

# Agreeing with moral disengagement statements: an experimental study with young adults

## *Concordância com afirmativas de desengajamento moral: um estudo experimental com jovens*

Guilherme Azzi **RUSSO**<sup>1</sup>  0000-0001-8763-5698

Roberta Gurgel **AZZI**<sup>2</sup>  0000-0003-0971-7852

Elias José de **LIMA JÚNIOR**<sup>3</sup>  0000-0003-3566-8679

Warley Guilger **CORRÊA**<sup>2</sup>  0000-0002-3322-5754

### Abstract

How do individuals react to statements that condone acts of corruption and violence? This article contributes to the understanding of how people respond to discourses related to antisocial practices by examining how (1) information about the person who made the statement, and (2) the use of additional information as justification for the antisocial statement affect individuals' agreement with this type of discourse. Relying on moral disengagement theory, we present an experimental study conducted with 487 high school students that assesses the level of agreement with statements made by Brazilian politicians that illustrate different mechanisms of moral disengagement. While overall results indicate low levels of agreement with these statements, comparisons across experimental groups suggest that providing case-based arguments as justification increases agreement with statements in favor of violence but reduces agreement with statements that condone acts of corruption.

**Keywords:** Adolescent; Cognition; Corruption; Moral development; Politics; Psychology, social.

### Resumo

*Como indivíduos reagem a declarações que aceitam atos de corrupção e violência? Este artigo contribui para a compreensão de como indivíduos respondem a discursos de práticas antissociais ao examinar como (1) a informação sobre o emissor*

▼ ▼ ▼ ▼ ▼

<sup>1</sup> Fundação Getúlio Vargas, Centro de Política e Economia do Setor Público, Escola de Administração de Empresas de São Paulo. R. Itapeva, 286, 10º andar, Bela Vista, 01332-000, São Paulo, SP, Brasil. Correspondence to: G.A. RUSSO. E-mail: <guilherme.russo@fgv.br>.

<sup>2</sup> Universidade Estadual de Campinas, Faculdade de Educação, Departamento de Psicologia Educacional. Campinas, SP, Brasil.

<sup>3</sup> Universidade Estadual Paulista Júlio de Mesquita, Instituto de Biociências, Programa de Pós-Graduação em Educação. Rio Claro, SP, Brasil.

Support: *Fundação de Amparo à Pesquisa do Estado de São Paulo* (nº 2018/08112-6).

▼ ▼ ▼ ▼ ▼

How to cite this article

Russo, G. A., Azzi, R. G., Lima, J. E., Jr., & Corrêa, W. G. (2021). Agreeing with moral disengagement statements: an experimental study with young adults. *Estudos de Psicologia* (Campinas), 38, e180041. <https://doi.org/10.1590/1982-0275202138e180041>



da frase e (2) o uso de dados adicionais como justificativa para a afirmação antissocial afetam sua concordância com este tipo de discurso. Utilizando-se da teoria de desengajamento moral, apresenta-se um estudo experimental conduzido com 487 estudantes de Ensino Médio, o qual avalia o grau de concordância com declarações feitas por políticos brasileiros que representam diferentes tipos de desengajamento moral. Enquanto os resultados gerais apontam um nível baixo de concordância com as declarações, as comparações entre os grupos experimentais sugerem que o uso de argumentos baseados em casos como justificação aumenta a concordância com afirmações em favor de violência, mas reduz a concordância com afirmações que aceitam atos de corrupção.

**Palavras-chave:** Adolescente; Cognição; Corrupção; Desenvolvimento moral; Política; Psicologia social.

How do individuals respond to discourses that condone antisocial practices of violence and corruption? Laws, rules, and norms shape how individuals behave through the process of socialization (Duarte, 2019). Socially-constructed references carry expectations about the consequences (positive and negative) of executing different actions. In turn, these expectations serve as parameters through which individuals self-regulate their behavior while looking for self-satisfaction and avoiding self-condemnation, legal and social sanctions (Bandura, 2016). Nevertheless, day-to-day life presents instances in which moral references (i.e., ethical standards) are simply not enough to avoid behavior that devalues or harms others. When individuals act in such ways, they tend to justify it to themselves that their actions do not contradict their personal morals.

The concept of moral disengagement speaks to this phenomenon of how individuals rationalize antisocial behavior, i.e. actions that undermine, hurt, or harm other human beings (Azzi, Lima Júnior, & Corrêa, 2017; Bandura, Azzi, & Tognetta, 2015). That is, moral disengagement seeks to explain why individuals who tend to behave according to moral standards based on respect for others justify moments in which they do not. The concept is best defined as a socio-cognitive process through which individuals justify an act that contradicts their personal moral standards in order to avoid self-censorship and self-depreciation (Bandura, 1999). But while one may argue that only a few bad apples carry out morally-disengaged behavior, decades of research suggest this type of behavior tends to take place with reasonable frequency in the daily life of ordinary individuals. Moral disengagement processes take place whenever one rationalizes an action that would entail a personal, legal, or social sanction. More specifically, moral disengagement happens when one behaves in contradiction with his or her moral standards, for instance, when one breaks a traffic law (Luiza Neto, 2009; Mognon & Santos, 2014), purposively slacks at work (Resende, 2015), litters, does not pay taxes appropriately, makes a joke about someone for their personal or physical characteristics (Farias, 2018; Jesus & Silva, 2018; Medeiros, Silveira, & Oliveira, 2018), interacts with peers (Molchanov, Almazova, & Kirsanov, 2019), or engages in cyberbully (Allison & Bussey, 2017; Bussey, Fitzpatrick, & Raman, 2015; Gini, Pozzoli, & Bussey, 2014; Luo & Bussey, 2019; Wang, Zhao, Yang, & Lei, 2019).

As these morally-disengaged practices occur in society, it is natural that individuals witness such behaviors often enough, forcing them to decide how to react. In other words, individuals often encounter speech and behavioral conduct that conflicts with socially acceptable moral standards. In fact, representatives in high political office often engage in moral disengagement reasoning in speeches about governmental actions (Cartledge, Bowman-Grieve, & Palasinski, 2015).

This study contributes to the understanding of how individuals respond to discourses that reinforce antisocial practices by examining how the authorship and informational contexts of such speeches affect the level of agreement with this type of reasoning. We present results from an experimental study conducted with high school students from the metropolitan region of *Campinas*, Brazil, designed to assess how individuals respond to statements about corruption and violence that represent examples of moral disengagement. More specifically, the study provides insights about individuals' level of agreement with moral disengagement speeches in three ways.

