actions, the stronger the explicative power of the measure.

Results from regression analyses attest that OMD and traditional MD are differently relevant depending on adolescents’ tendencies to engage in aggressive and harmful conduct. Specifically, traditional MD is more relevant for individuals with high externalising problems. For these adolescents, traditional MD significantly explains cyberbullying over and above online MD. This indicates that, for this group, online MD is an ‘extension of offline MD’. In other words, for individuals with high externalising problems offline, MD it is likely
to be sufficient to explain engagement in cyberbullying, since this misconduct might just be an additional expression, within the online context, of their aggressive behaviour and does not necessitate specific justification mechanisms. On the other hand, online MD is particularly informative of cyberbullying for those adolescents who do not have, or are not at risk of, more general externalising problems. These adolescents may be more influenced by characteristic features and conditions of the online environment to morally disengage and behave in ways that they would avoid in face-to-face interactions.

Results from the cluster analyses support findings from the regressions, while also capturing a more articulated and comprehensive picture. First, consistent with the results from regression and the simple slope analyses, for adolescents characterised by high levels of externalising problems, the online MD is not relevant, while the traditional MD seems to play a key role. However, the cluster analyses provided a more nuanced description. Indeed, further to a configuration in which externalising behaviour is, as expected, associated with traditional MD (Cluster 1 *Externalising Traditionally Disengaged*, 14.2% participants, mainly male), there is another one (Cluster 2 *Externalising Not Disengaged*, 20.2%, mainly female), characterised by levels below average for both online and traditional MD. Both these clusters are equally involved in cyberbullying, but they are not the one showing the highest level of perpetration. Also for participants with intermediate level of externalising behaviour (Cluster 3 *Online Disengaged*, 7.4%, younger), findings from regression and simple slope analyses are confirmed. Indeed, this cluster has a high level of both online and traditional MD, however the former has the highest level across all the clusters. Consistent with our reasoning, this is the cluster with the highest engagement in cyberbullying perpetration. Finally, when considering participants with low levels of externalising behaviour, results from cluster analysis provide support to findings from regression, although suggesting a more
complex picture. Along with a conspicuous cluster of adolescents with low MD, both online and traditional (Cluster 4 *All Good*, 35.2%), it is also possible to identify an additional cluster (Cluster 5 *Unsuspected*, 22.9%) with equally low levels of externalising but with intermediate level of online MD. Worryingly, this latter cluster presents the same engagement in cyberbullying as Cluster 1 *Externalising Traditionally Disengaged* and 2 *Externalising Not Disengaged*.

Overall, focusing the attention only on externalising behaviour to understand cyberbullying perpetration might indeed be misleading. The clusters with the highest levels of externalising behaviours (Cluster 1 and 2), potentially calling a greater attention from teachers and professionals, engage in cyberbullying but significantly less than the cluster with lower level of externalising tendencies (Cluster 3). Possibly more importantly, considering externalising behaviour as a key ‘flagger’ of risk of perpetration of cyberbullying would result in the exclusion of a considerably large group of adolescents (Cluster 4, about one in four) who despite not showing problematic behaviour in face to face interaction, engage in cyberbullying as much as their highly externalising classmates.

