ORIGINAL ARTICLE

Is Online Disinhibition Related to Cyberdating Abuse Perpetration through Moral Disengagement? The Moderating Role of Gender, Sexism, and Cybervictimization

M. Dolores Sánchez-Hernández[1] · M. Carmen Herrera[1] · Francisca Expósito[1]

Abstract

Evidence suggests that online disinhibition enhances the likelihood of perpetrating cyberbullying by increasing moral disengagement; however, these psychological mechanisms have not been examined in the context of cyberdating abuse. In the current study ($N = 362$), we examined whether online disinhibition would predict more frequent direct cyberaggression toward a partner through greater moral disengagement, and explored the moderating role of gender, sexism, and past experiences of cyberdating abuse victimization. The results indicated that online disinhibition was positively correlated with moral disengagement, which in turn predicted more frequent direct cyberaggression toward partners. In addition, participants' gender and past experiences of cyberdating abuse victimization moderated this relationship: (a) more online disinhibition was associated with greater moral disengagement in men (vs. women), which in turn predicted more direct cyberaggression toward partners and (b) more online disinhibition was linked to greater moral disengagement, which in turn predicted more direct cyberaggression perpetration toward partners among individuals with frequent past victimization experiences (vs. low past victimization experiences). These findings highlight online disinhibition and moral disengagement as potential risk factors that may heighten direct cyberaggression against partners, as well as enhance our understanding of the circumstances determining its occurrence. Scholars and practitioners may use this work to develop and test psychoeducational programs to prevent cyberdating abuse through mitigating the occurrence of these disinhibiting factors in romantic.

Keywords Cyberdating abuse · Online disinhibition · Moral disengagement · Romantic relationships · Gender differences · Sexism · Victimization

Introduction

As a form of virtual intimate partner violence (IPV), cyberdating abuse is a multidimensional construct that includes various types of behaviors enacted against partners through digital media (Gámez-Guadix et al., 2018). Although there is a lack of consensus regarding cyberdating abuse terminology, operationalization, and measurement, which makes it difficult to determine the real incidence of cyberdating abuse (Soto & Ibabe, 2022), the international literature reveals alarming levels of this violence among young people, with reports of victimization rates ranging from 5.8% to 92% and reports of perpetration rates ranging from 6% to 93.7%

(Caridade et al., 2019). Soriano-Ayala et al. (2023) recently observed that in a sample of adolescents aged 13–20 years, 53.3% acknowledged perpetrating cyberdating abuse against their partners in the past year, while 62.1% reported experiencing it. Several studies have linked more experiences of cyberdating abuse with lower self-esteem and self-efficacy, and higher anxiety and depression symptomatology, emotional distress, and involvement in risky behaviors (e.g., substance use or contraceptive nonuse; Borrajo & Gámez-Guadix, 2016; Hancock et al., 2017; Hinduja & Patchin, 2021; Lu et al., 2018; Wright, 2016).

According to Borrajo et al. (2015b), cyberdating abuse can be classified into *direct cyberaggression*—deliberate behaviors intended to harm the partner (e.g., sending insulting and/or humiliating messages to the partner through digital media or sending intimate photos, images and/or videos of the partner to other people without his or her permission); and *cybercontrol*—online behaviors aimed at controlling the partner (e.g., checking

✉ M. Carmen Herrera
  mcherrer@ugr.es

1   Department of Social Psychology, Faculty of Psychology, University of Granada, Campus Universitario Cartuja, S/N, 18071 Granada, Spain

a partner's SNS, WhatsApp, or mail accounts without his or her permission or calling excessively to determine the whereabouts of the partner and whom they are with). Both types of cyberdating abuse involve cyberviolence within romantic relationships with detrimental effects for the individuals and the relationship (Deans & Bhogal, 2019); however, there are differences between them. Direct cyberaggression tends to be more explicit with the intention to harm the partner and thus easier to identify as a form of IPV, whereas cybercontrol is more indirect and tends to be accepted and normalized among young people (Donoso-Vázquez et al., 2018; Stonard et al., 2017). The international literature suggests that rates of direct aggression perpetration range from 10.6% (Borrajo et al., 2015a) to 14.7% (Caridade et al., 2019), and rates of cybercontrol perpetration range from 49.6% (Van Ouytsel et al., 2017) to 88.4% (Borrajo et al., 2015b). Similarly, the prevalence of direct cyberaggression victimization ranges from 14% (Borrajo et al., 2015b) to 31.7% (Gámez-Guadix et al., 2016) whereas cybercontrol victimization ranges from 65% (Van Ouytsel et al., 2017) to 81% (Gámez-Guadix et al., 2016).