First, it evaluates the overall level of agreement with public statements made by politicians that characterize different mechanisms of moral disengagement. Second, it measures the extent to which (1) identifying the person who made the statement as a politician and (2) providing additional case-based arguments for the moral disengagement reasoning affect the level of agreement with these statements. The comparisons across these different scenarios are based on the random assignment of participants to three experimental conditions. Third, the article examines the heterogeneous effects of the experimental stimuli conditional on attitudes toward the military and individuals' religiosity. Overall, the results show a low level of agreement with the moral disengagement statements. Yet, they indicate that providing background information about the case presented in the moral disengagement statement tends to increase agreement with statements that support violence but reduce agreement with discourses intended to justify acts of corruption. Finally, the test of heterogeneous effects suggests that there may be important differences in how individuals respond to justification based to their pre-existing attitudes.

## Modeling and mechanisms of moral disengagement

According to social cognitive theory, individuals develop a group of moral references by being exposed to models within processes of social learning. Models can originate from a variety of sources and belong to different domains, i.e., social circles such as family, friends, and colleagues, but also publicly-known figures such as artists, journalists, athletes, and politicians. In this modeling process, individuals learn a series of moral standards that later serve as a compass for what is right and wrong, and ultimately, for how to behave in society (Caprara & Cervone, 2000; Yilmaz, Yilmaz, & Demir-Yilmaz, 2019). Making use of these learned references that carry expectations about the consequences of different avenues of action, individuals observe and appraise the social context to determine in which type of behavior they ought to engage (Bandura, 2016). For the most part, individuals seek to behave according to their moral standards in a way that results in personal satisfaction and avoids social and personal condemnation (Houwing & Bussey, 2017). Yet, with relative frequency, individuals conduct practices that contradict their own moral standards. And in order to avoid self-censure for violating these standards, they justify these attitudes and behavior through a process called moral disengagement (Bandura, 2016). In other words, as people want to avoid the self-condemnation and self-devaluation derived from moral-conflicting practices, they create justifications that turn the incompatible action into acceptable or less condemnable through moral disengagement. According to Bandura (1999, 2016), the process of moral disengagement can occur through eight different types of mechanisms clustered within four loci: Behavioral, Effects, Agency, and Victim.

At the *behavioral locus*, individuals seek to turn a harmful conduct into an acceptable behavior by framing the action as positive, or at least as less negative, by means of three mechanisms: (a) Moral justification: when one argues that the harmful conduct was executed in the name of benefiting society or based on other morals that are valued socially; (b) Advantageous comparison: when one compares a reprehensible action with conducts that are deemed worse than the original action; and (c) Euphemistic language: when one tries to make an antisocial conduct acceptable by calling it something less culpable, like using the term "collateral damage" for the death of innocent people.

At the *effects locus*, individuals attempt to argue that a harmful practice did not have detrimental effects by denying, distorting, minimizing, or disregarding the consequences of antisocial acts, like claiming that verbal offense does not lead to psychological harm.

At the *agency locus*, individuals attempt to deny, minimize, or obfuscate the responsibility of the harmful action's agent by means of two mechanisms: (a) Displacement of responsibility: when one claims the agent of an antisocial action cannot be blamed for such conduct because he or she was forced by a

superior order or circumstance; and (b) Diffusion of responsibility: when one argues that, because there is widespread responsibility for an antisocial conduct, one should not be blamed for playing only a small part.

At the *victim locus*, individuals attempt to justify a reprehensible conduct by placing the responsibility onto the victim due to his or her behavior or features, by means of two mechanisms: (a) Dehumanization: when one depicts the victim as destitute of human characteristics or attributes animal-like qualities in order to justify the harmful conduct or speech; and (b) Blame attribution: when one attributes the responsibility of a violent practice because of the victim's past behavior or characteristics, like assigning blame for sexual assault because of one's behavior or outfit.

Based on this theoretical framework, analysis of antisocial behavior should take into account the various contextual elements to understand why and how individuals make use of moral disengagement mechanisms. That is, situations with moral implications are a product of many "ingredients" that vary in importance according to the circumstances, such as the nature of the transgression, the frequency of its occurrence, the degrees of variation and form, the environment in which it took place, the situational and personal motivating factors, the immediate and long-term consequences, the nature of the damage (personal or material), if it was targeted at institutions or individuals, the characteristics of the agents (age, sex, ethnicity, and social status), and the characteristics of the victim and his or her perceived culpability (Bandura, 2016).

In addition to having an impact on the probability that a morally-disengaged behavior will take place, the context and personal demands of the agent should also affect the extent to which bystanders who witness such conduct reprehend it or not. In others words, the response from individuals who encounter actions or discourses that conflict with socially-shared moral values should also be a function of the context in which the antisocial action or speech occurs (Bandura, 2016). For instance, a content analysis of former U.S. President George Bush's speeches on war and terrorism indicates that his use of moral disengagement reasoning was coupled with other complex ideas such as patriotism, religious metaphors, and military success to justify violent acts that contradict socially-shared values (Cartledge et al., 2015). And because political leaders are perceived to be individuals who should be held to a higher moral standard, judgements of their actions may be more rigorous than if the exact same behavior was conducted by an ordinary citizen. Based on this idea, we derive a hypothesis that when the author of a statement that justifies an antisocial behavior with moral disengagement reasoning is identified as a politician, the level of agreement with the statements should be lower vis-à-vis a scenario when the authorship was not identified.

In addition to the status of the person who made the statement, we hypothesize that the context under which moral disengagement is used to justify actions of violence and corruption may affect the level of agreement from those who observe it. More specifically, being given descriptions of the circumstances can substantiate moral disengagement processes and a more detailed narrative of a case should increase the probability that the behavior is perceived as acceptable. As such, we derive a second hypothesis: the addition of a case-based explanation, such as extra sentences describing an event, should increase the level of agreement with statements that support harmful practices.

## Method

### Instruments

To assess the level of acceptance to speech that deploys moral disengagement, and the impact that information about the speaker and the background have on individuals' assessments, we produced a survey instrument to measure the level of agreement with six statements made by Brazilian politicians in the media, all of which characterize examples of moral disengagement.

To select the topics of these statements, we analyzed results from public opinion surveys that asked young adults what they believed were the main problems in the country and in their lives. Results from these polls indicate that, although health care and economic issues are perceived to be important problems by a considerable percentage of participants, larger shares of young adults tend to view violence and corruption as the main issues in their lives (Abramo & Branco, 2005; Waiselfisz, 2015).