These findings are consistent with well-established theoretical models that underline the importance of considering the social context to understand individual behaviour, such as social-ecological framework (Cross et al., 2015), social cognitive theory (Bandura, 1991; 2016), and contextual ecological user experience framework (Shin, 2017; Shin & Biocca, 2017). These models highlight the role of social-cognitive processes operating at an online level to understand cyberbullying. Although the literature on these aspects is still limited, recent studies (Kowalski et al., 2014; Runions & Bak, 2015) have suggested the need to explore technological dimensions and online influences that could provide opportunities for activating disinhibitory processes to engage in deviant online behaviour. Indeed, the role
played by online MD for ‘good guys’ points the attention towards the technological affordances that could increase MD in technologically-mediated contexts. Runions and Bak (2015) claimed that the activation of MD mechanisms may be facilitated by online characteristics, such as the paucity of social emotional cues, the ease of disseminating communication via social networks, the increase of media attention to online aggressive forms and the lack of space-time constraints delimiting aggressive manifestations. In addition, there is an extensive literature underlining the insidious nature of cyberbullying (e.g. Robson & Witenberg, 2013) that could be facilitated by its occurrence with no space and time barriers (Menesini et al., 2012; Patchin & Hinduja, 2006). Kowalski and colleagues (2014) claimed that features of the online environment, and anonymity in particular, ‘significantly opens up the pool of potential perpetrators of cyberbullying, compared to traditional bullying’ (p. 1107). More worryingly, the phenomenon might be further exasperated when considering the availability of app and technologies allowing people to preserve their anonymity during their online interactions (e.g., Honesty app such as Sarahah). Moreover, cyberbullies may not be aware of the consequences of their actions, and/or may completely misinterpret or disregard victims’ perception of hurtful messages (Dehue et al., 2008). In sum, research suggests that the characteristics of online environments facilitate the recourse to MD mechanisms for the majority of adolescents (Bauman, 2010), which in turn facilitates engagement in cyberbullying (Cross et al., 2015). Online MD may capture adolescents’ ‘permeability’ to the activation of individual disinhibitory mechanisms facilitating engagement in aggressive behaviour in an online context, which is already characterised by affordances that may jeopardise the exercise of moral control. To this regard, recent studies have highlighted the necessity to consider an individual by context approach to explain bullying and cyberbullying (Authors, submitted; Cross et al. 2015). By integrating variable- and person-oriented approaches, this study represents a step ahead in
this direction. Overall, findings show how contextual variables, by disinhibiting moral context specific constraints and limitations, are particularly relevant for adolescents with no (cluster 5) or limited externalising problems (cluster 3). On the contrary, personal characteristics can better explain the mediating role of traditional MD and in turn the use of cyberbullying behaviour for the group with high externalising problems (cluster 1).

Notwithstanding the novelty of the results of the presented study, the findings need to be understood within the methodological contexts of the study. Firstly, the study comprised self-reported measures, which are subject to social desirability biases. In this study we have been unable to understand causal links between (O)MD, externalising behaviours and cyberbullying; we therefore recommend future research using experimental and longitudinal designs to investigate this. Cyberbullying assessment might indeed be biased by under-reporting, due to shame, fear, or more generally self-protection. A real-world, longitudinal study would allow the researchers to overcome these limitations and to capture a more precise picture of the phenomenon, including the role of gender and age differences. In particular, such an approach could better capture whether OMD and MD are distinct across time and/or whether one predicts the other. Finally, other variables, for instance emotional activation and regulation, can play a role in the deactivation of moral processes, and they should be examined in future studies. In particular an increasing emotional activation may impede cognitive control: individuals would then act impulsively and without considering the potential consequences. In the online setting, the expression and diffusion of emotions is very likely, as suggested by the literature on online emotional contagion (Kramer, Guillory & Hancock, 2014). This can be amplified by the conformism, that in adolescence can be particularly common given the importance of peers during this particular stage of life, and by the difficulties that adolescents generally experience in managing their impulses. This mix of
elements can represent a fertile ground where online MD can grow. Future studies should specifically explore other potential leverages, in addition to online and traditional MD, to be included in programmes promoting cyberbullying prevention.

Nevertheless, the findings of this study have important implications for interventions directed at preventing cyberbullying. Although researchers have clearly indicated that such interventions should focus on reducing MD and enhancing personal responsibility, very few efforts have been made to define specific components of such programmes and to test their efficacy (e.g., Hymel & Bonanno, 2014). Some anti-bullying interventions have addressed ‘self-serving cognitive distortions’, ‘thinking errors’ and perceptions of peer victimisation (Hymel & Bonanno, 2014), and moral reasoning (Schultze-Krumbholz, Zagorscak, & Scheithauer, 2018). Other studies suggest that the focus should be moral identity processes (Wang, Lei, Liu, & Hu, 2016), and on moral rules and standards (Caravita & Gini, 2010). The current study provides further support for the relevance of including a specific component focused on moral engagement and moral self-justifying mechanisms to prevent cyberbullying. In particular, a general intervention on moral processes and on moral understanding seems to be particularly appropriate for adolescents who are high in externalising behaviours. More importantly, an intervention on moral processes specifically targeting the online environment needs to be promoted for the ‘good fella’. Indeed, for the majority of adolescents, cyberbullying prevention programmes should address context-specific processes. To this end, they should work on the affordances of the online setting (e.g., anonymity, spatial and temporal deferral of consequences) that may facilitate the activation of online MD.
References


Authors (submitted). Anonymised for review


Acknowledgments

Authors would like to thank Miss Rebecca Crookes for her support in proof-reading the manuscript.
Figure 1. Simple slope analysis - Plotted slopes of the interaction between MD and Externalizing behaviours, controlling for OMD.
Figure 2. Final 5-cluster solution and associated engagement in cyberbullying perpetration.