Although a substantial body of literature has focused on examining predictors of cyberdating abuse perpetration (e.g., Branson & March, 2021; Mahoney et al., 2022), to date, no known studies have examined whether direct cyberaggression and cybercontrol can be differentiated by the psychological mechanisms predicting them. Specifically, research on cyberbullying perpetration has noted that online disinhibition, which refers to a state of uninhibited engagement in behaviors in the online environment (Suler, 2004), predicts more frequent perpetration (e.g., Falla et al., 2021; Wang et al., 2022). Moreover, drawing on Bandura's (1986, 1989) social cognitive theory, prior research demonstrates that the positive association between online disinhibition and cyberbullying perpetration is mediated through the activation of moral disengagement mechanisms (Wang & Ngai, 2020), which are considered to be important psychological processes contributing to the maintenance and justification of violent behaviors in different contexts (e.g., Bussey et al., 2015; Caprara et al., 2014). To our knowledge, no studies have examined the role of online disinhibition and moral disengagement in relation to direct cyberaggression and cybercontrol behaviors. The current study examined whether online disinhibition is indirectly associated with more frequent perpetration of both forms of cyberdating abuse via increases in moral disengagement. In addition, we examined three potential moderators of these associations to account for individual and relational variability, including gender, sexist attitudes, and prior cyberdating abuse victimization.

The Role of Online Disinhibition

Researchers have coined the term online disinhibition to refer to the psychological state in which individuals feel more liberated, uninhibited, and predisposed to perform certain behaviors in the online environment (Cheung et al., 2020). This mechanism operating in digital interactions explains why people express themselves and behave differently online compared to the offline environment (Longden, 2014). Six factors of the virtual environment interact with each other and can contribute to a state of online disinhibition (Suler, 2004): (a) *dissociative anonymity* (the degree to which people believe that they can change or hide their true identity in the online environment), (b) *invisibility* (the tendency to perceive that one cannot be physically seen by others in the online environment), (c) *asynchrony* (the perception that digital communication allows for delayed responses in interpersonal interactions), (d) *solipsistic introjections* (the tendency to interpret ambiguous online messages based on one's own expectations or beliefs), (e) *dissociative imagination* (the degree to which a person views the online environment as an imaginary world that has no link to reality), and (f) *authority minimization* (the belief that the influence of authority figures in real life is diminished or absent in the online environment).

According to Suler (2004), online disinhibition is considered *benign* when individuals feel more relaxed and willing to reveal their emotions, fears, likes, dislikes, preferences, or kindness to others in the online environment, whereas it is defined as *toxic* when individuals manifest inappropriate behaviors in the online environment (e.g., anger, insults, threats, criticism) in which they would not engage in a face-to-face context. This toxic online disinhibition is precisely what becomes a relevant precipitant of inappropriate online behaviors among adolescents and young adults, such as cyberbullying perpetration (Sanfilippo et al., 2017; Wang et al., 2021). However, the influence of this variable has not been evaluated in the context of cyberdating abuse. We believe that these factors, which are unique to the online environment, could be positively related to the perpetration of cyberdating abuse behaviors, mainly those that involve an intention to inflict harm against a partner (i.e., direct cyberaggression).

The Role of Moral Disengagement

Bandura's (1986, 1989) social-cognitive theory postulates that to understand human development and behavior, it is essential to examine how personal and environmental factors interact with each other. In this line, previous research has suggested that specific characteristics of a virtual context (e.g., anonymity, invisibility, and authority minimization) may operate in conjunction with personal factors, such as moral disengagement, to contribute to the rise of online deviant behaviors, such as cyberbullying (Runions & Bak, 2015).

Moral disengagement refers to the activation of cognitive processes of internal self-regulation by which individuals disengage from their moral responsibilities. Bandura (1999) theorized that internal moral control can be disconnected from blameworthiness by (a) reinterpreting one's behavior so that it is not considered immoral, (b) obscuring action agency (i.e., capacity to act intentionally) to minimize one's responsibility for causing harm, (c) distorting or ignoring the harmful consequences arising from one's actions, and/or (d) disregarding the victims of abuse by devaluing them as human beings and blaming them for the events.