As a next step, we put together a list of 22 controversial statements from Brazilian politicians that referred to corruption and violence. The list of statements and respective sources can be found in the supplementary information online appendix C at <[https://github.com/guirusso/Moral\\_Disengag\\_Exp\\_EstPsi2021/blob/master/MoralDisengagStatements\\_SupplementalFile.pdf](https://github.com/guirusso/Moral_Disengag_Exp_EstPsi2021/blob/master/MoralDisengagStatements_SupplementalFile.pdf)> (Russo, Azzi, Lima, & Corrêa, 2020). Out of the 22 statements, we selected six: three related to corruption and three to violence, in a way that comprised a range of political parties and levels of office. This list of statements also embodied varying mechanisms of moral disengagement, and stood in clear contradiction with the acceptable norms defined by the Brazilian law. The advantage of using actual statements is an increase in the level of ecological validity (Pasquali, 2017). The disadvantage of this approach in comparison to hypothetical statements is that we had to omit a few words to avoid recognition of the politician who made the statements, which could impact the respondents' level of agreement for reasons other than content itself.

For each statement, we also added a synthesis sentence that underscored the reasoning behind the moral disengagement. For example, for the first statement: "76 killed by the police is not a lot for me, the more criminals wiped out, the better. It will be less thugs causing trouble and avoiding the death of family men," we added the synthesis sentence: "Criminals ought to be killed by the police." Table 1 presents the six statements followed by their respective synthesis sentences and, in the second column, a short description of the moral disengagement mechanism. The statements and respective synthesis sentences represent six out of the eight mechanisms proposed by Bandura (1999, 2016).

After being exposed to the statements, respondents were asked to indicate their level of agreement with each statement based on a five-point Likert "completely agree-completely disagree" scale. But, in order

Table 1  
List of statements by politicians and their respective moral disengagement mechanisms

	Statements	Moral Disengagement Mechanism
1	76 killed by the police is not much for me, the more criminals wiped out, the better. It will be less thugs causing trouble and avoiding the death of family men. Criminals ought to be killed by the police.	Moral justification – by "avoiding the death of family men", killing becomes justified. Dehumanization – by dehumanizing the target with the word "criminal", killing becomes justified.
2	We are not going to fight violence by releasing doves in Copacabana or "hugging" the Rodrigo de Freitas Lagoon. Violence ought to be fought with violence.	Palliative comparison – by qualifying non-violence means as inefficient, acts of violence become justified as means to avoid more human suffering.
3	During 2003 and 2004, the party used unaccounted resources to pay campaign debts. We all know how a political campaign is run. No one should be punished for illegal actions if everyone does the same.	Euphemistic language – by calling it unaccounted resources, illegal campaign financing becomes justified. Diffusion of responsibility – if everyone does it, doing it becomes morally justified.
4	Nowadays, 16-year-old girls get boob jobs, have large backsides, put on mini-skirts, and flirt. Then the guy falls for it, goes to the motel, has intercourse, and goes to jail? One should not be blamed for harassing someone who is wearing provocative outfits.	Attribution of blame to the victim – by attributing blame to young girls, sexual harassment of minors becomes justified.
5	There was, eventually, malpractice with campaign resources but not systematic payments for political support in Congress. Doing something illegal once is not as bad as doing it often.	Palliative comparison – by comparing malpractices with campaign resources to something that seems more serious, illegal political practices become justified.
6	Trust me, given the amount a councilman is paid, one barely makes it through the year. Individuals who make little money cannot be blamed for theft.	Moral justification – by arguing it is a making-a-living issue, attaining financial resources illegally becomes justified.

to examine the impact of the authorship and the longer description of the case, we created three different versions of the statements, and participants were randomly assigned to one of the three versions. In the first experimental condition (“Politicians”), the question wording reported that the statements were made by Brazilian politicians and extracted from the media, while in the second condition (“Media”), it was not revealed that politicians were the authors of the statements. The wording for the first condition was, as follows: “Below, you will find statements made by Brazilian politicians and extracted from the media about different topics. For each statement, please tell us the extent to which you agree or disagree with the stated opinion”. The wording of the second condition did not include “made by Brazilian politicians”.

In the third experimental condition (“Synthesis”), respondents were not only not informed about the authorship, but also not shown the original statement by the politician that provided the background for the moral disengagement reasoning. Hence, those in the third condition were only presented with the synthesis sentence of each item. See the online appendix D at [https://github.com/guirusso/Moral\\_Disengag\\_Exp\\_EstPsi2021/blob/master/MoralDisengagStatements\\_SupplementalFile.pdf](https://github.com/guirusso/Moral_Disengag_Exp_EstPsi2021/blob/master/MoralDisengagStatements_SupplementalFile.pdf) (Russo et al., 2020), for a table with the differences between experimental conditions.

The benefit of randomizing the wording of the questions for various respondents is a fundamental increase in internal validity. The experimental design provides a higher potential for causal inference given that, if the randomization is successful, differences in the dependent variable between the three experimental groups should be attributed only to the variation controlled by the researcher, and not to other differences across groups (Campos, 2015).

## Procedures and Participants

The study was conducted with high school students from the metropolitan region of *Campinas*. The selection of high schools where the study was implemented was done purposively based on their proximity to the city of *Campinas* and previous contact with school staff. The reasons for the selection of students as participants of our study was threefold. First, our study was part of a larger research project that aimed at understanding the political behavior and attitudes toward democracy of the next generation. Second, young adults naturally have, on average, less knowledge of statements made by political leaders in the past, which should reduce the probability of recognizing the authorship of the sentences. Third, because younger individuals have less crystalized attitudes toward the political system, they should be less likely to be impacted by the information that the authors were politicians for reasons other than their status as elected leaders.