![Cluster Analysis](image)

**Notes.** md_on=Online Moral Disengagement; md_t=Traditional Moral Disengagement; ext=Externalising behaviour; cl_1=Externalising Traditionally Disengaged; cl_2=Externalising Not-Disengaged; cl_3=Online Disengaged; cl_4=All Good; cl_5=Unsuspected, N = cluster size, M_{age}=mean age; SD_{age}=standard deviation for age. Plotted cluster centroids were previously standardised. In the graph on cluster analysis, same letters refer to same dimensions (i.e. a=Online Moral Disengagement; b=Traditional Moral Disengagement; c=Externalising behaviour); different numbers indicate significant differences across clusters. In the graph on cyberbullying different subscripts indicate significant differences across clusters.
Highlights

- Online MD is a domain-specific construct related but distinct from traditional MD
- Online MD significantly contributes to understanding misbehaviour in online contexts
- Externalising problems moderate the relation between MD and cyberbullying
- The interplay among both MDs, and externalising is key to understand cyberbullying
Table 1. Descriptive statistics for the OMD items

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Sk</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ONLINE MORAL DIENGAGEMENT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Assuming the identity of a classmate/friend online is just a game among friends (<em>euphemistic labelling</em>)</td>
<td>1.45</td>
<td>0.747</td>
<td>2.276</td>
<td>5.132</td>
</tr>
<tr>
<td>2. Sharing a video about someone else is just a way of paying attention to them (<em>disregarding/distorting consequences</em>)</td>
<td>1.39</td>
<td>0.626</td>
<td>2.482</td>
<td>6.606</td>
</tr>
<tr>
<td>3. If a child is groomed online by people with bad intentions, (s)he is responsible for not having been able to recognise them (<em>attribution of blame</em>)</td>
<td>1.93</td>
<td>1.133</td>
<td>1.197</td>
<td>0.950</td>
</tr>
<tr>
<td>4. It is right to share someone’s intimate images to highlight a problem (<em>moral justification</em>)</td>
<td>1.37</td>
<td>0.559</td>
<td>2.460</td>
<td>6.503</td>
</tr>
<tr>
<td>5. It is right to slander online a person who behaves like a beast (<em>de-humanisation</em>)</td>
<td>1.94</td>
<td>1.290</td>
<td>1.165</td>
<td>0.559</td>
</tr>
<tr>
<td>6. It is not that serious to insult someone on a social network because doing it in person would be worse (<em>advantageous comparison</em>)</td>
<td>1.27</td>
<td>0.497</td>
<td>3.220</td>
<td>11.288</td>
</tr>
<tr>
<td>7. If teachers do not monitor technology at school, children cannot be blamed if they use their smartphone to belittle a classmate (<em>displacement of responsibility</em>)</td>
<td>1.58</td>
<td>0.972</td>
<td>1.924</td>
<td>3.305</td>
</tr>
<tr>
<td>8. Teasing someone online is not so serious if everyone does it (<em>diffusion of responsibility</em>)</td>
<td>1.15</td>
<td>0.260</td>
<td>3.935</td>
<td>16.477</td>
</tr>
<tr>
<td><strong>MORAL DIENGAGEMENT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive Restructuring</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>item1</td>
<td>2.45</td>
<td>1.722</td>
<td>0.654</td>
<td>-0.691</td>
</tr>
<tr>
<td>item2</td>
<td>1.89</td>
<td>1.143</td>
<td>1.260</td>
<td>1.014</td>
</tr>
<tr>
<td>item3</td>
<td>1.24</td>
<td>0.493</td>
<td>3.609</td>
<td>13.921</td>
</tr>
<tr>
<td>item4</td>
<td>3.11</td>
<td>1.685</td>
<td>-0.011</td>
<td>-1.098</td>
</tr>
<tr>
<td>item5</td>
<td>1.24</td>
<td>0.447</td>
<td>3.750</td>
<td>15.875</td>
</tr>
<tr>
<td>item6</td>
<td>1.57</td>
<td>0.925</td>
<td>2.011</td>
<td>3.775</td>
</tr>
<tr>
<td>Minimizing one's agentive role</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>item7</td>
<td>2.03</td>
<td>1.513</td>
<td>1.066</td>
<td>0.116</td>
</tr>
<tr>
<td>item8</td>
<td>1.99</td>
<td>1.272</td>
<td>1.051</td>
<td>0.311</td>
</tr>
<tr>
<td>item9</td>
<td>2.41</td>
<td>1.479</td>
<td>0.595</td>
<td>-0.550</td>
</tr>
<tr>
<td>D disregard/Distorting the consequences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>item10</td>
<td>1.59</td>
<td>0.943</td>
<td>1.828</td>
<td>2.949</td>
</tr>
<tr>
<td>item11</td>
<td>1.74</td>
<td>1.077</td>
<td>1.421</td>
<td>1.321</td>
</tr>
<tr>
<td>Blaming/Dehumanising the victim</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>item12</td>
<td>1.95</td>
<td>1.317</td>
<td>1.193</td>
<td>0.643</td>
</tr>
<tr>
<td>item13</td>
<td>1.58</td>
<td>0.783</td>
<td>1.816</td>
<td>3.443</td>
</tr>
<tr>
<td>item14</td>
<td>1.38</td>
<td>0.679</td>
<td>2.565</td>
<td>6.618</td>
</tr>
</tbody>
</table>
Table 2. Multivariate Hierarchical Regression Model predicting cyberbullying