Attending to the online context, some researchers have found that the possibility of maintaining anonymity and the feeling of impunity contribute to the moral justification and triggering of cyberviolence acts (e.g., Barlett & Helmstetter, 2018; Stonard, 2021). Likewise, invisibility in cyberspace has been associated with a lack of direct feedback about the emotional consequences victims suffer and a reduced ability to empathize with them, which often leads to more explicit cyberaggression behaviors by perpetrators (Smith, 2019). Thus, people are more likely to feel liberated from moral principles because of the increased psychological distance between the actor and the victim and between inappropriate actions and their harmful consequences (Naquin et al., 2010). The sense of asynchrony may also promote moral disengagement (Runions et al., 2013). For example, hurtful or offensive information posted online (e.g., photos, videos, or comments) may remain unchanged over time. Consequently, victims can revisit the offensive content at any time, and such content can exacerbate their negative emotions (i.e., anger and rage), inducing a desire to get revenge and increasing the likelihood of perpetrating cyberviolence (Runions & Bak, 2015). Researchers have recently observed that moral disengagement among cybervictims may lead to the internalization of violence as an appropriate strategy to resolve conflicts (Wachs et al., 2022).

In sum, the idiosyncrasy of the virtual environment may promote moral disengagement to justify violent acts (Bandura, 2016). This is supported by studies conducted on several online deviant behaviors, such as cyberbullying (e.g., Paciello et al., 2020), online racism (e.g., D'Errico & Paciello, 2018), software piracy (e.g., Lowry et al., 2017), and online hate speech (e.g., Wachs et al., 2022). However, rather than demonizing the online environment, it is necessary to understand how these sociocognitive processes lead to online deviant behaviors, such as cyberdating abuse, to prevent them and mitigate their potential negative effects. Specifically, moral disengagement has been considered a mediator of the effects of personal-level predictors in morally problematic behavior (e.g., Caprara et al., 2014; Paciello et al., 2013). For example, Wang and Ngai (2020) recently examined whether the positive relationship between the psychological tendency toward online disinhibition and

cyberbullying perpetration was mediated by moral disengagement in a sample of adolescents. They found that certain online disinhibition factors (i.e., anonymity and asynchrony) were related indirectly to cyberbullying via increases in moral disengagement. That is, the disinhibiting factors of anonymity and asynchrony were related to greater moral disengagement, which, in turn, was associated with greater engagement in cyberbullying. Following Bandura's (1986, 1989) social-cognitive framework, we examine how factors of online disinhibition (anonymity, invisibility, asynchrony, etc.) may operate jointly with moral disengagement to predict cyberdating abuse perpetration, which remains an unexplored area in the literature.

Gender, Sexism, and Past Experiences of Cyberdating Abuse Victimization as Moderators

Notably, the disinhibitory mechanisms of online disinhibition and moral disengagement do not consistently relate to the perpetration of cyberviolence (Moore, 2015); they may be moderated by other individual and interpersonal factors. Hence, using Bandura's (1986, 1989) social cognitive theory as a basis, it is necessary to understand under what circumstances these psychological processes operate to broadly examine the context in which cyberdating abuse takes place. Gender is one potential moderating variable in this context. Prior research has indicated that young men experience greater toxic online disinhibition (e.g., Wang & Ngai, 2020; Wang et al., 2021) and greater moral disengagement (e.g., Navas et al., 2021; Sánchez-Jiménez & Muñoz-Fernández, 2021) than young women. Moreover, past research examining gender norms, male privilege, and gender inequality has demonstrated that men are more likely to justify and perpetrate multiple forms of violence against women, including IPV (e.g., Glick & Fiske, 2001; Moya, 2003). Therefore, we would expect stronger associations among online disinhibition, moral disengagement, and cyberabuse dating among men compared to women.