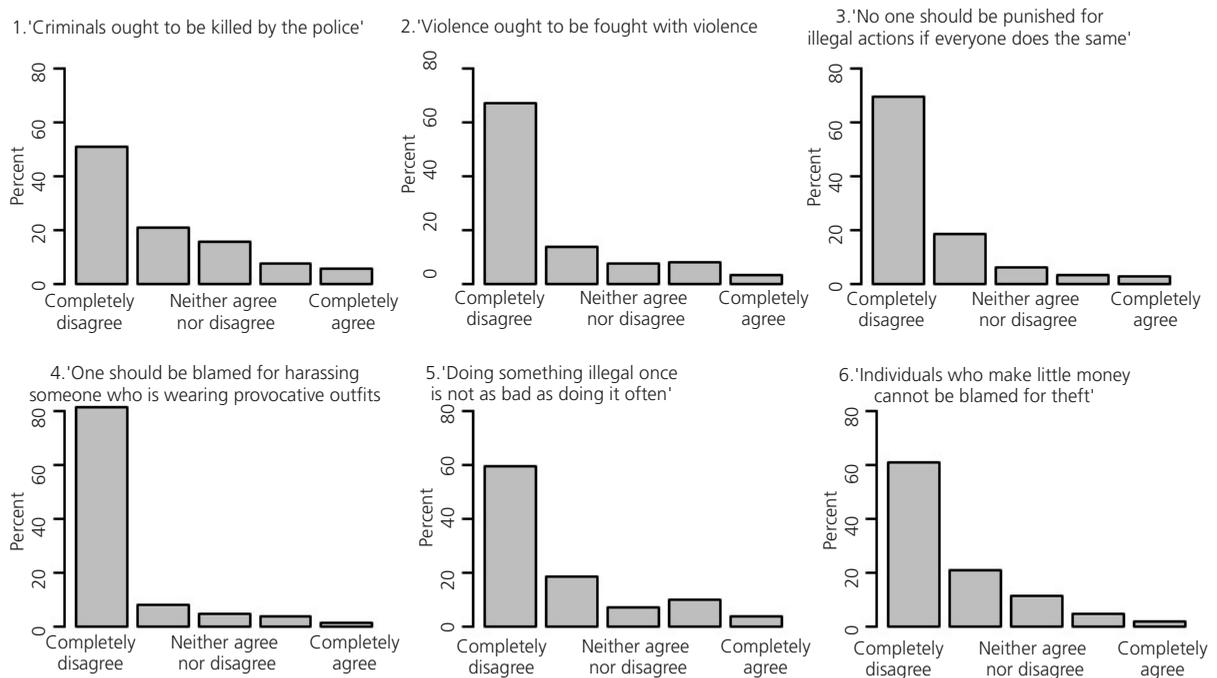
In each school, a member of the research team presented the motivation, goals, and instruments of the study to the school principal and asked for authorization to administer the study in a later date. After approval from the administration, the study was announced to the students along with a presentation of the informed consent form *Termo de Consentimento Livre e Esclarecido*. Students received two copies of the form, one to keep and one to be signed and handed in to the research team. The consent document presented information about the researchers’ institutional affiliation, the objectives, the absence of risks and compensation, as well as the anonymity and confidentiality, but did not provide the specific wording of survey items. Eighteen-year-olds were allowed to sign the document, while underage students had to have the form examined and signed by their parents or legal guardians. Before any data was collected, researchers visited the school a third time to collect the consent forms. In the fourth visit, students were asked to fill out the survey instrument on computers or on paper depending on the equipment availability in the school. Students responded to the questionnaire individually in classrooms or computer labs during regular class times. In the schools where students used a computer, random assignment to the experimental conditions was performed through the online survey platform Qualtrics, whereas in the schools where paper questionnaires were used, random assignment was manually implemented during the stacking of the questionnaires, but according to the randomization performed by a computer.

A total of 487 students between the ages of 15 and 21, from 11 public and private high schools from six municipalities (*Campinas, Indaiatuba, Paulínia, Sumaré, Hortolândia, and Jundiaí*) participated in the study. All respondents were sophomores, juniors, or seniors (2nd through 4th year) in technical or regular high schools. Roughly 61.0% of the students were female, and 70.0% were 16 or 17-year-olds. An overwhelming 96.7% of students reported using the internet daily, and 79.2% utilized the internet to follow political news. For more details about the sample, please see the online appendix E at <[https://github.com/guirusso/Moral\\_Disengag\\_Exp\\_EstPsi2021/blob/master/MoralDisengagStatements\\_SupplementalFile.pdf](https://github.com/guirusso/Moral_Disengag_Exp_EstPsi2021/blob/master/MoralDisengagStatements_SupplementalFile.pdf)> (Russo et al., 2020).

The randomization process split the sample into groups of similar size and composition. The group assigned to the experimental condition “Politicians” was comprised of 163 students (57.1% female, 71.2% 16 or 17-year-olds, and 62.0% who responded to the paper questionnaire), the one assigned to “Media” was composed of 158 students (66.9% female, 69.9% 16 or 17-year-olds, and 67.7% who responded to the paper questionnaire), and the “Synthesis” group was composed of 149 students (57.7% female, 68.2% 16 or 17-year-olds, and 66.4% who responded to the paper questionnaire). All of the research procedures followed principles of research ethics and were approved by the IRB at the *Universidade Estadual de Campinas* (Unicamp, State University of *Campinas*), nº 51531115.5.0000.5404.

## Results

Before showing the differences between experimental condition groups, Figure 1 presents the distribution of responses to each of the statements among those in the third condition (“Synthesis”), which serve as the baseline (control) for the comparisons that follow.



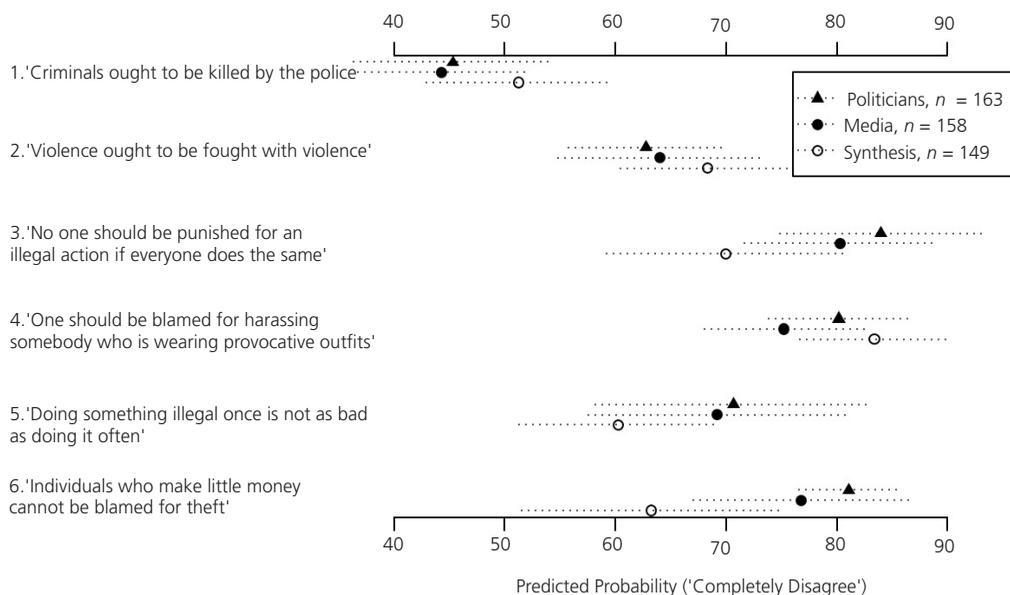
**Figure 1.** Level of agreement with statements among respondents from experimental condition “synthesis.”

Note: The y-axis in each of the six graphs presents the percentage of the 149 respondents randomly assigned to the “Synthesis” experimental group who (dis)agree with statement. The five bars in each graph represent the response options in the following order: “completely disagree”, “somewhat disagree”, “neither agree nor disagree”, “somewhat agree”, and “completely agree.”