<table>
<thead>
<tr>
<th></th>
<th>STEP 1</th>
<th></th>
<th>STEP 2</th>
<th></th>
<th>STEP 3</th>
<th></th>
<th>ΔR²</th>
<th>F (2; 855)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SE</td>
<td>p</td>
<td>B</td>
<td>SE</td>
<td>p</td>
<td>B</td>
<td>SE</td>
</tr>
<tr>
<td>Age</td>
<td>.006</td>
<td>.002</td>
<td>.009</td>
<td>.005</td>
<td>.002</td>
<td>.007</td>
<td>.005</td>
<td>.002</td>
</tr>
<tr>
<td>Gender</td>
<td>-.015</td>
<td>.007</td>
<td>.042</td>
<td>-.001</td>
<td>.007</td>
<td>.777</td>
<td>-.002</td>
<td>.007</td>
</tr>
<tr>
<td>MD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.023</td>
<td>.008</td>
<td>.004</td>
<td>.020</td>
<td>.009</td>
<td>.015</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OMD</td>
<td>.055</td>
<td>.008</td>
<td>.000</td>
<td>.057</td>
<td>.009</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXT</td>
<td>.065</td>
<td>.010</td>
<td>.000</td>
<td>.062</td>
<td>.010</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MD*EXT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STEP 1</td>
<td>.013</td>
<td>.030</td>
<td>.014</td>
<td>.030</td>
<td>.013</td>
<td>.014</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes. Unstandardized estimates; MD= Moral Disengagement; OMD=Online Moral Disengagement; EXT=Externalising problems
Table 3. Simple Slope Analysis. The association between MD and cyberbullying in the three groups: Low, Medium and High Externalizing (controlling for OMD).

<table>
<thead>
<tr>
<th>Cyberbullying</th>
<th>Low Externalising (N=97; 12%)</th>
<th>Medium Externalising (N=585; 68%)</th>
<th>High Externalising (N=175; 20%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>P</td>
<td>B</td>
</tr>
<tr>
<td>MD</td>
<td>.209</td>
<td>.066</td>
<td>.102</td>
</tr>
<tr>
<td>OMD</td>
<td>.411</td>
<td>.000</td>
<td>.261</td>
</tr>
<tr>
<td>R²</td>
<td>.31</td>
<td>.11</td>
<td>.11</td>
</tr>
<tr>
<td>F</td>
<td>F(2, 96) = 22.631; p=.000</td>
<td>F(2, 582) = 36.070; p=.000</td>
<td>F(2, 175) = 11.006; p=.000</td>
</tr>
</tbody>
</table>

Notes. Unstardardized estimated. Significant values are bolded. MD=Moral Disengagement; OMD=Online Moral Disengagement.