A second potential moderator is endorsement of sexist attitudes that assume the inferiority of women, which may interact with gender to influence how people perceive violence (Karasavva et al., 2022; Klement et al., 2019) and, consequently, how they make moral judgments (Bandura, 2016). According to ambivalent sexism theory (Glick & Fiske, 1996, 2001), sexist attitudes are ambivalent, consisting of both hostile and benevolent feelings and behaviors. Hostile sexism consists of an antagonistic attitude toward women, primarily those who challenge male supremacy and thus do not adhere to normative roles of femininity. In contrast, benevolent sexism adopts a positive view of respect and adoration towards women who assume pre-established roles, while also considering them as in need of male protection. However, these interrelated attitudes toward women,

although seemingly positive, remain sexist, as they have a stereotypical and restricted view of women that perpetuates gender hierarchies (Glick & Fiske, 2001).

Previous research has shown that both hostile and benevolent sexism are implicated in the justification and acceptability of IPV against women (e.g., Martín-Fernández et al., 2018; Sánchez-Hernández et al., 2020), and they are strongly associated with the perpetration of direct cyberaggression toward a partner (e.g., insults, threats, or humiliation through digital media), particularly among men (Linares et al., 2021; Martínez-Pecino & Durán, 2019; Rodríguez-Domínguez et al., 2018). Indeed, some authors suggest that both sexist attitudes and moral disengagement serve to normalize IPV against women and exonerate the perpetrators (Page & Pina, 2015). In this regard, several studies have observed a positive association between sexist attitudes (both hostile and benevolent) and moral disengagement in situations of IPV against women (Navas et al., 2021; Rollero & De Piccoli, 2020) and dating abuse (Sánchez-Jiménez & Muñoz-Fernández, 2021). In particular, Sánchez-Jiménez and Muñoz-Fernández (2021) discovered that teenagers exhibiting high levels of sexism (both hostile and benevolent) and moral disengagement tended to engage in more psychological dating aggression. Given the above, it would be reasonable to assume that men with strong sexist attitudes could more easily conform to the moral norms accepted by themselves and their environment and exercise direct cyberaggression toward partners. This research aims to test whether online disinhibition may be related to higher levels of moral disengagement primarily in men with high sexist attitudes, and, consequently, with a higher frequency of direct cyberaggression toward partners.

The last moderator is past experiences of cyberdating abuse, as having previously been a victim or perpetrator of cyberdating abuse increases the risk of perpetrating or suffering this type of violence in the future, regardless of the participant's gender (Fernández-González et al., 2020). Specifically, Villora et al., (2019a, b, 2021) found that cyberdating abuse perpetration and victimization were highly and positively associated with direct cyberaggression and cybercontrol in samples of university students. Thus, in the online environment, victims can easily engage in online reactive violence and take on the role of aggressors (Smith et al., 2018). Given the above, we considered the possibility that experiences of cybervictimization moderate the association between these psychological disinhibitory mechanisms (i.e., moral disengagement and online disinhibition) and the perpetration of cyberdating abuse. In this respect, recent research has found that moral disengagement is positively associated with victimization and reciprocal violence in the IPV context (e.g., Cuadrado-Gordillo & Fernández-Antelo, 2019). Likewise, it has been observed that factors associated with online disinhibition may increase the likelihood that victims of cyberbullying may become cyberperpetrators,

regardless of gender (Wong-Lo & Bullock, 2014). Therefore, it would be reasonable to expect that past experiences of cybervictimization within relationships may exacerbate the relationship between these psychological disinhibitory mechanisms (i.e., moral disengagement and online disinhibition) and cyberdating abuse perpetration.

Current Study

The current study examines whether online disinhibition is positively associated with cyberdating abuse behaviors via increases in moral disengagement. According to our theoretical model, we expected that online disinhibition would positively predict moral disengagement (Hypothesis 1a) and moral disengagement would positively predict direct cyberaggression (i.e., explicit and intentional cyberdating abuse manifestations), but not cybercontrol (Hypothesis 1b). Likewise, building on the work of Wang and Ngai (2020), we also expected an indirect effect of online disinhibition on the perpetration of direct cyberaggression via moral disengagement. That is, higher levels of online disinhibition would be associated with higher levels of moral disengagement, which, in turn, would be related to more frequent direct cyberaggression, but not cybercontrol (Hypothesis 1c).