Generally, the levels of agreement with the statements were low. Nevertheless, there are important differences between the statements. The percentage of young adults who expressed complete disagreement with the first statement is 30 percentage points lower than the level presented with the fourth. In fact, statements that make use of the moral justification mechanism (1 and 6) presented the lowest levels of disagreement, followed by the ones that make use of advantageous comparison (2 and 5), and contrast with the widespread disagreement with statements 3 and 4.

The second objective of the study was to evaluate the impact of (1) informing that politicians were the authors of the statements and (2) presenting the background statement to the moral disengagement synthesis sentence. In the previous section, we showed that there are few demographic differences between the experimental groups; yet, to avoid potential spurious effects, we compared the levels of agreement between groups using multivariate ordinal logistic regression models performed with the Stata Statistical Software: Release 13, (Stata Corp, 2013), that include sex, education of the mother, religious affiliation, and questionnaire mode (online vs. paper) as controls. To take into account the homogeneity of respondents, observations are also clustered by school. The regression results are presented in the appendix H at <[https://github.com/guirusso/Moral\\_Disengag\\_Exp\\_EstPsi2021/blob/master/MoralDisengagStatements\\_SupplementalFile.pdf](https://github.com/guirusso/Moral_Disengag_Exp_EstPsi2021/blob/master/MoralDisengagStatements_SupplementalFile.pdf)> (Russo et al., 2020), and in Figure 2 we present the predicted probabilities of a “completely disagree” response across the experimental conditions, while holding other predictors constant at their means (Campos, 2015).

The impact of informing that politicians were the authors of the statements can be seen by the difference between the filled triangles, which present the predicted probabilities for those in the condition “Politicians”, and the filled circles, which present the predicted probabilities for the condition “Media”. While the results show that, on average, there was not a statistically significant effect associated with informing the authorship, in two out of three statements related with politicians’ behavior and corruption (3 and 6),



**Figure 2.** Predicted probabilities of “completely disagree” by experimental condition.

Note: The filled triangle, filled circle, and empty circle represent the average predicted probability of a “Completely disagree” for conditions “Politicians”, “Media”, and “Synthesis”, respectively, while holding all other regression predictors at their means. The dashed lines indicate the confidence interval around the estimates with an alpha-level of 0.05. For the entire regression results and significance of beta predictors, see the online appendix H at <[https://github.com/guirusso/Moral\\_Disengag\\_Exp\\_EstPsi2021/blob/master/MoralDisengagStatements\\_SupplementalFile.pdf](https://github.com/guirusso/Moral_Disengag_Exp_EstPsi2021/blob/master/MoralDisengagStatements_SupplementalFile.pdf)> (Russo et al., 2020).

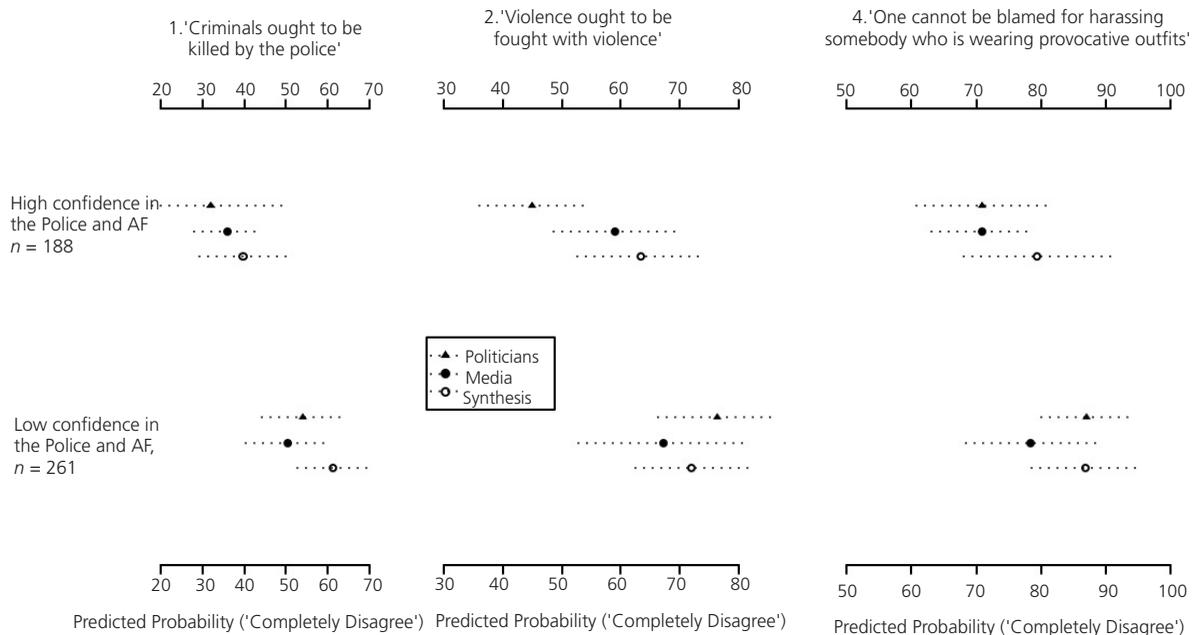
the predicted probability of complete disagreement was five percentage points higher among those who were informed the authorship. The level of disagreement was also higher among those in the condition “Politicians” for statement 4, which attributes blame for harassment to the victim.

The impact of presenting the case-based argument, measured by the comparison of the filled circles (“Media”) with the empty circles (“Synthesis”), is substantially larger, although the direction of the effect seems to be conditional on the statement’s topic. For the three statements pertaining to violence (1, 2, and 4), the background information decreased the probability of a complete disagreement, which suggests the mechanisms of moral justification, advantageous comparison, and dehumanization were effective. For the corruption statements (3, 5, and 6), however, the effect was in the opposite direction. Compared to the responses offered by individuals who only received the synthesis sentences, the level of disagreement with the speech was higher when the justifying information about the case for such conduct was presented. These differences are both statistical and substantively significant. Comparing to the predicted probabilities of the sixth statement, we find the difference between conditions to be larger than 13 percentage points. It is important to note that this result may be inflated by the fact that the specific synthesis sentence could prime individuals to think about lower-income individuals, whereas the background statement explicitly discusses the remuneration for councilmen. Still, this analysis indicates that including more details in a speech that portrays an antisocial conduct as more acceptable can have heterogeneous effects (i.e. positive or negative) on the level of agreement depending on the context and issue domain. In other words, an additional statement to explain the conduct had a positive effect for violence, but a negative effect for statements describing corruption practices.

We also explored the mechanisms through which young adults develop higher levels of agreement with this type of discourse by examining the extent to which the effect of the stimuli varies across individuals who have distinct attitudes. For instance, given that the agreement with statements that promote violence is linked to the justification of police action as a means to prevent and reduce crime, we expected that those who have more favorable views toward law enforcement, operationalized here with trust items on the military police and the Armed Forces, should express higher agreement toward speech that relates to violence, and also higher levels of agreement when the statement originated from a politician, given that it personalizes the discourse (Soares, Barbosa, Matos, & Mendes, 2018). On the other hand, those who have low levels of trust in these institutions should not perceive the speech as more adequate when knowing a politician professed it. In fact, it is possible that those who have lower levels of trust in these institutions become even more unreceptive to the idea given they do not find legitimacy in the speech of a politician who rationalizes forceful actions by the police.

To test these hypotheses, we created an additive measure (from 2 to 14 points) based on responses to the items asking about the level of trust in the Military Police (*Policia Militar*) and the Armed Forces, each measured with a seven-point scale in which 1 means “Not at all” and 7 means “A lot”. The exact wording of the question is presented in the online appendix A at [https://github.com/guirusso/Moral\\_Disengag\\_Exp\\_EstPsi2021/blob/master/MoralDisengagStatements\\_SupplementalFile.pdf](https://github.com/guirusso/Moral_Disengag_Exp_EstPsi2021/blob/master/MoralDisengagStatements_SupplementalFile.pdf) (Russo et al., 2020). Respondents with scores lower than 9 were categorized as having a low trust in the Police and the Armed Forces, and those with a score 9 or higher as showing high trust. This variable was then introduced in the regression models and interacted with the experimental condition variables. That is, we analyze the level of agreement with the statement across experimental conditions among young adults who expressed high levels of trust in these institutions and among those who reported lower levels of trust. Similarly to the previous analysis, Figure 3 presents the predicted probabilities of a response “Completely disagree” for each condition.

Among those who expressed a high level of confidence in law enforcement institutions, the predicted probability for those assigned to the experimental condition “Politicians” was lower than those assigned to the condition “Media” for statements 1 and 2, and similar for statement 4. Yet, among those who trust

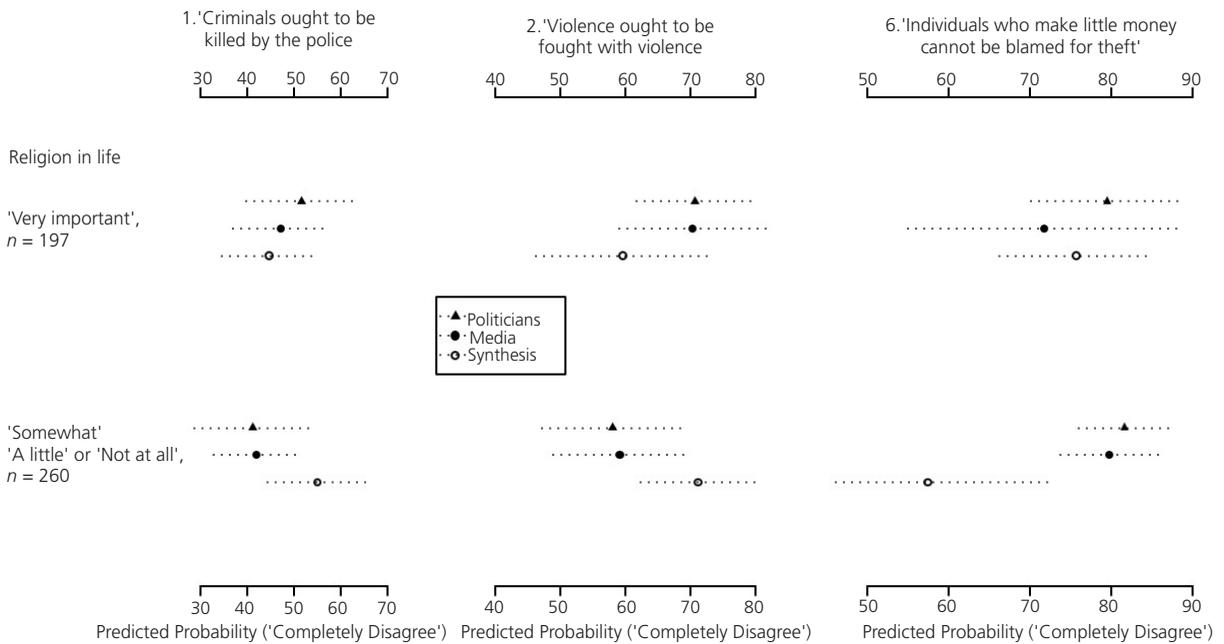


**Figure 3.** Predicted probabilities of “completely disagree” by experimental condition and level of trust in the military police and armed forces. Note: The filled triangle, filled circle, and empty circle represent the average predicted probability of a “Completely disagree” for conditions “Politicians”, “Media”, and “Synthesis”, respectively, while holding all other regression predictors at their means. “AF” in the figure stands for Armed Forces. The dashed lines indicate the confidence interval around the estimates with an alpha-level of 0.05. The number of high confidence respondents in the experimental conditions were 62 (Politicians), 60 (Media) and 66 (Synthesis), and lower level of confidence: 95 (Politicians), 91 (Media) and 75 (Synthesis). For the entire regression results and significance of beta predictors, see the online appendix J at <[https://github.com/guirusso/Moral\\_Disengag\\_Exp\\_EstPsi2021/blob/master/MoralDisengagStatements\\_SupplementalFile.pdf](https://github.com/guirusso/Moral_Disengag_Exp_EstPsi2021/blob/master/MoralDisengagStatements_SupplementalFile.pdf)> (Russo et al., 2020).

these institutions less, the effect of informing the authorship was positive. Differently stated, within the less trusting group, the probability of complete disagreement is higher among those who were informed of the authorship of the statement compared to those who were not.

In a second test of heterogeneous effects, we examined the impact of the different stimuli conditional on varying level of religiosity. Participants were classified based on their responses to a question on how important religion was in their lives. Those who said religion was “Very important” were classified as showing high religiosity, while those who said “Somewhat”, “A little” or “Not at all” were classified as having low religiosity. Similarly, to the previous analysis, Figure 4 presents the predicted probabilities for statements 1, 2, and 6 for each condition among those with high and low levels of religiosity.