We also examined the moderating role of gender, sexism, and past experiences of cyberdating abuse victimization on the aforementioned association. Specifically, we hypothesized that high online disinhibition would be primarily associated with higher levels of moral disengagement in men (vs. women), which, in turn, would be related to a higher frequency of direct cyberaggression against the partner (Hypothesis 2). Likewise, we hypothesized that sexism (both hostile and benevolent) would interact with gender to moderate the aforementioned relationship. That is, high online disinhibition would be primarily associated with higher levels of moral disengagement in men with high hostile sexism (vs. low hostile sexism; Hypothesis 3a) and high benevolent sexism (vs. low benevolent sexism; Hypothesis 3b), which, in turn, would be related to a higher frequency of direct cyberaggression against the partner (see Fig. 1, Panel A). Finally, we hypothesized that past experiences of cyberdating abuse victimization would exacerbate both the indirect effect of online disinhibition on the perpetration of direct cyberaggression via moral disengagement and its direct effect: (a) online disinhibition would be associated with higher moral disengagement, which, in turn, would be related to a higher frequency of direct cyberaggression perpetration toward one's partner mainly in highly victimized (vs. less victimized) individuals (Hypothesis 4a), and (b) online disinhibition would be predictive of increases in the frequency of direct cyberaggression perpetration mainly in highly victimized (vs. less victimized) participants (Hypothesis 4b; see Fig. 1, Panel B).

Panel A
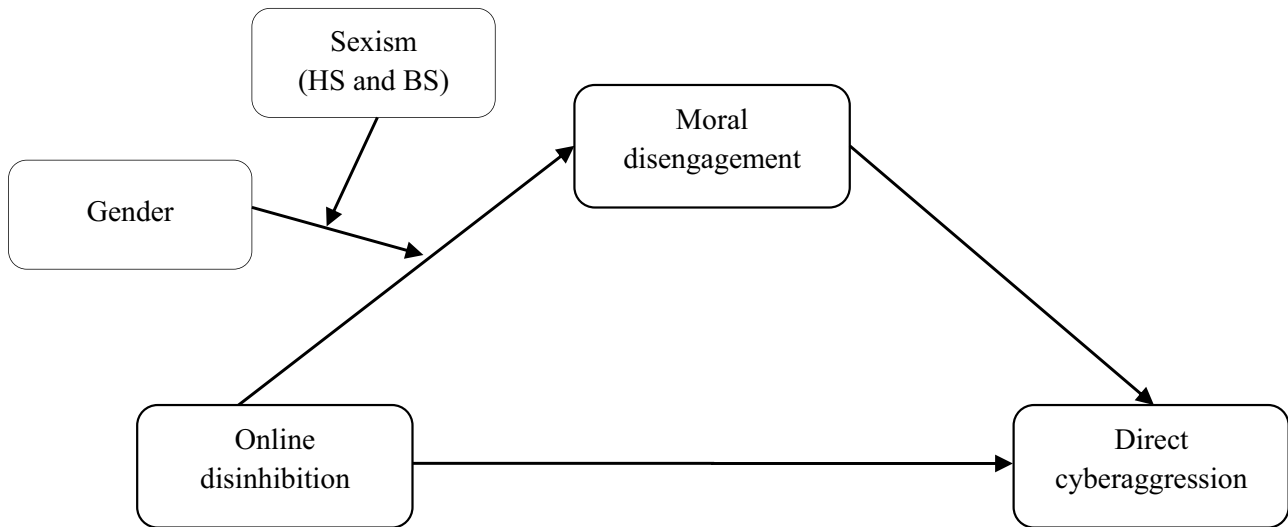
Panel B
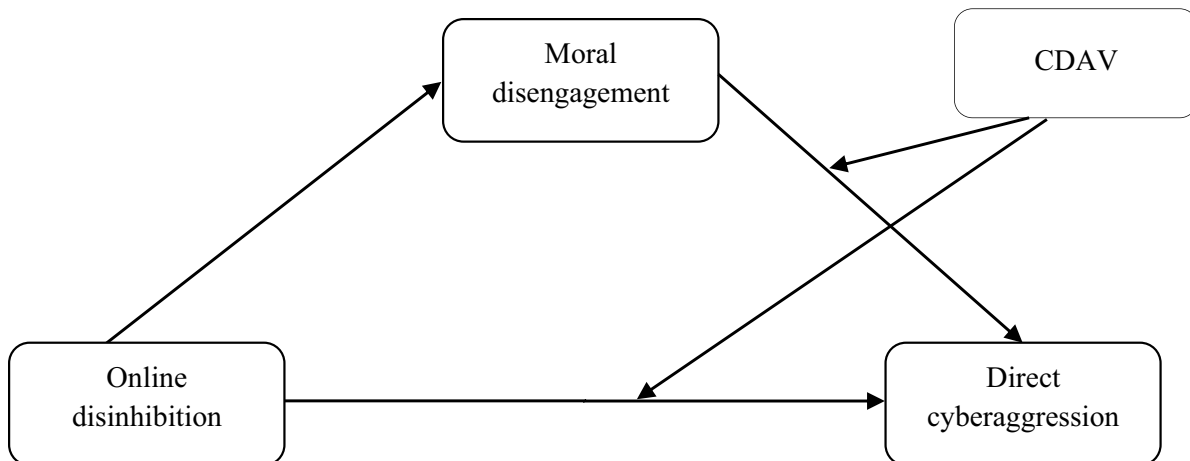
Note. HS = hostile sexism; BS = benevolent sexism; CDAV = cyberdating abuse victimization