Results from Figure 4 indicate that the inclusion of the case-based argument had a different impact among those with high religiosity compared to those of lower religiosity. For the more religious, the background statement seems to have increased levels of disagreement for sentences 1 and 2. But among those with lower levels of religiosity, the presentation of more information seems to have lowered disagreement. In other words, while the presentation of the reasoning for violence increased disagreement among those with high religiosity, those with lower levels of religiosity expressed higher agreement. For statement 6, nonetheless, the result was reversed. The background statement significantly lowered the level of disagreement among those with low religiosity, but slightly increased among those with high religiosity. While we believe that detailed theory and empirical research is needed to understand the specific mechanisms for these results, these analyses strongly indicate that the impact of additional information can vary according to prior beliefs and values.



**Figure 4.** Predicted probabilities of “completely disagree” by experimental condition and level of religiosity.

Note: The filled triangle, filled circle, and empty circle represent the average predicted probability of a “Completely disagree” for conditions “Politicians”, “Media”, and “Synthesis”, respectively, while holding all other regression predictors at their means. The dashed lines indicate the confidence interval around the estimates with an alpha-level of 0.05. The number of high religiosity respondents in the experimental conditions were 71 (Politicians), 72 (Media) and 54 (Synthesis), and lower religiosity: 80 (Politicians), 91 (Media) and 89 (Synthesis). For the entire regression results and significance of beta predictors, see the online appendix J at <[https://github.com/guirusso/Moral\\_Disengag\\_Exp\\_EstPsi2021/blob/master/MoralDisengagStatements\\_SupplementalFile.pdf](https://github.com/guirusso/Moral_Disengag_Exp_EstPsi2021/blob/master/MoralDisengagStatements_SupplementalFile.pdf)> (Russo et al., 2020).

## Discussion

In this article, we contribute to the moral disengagement literature by testing two hypotheses about the conditions under which individuals accept morally-disengaged rhetoric. First, we hypothesized that when an individual knows that a speech that supports antisocial practices was made by a politician, as opposed to not knowing the identity of the author, levels of acceptance to antisocial speech should be lower. Second, we argued that when presented with case-based information that supports the antisocial rhetoric, acceptance to the moral disengagement reasoning should be higher. We presented the results of an experimental study conducted with high school students that assesses the level of agreement with statements made by Brazilian politicians that condone acts of corruption and violence through different mechanisms of moral disengagement. The results provide very limited non-statistically significant evidence for the first hypothesis – that informing politicians were the authors of these statements would increase disagreement.

This may be the case because of limitations of the study. First, it is possible that individuals could easily deduce the statements originated from politicians given their topics and tones. Second, some of the statements and their authors were relatively well-known by the general public, which could have influenced the potential impact of the authorship stimuli. Third, given that the stimuli was presented as a bundle rather than individually randomized items, it is possible one or two items affected the rest. Fourth, it is important to note that because our measures show very low levels of agreement with the statements, we have limited statistical power. Finally, it is possible that the results were affected by social desirability bias. But, in spite of the lack of statistically significant results, for the three statements that pertained to corruption, we still find a higher level of disagreement among those who were told politicians made the statements in comparison to their counterparts who read the background statement but were not informed about the authorship.

As for the second hypothesis, the results provide evidence for the expected effect when we analyze the statements about acts of violence. That is, being exposed to more information about the case increases the levels of acceptance. However, we also find a diametrically opposed effect for the corruption statements. In other words, while the contextualizing information about the morally disengaged speech pertaining to acts of violence increased the likelihood of a positive response, the contextualization of speech that justified acts of corruption decreased the levels of agreement. These results are, to a certain extent, surprising given that they indicate that a longer attempt to explain a moral disengagement reasoning can backfire. That is, the direction of the process in which individuals evaluate antisocial practices can change depending not only on specific elements of the moral circumstance, but also on the larger domain of the issue. According to our results, more information about the circumstance around an act of corruption only increases disagreement with a moral disengagement reasoning.

This finding has important implications for current Brazilian politics. It suggests that attempts to contextualize practices of corruption, be it by blaming the system, emphasizing its pervasiveness, or the inherent nature for the day-to-day work in politics, tend to increase levels of disagreement with this type of action. On the other hand, providing reasoning for speech that supports violence, be it through dehumanization, blaming the aggression on the victims, or moral justification, tends to decrease the level of disagreement.

It is important to note, however, that the acts of corruption presented in the statements do not have a positive ramification such as increased efficiency or any other benefit. In contrast, the acts of violence discussed in the statements could be interpreted by a few as having some positive effects for society, e.g., less criminals on the streets. But rather than contrast, this comparison of issue domains echoes our conclusion about the difference between the two domains given political corruption is by definition a private gain in detriment of the public interest, while violence can be argued, at times, to be perpetrated in defense of common interests.

Finally, we performed an analysis of heterogeneous effects to highlight potential mechanisms. The results indicate that prior beliefs such as attitudes toward the police and personal religiosity may condition the appraisal of speeches that use moral disengagement. We believe this finding may contribute to future research on how individuals use prior moral standards and other attitudes to evaluate moral disengagement discourse.

Future studies may also explore if the results hold when applied to a different sample, with other antisocial speech, or with varying contextual ingredients such as descriptions of the behavior or speech, or characteristics of the agent, the impact and those affected by it. While researchers may use aggregate data to find these effects, we believe individual-level data from in-depth interviews, large-n surveys, and experimental research with subjects may yield more informative results, especially pertaining to mechanisms. Future research may also be well advised to carefully consider how to improve measurement and the power of stimuli.

## Acknowledgment

We thank our colleagues from the former research center NEAPSI who assisted in the study, as well as the schools, parents, and especially, the students who graciously accepted to participate in the study.

## Contributors

G. A. RUSSO contributed in the research design, data collection, analysis, and interpretation, as well as in the writing and revisions. R. G. AZZI and J. E. LIMA JÚNIOR contributed in the research design, data collection, writing, and revisions. W. G. CORRÊA contributed in the research design and data collection.

## References

- Abramo, H. W., & Branco, P. P. M. (2005). *Retratos da juventude brasileira: análises de uma pesquisa nacional*. São Paulo: Fundação Perseu Abramo. Recuperado de [http://www.uel.br/prograd/gepe/materiais/retratos\\_juventude\\_brasileira.pdf](http://www.uel.br/prograd/gepe/materiais/retratos_juventude_brasileira.pdf)
- Allison, K., & 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>
- Azzi, R. G., Lima Júnior, E. J., & Corrêa, W. G. (2017). *Agência moral na visão da teoria social cognitiva*. Porto Alegre: Letra 1.
- Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, *3*(3), 193-209. [https://doi.org/10.1207/s15327957pspr0303\\_3](https://doi.org/10.1207/s15327957pspr0303_3)
- Bandura, A. (2016). *Moral disengagement: how people do harm and live with themselves*. New York: Worth Publishers.
- Bandura, A., Azzi, R. G., & Tognetta, L. A. (2015). *Desengajamento moral: teoria e pesquisa a partir da teoria social cognitiva*. Campinas: Mercado de Letras.
- Bussey, K., Fitzpatrick, S., & Raman, A. (2015). The role of moral disengagement and self-efficacy in cyberbullying. *Journal of School Violence*, *14*(1), 30-46. <https://doi.org/10.1080/15388220.2014.954045>
- Campos, L. F. L. (2015). *Métodos e técnicas de pesquisa em psicologia*. Campinas: Alínea.
- Caprara, G. V., & Cervone, D. (2000). *Personality: determinants, dynamics, and potentials*. New York: Cambridge University Press.
- Cartledge, S., Bowman-Grieve, L., & Palasinski, M. (2015). The mechanisms of moral disengagement in George W. Bush's "War on Terror" rhetoric. *The Qualitative Report*, *20*(11), 1905-1921. Retrieved from <https://nsuworks.nova.edu/tqr/vol20/iss11/14/>
- Duarte, S. C. (2019). Classe, técnica e moral: o Estado em perspectiva sociológica. *Revista Interscientia*, *7*(1), 83-101. Recuperado de <https://periodicos.unipe.br/index.php/interscientia/article/view/1005/616>
- Farias, J. E. M. (2018). *Influência da manipulação de jeitinho e de privacidade no engajamento em comportamento desonesto* (Dissertação de mestrado não-publicada). Universidade de Brasília. Recuperado de [https://repositorio.unb.br/bitstream/10482/31893/1/2017\\_%c3%a9ssicaEstherMachadoFarias.pdf](https://repositorio.unb.br/bitstream/10482/31893/1/2017_%c3%a9ssicaEstherMachadoFarias.pdf)
- Gini, G., Pozzoli, T., & Bussey, K. (2014). 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>
- Houwing, F., & Bussey, K. (2017). Moral disengagement and the propensity to endorse physical punishment practices. *Journal of Child and Family Studies*, *26*(4), 1206-1218. <https://doi.org/10.1007/s10826-016-0643-y>
- Jesus, S. E., & Silva, K. R. X. P. (2018). Um material didático sobre desengajamento moral e violência de gênero. *Revista Contemporânea de Educação*, *13*(28), 875-893. Recuperado de <https://revistas.ufjr.br/index.php/rce/article/view/15759/pdf>
- Luiza-Neto, I. (2009). *Justificativas de motoristas para infrações de trânsito: esboçando um modelo* (Dissertação de mestrado não-publicada). Universidade de Brasília. Recuperado de [https://repositorio.unb.br/bitstream/10482/4806/1/2009\\_IngridLuizaNeto\\_Disserta%c3%a7%c3%a3o.pdf](https://repositorio.unb.br/bitstream/10482/4806/1/2009_IngridLuizaNeto_Disserta%c3%a7%c3%a3o.pdf)
- Luo, A., & Bussey, K. (2019). The selectivity of moral disengagement in defenders of cyberbullying: contextual moral disengagement. *Computers in Human Behavior*, *93*, 318-325. <https://doi.org/10.1016/j.chb.2018.12.038>
- Medeiros, C. R. O., Silveira, R. A., & Oliveira, L. B. (2018). Mitos no desengajamento moral: retóricas da Samarco em um crime corporativo. *Revista de Administração Contemporânea*, *22*(1), 70-91. <http://dx.doi.org/10.1590/1982-7849rac2018160310>
- Mogon, J. F., & Santos, A. A. A. (2014). Autoeficácia para dirigir, desengajamento moral e impulsividade em motoristas. *Psico-USF*, *19*(3), 457-466. <https://doi.org/10.1590/1413-82712014019003008>
- Molchanov, S. V., Almazova, O. V., & Kirsanov, K. A. (2019). Types of adolescent moral disengagement strategies. *The European Proceedings of Social & Behavioural Sciences*, *58*(7), 442-449. Retrieved from <https://www.futureacademy.org.uk/files/images/upload/PSYRGGU2019F058.pdf>
- Pasquali, L. (2017). Validade dos testes. *Examen: Política, Gestão e Avaliação da Educação*, *1*(1), 14-48. Recuperado de <https://examen.emnuvens.com.br/rev/article/view/19>
- Resende, M. M. (2015). *Jeitinho brasileiro tem jeito? O efeito do jeitinho brasileiro e da identidade moral no comportamento ético nas organizações* (Dissertação de mestrado não-publicada). Universidade de Brasília. Recuperado de [https://repositorio.unb.br/bitstream/10482/19932/1/2015\\_Mar%c3%adliaMesquitaResende.pdf](https://repositorio.unb.br/bitstream/10482/19932/1/2015_Mar%c3%adliaMesquitaResende.pdf)

- Russo, G. A., Azzi, R. G., Lima, J. E., Jr., & Corrêa, W. G. (2020). *Supplementary information appendix: agreeing with moral disengagement statements. An experimental study with young adults*. California: GitHub. Retrieved from [https://github.com/guirusso/Moral\\_Disengag\\_Exp\\_EstPsi2021/blob/master/MoralDisengagStatements\\_SupplementalFile.pdf](https://github.com/guirusso/Moral_Disengag_Exp_EstPsi2021/blob/master/MoralDisengagStatements_SupplementalFile.pdf)
- Soares, M., Barbosa, M., Matos, R., & Mendes, S. M. (2018). Public protest and police violence: moral disengagement and its role in police repression of public demonstrations in Portugal. *Peace and Conflict: Journal of Peace Psychology*, 24(1), 27-35. <https://doi.org/10.1037/pac0000277>
- StataCorp (2013). *Stata Statistical Software: release 13*. College Station: StataCorp LP.
- Waiselfisz, J. J. (2015). *Mapa da violência 2015: mortes matadas por armas de fogo*. Brasília: Juventude Viva. Recuperado de <http://www.mapadaviolencia.net.br/pdf2015/mapaViolencia2015.pdf>
- Wang, X., Zhao, F., Yang, J., & Lei, L. (2019). School climate and adolescents' cyberbullying perpetration: a moderated mediation model of moral disengagement and friends' moral identity. *Journal of Interpersonal Violence, Forthcoming*, 00(0)1-22. <https://doi.org/10.1177/0886260519860089>
- Yılmaz, M., Yılmaz, U., & Demir-Yılmaz, E. N. (2019). The relation between social learning and visual culture. *International Electronic Journal of Elementary Education*, 11(4), 421-427. <https://doi.org/10.26822/iejee.2019450837>

Received: April 23, 2018  
Final version: May 22, 2020  
Approved: June 3, 2020