Fig. 1 Proposed Conceptual Models: Moral Disengagement as Mediator of the Online Disinhibition–Direct Cyberaggression Perpetration Link, Moderated by Gender and Sexism (Panel **A**); Moral Disengage-ment as Mediator of the Online Disinhibition–Direct Cyberaggression Perpetration Link, Moderated by Past Experiences of Cyberdating Abuse Victimization (Panel **B**)

Method

Participants and Procedure

From the initial sample collected ($N = 605$), 227 (37.52%) respondents were removed because they did not complete the full questionnaire and 14 (2.31%) because they failed attention check items. Moreover, we excluded two participants who identified their gender as "other" from our analyses, because cyberdating abuse experiences may differ between partner categories defined by gender identity (i.e., cisgender vs. non-cisgender; Butler et al., 2023), and our work is focused

on examining differences in binary gender. Thus, the final sample consisted of 362 participants of Spanish nationality ($M_{age}=25.10$, $SD=4.45$, range 18–35 years; 45.9%, $n=166$ men; 54.1% women, $n=196$). All of them had a heterosexual orientation, of which 84 (23.2%) reported being single (52.4%; $n=44$ men, and 47.6%; $n=40$ women) and 278 (76.8%) were in a romantic relationship (43.9%; $n=122$ men, and 56.1%; $n=156$ women) at the time of data collection.

We implemented a cross-sectional survey design. Specifically, we used the LimeSurvey research platform to develop an online survey containing the variables of interest. Following a snowball sampling, we distributed the survey through an open-access link on several online social media platforms (i.e., email and social network sites: Facebook, Twitter, and WhatsApp). To participate in our study, respondents had to fulfill the following criteria: (a) having Spanish nationality, (b) having a heterosexual orientation, (c) having been in a past or current romantic relationship, and (d) being between 18 and 35 years of age. All inclusion criteria were confirmed by a set of checking questions at the end of the survey. We delimited this age range because the largest percentages of frequent internet and social-network users in Spain correspond to young people—92.6% of individuals aged 16 to 24 and 82.2% of those aged 25 to 34 (Statista, 2022)—and the highest levels of cyberdating abuse victimization and perpetration occur among adolescents and young adults (for a review, see Caridade et al., 2019). Moreover, previous authors (e.g., Oleszkowicz & Misztela, 2015; Sánchez-Hernández et al., 2020) have delimited the range of young adults between 18 and 35 years because the traditional markers of adulthood (e.g., marriage, parenthood, financial independence, and homeownership) are typically achieved by around the age of 35 (Arnett, 2000).

Data were collected during May and June 2021. Before completing the questionnaire, we provided information to the participants about the general study's goal (i.e., "to examine different psychological and relational processes that could be influencing the development of couple relationships") and its anonymous and voluntary nature. First, they had to give their consent to voluntarily participate in our research, according to the Declaration of Helsinki, and then fill in a single questionnaire based on their personal opinions and experiences. They were not paid for their participation. Participants took approximately 20 min to complete the task. This study was conducted after receiving the approval of Ethics Committee of University of Granada. The data and materials are available on the Open Science Framework [OSF].

Measures

Online Disinhibition

We used the Online Disinhibition Scale (ODS; Cheung et al., 2020), which is focused on evaluating different factors associated with feeling and behaving in an uninhibited way in the virtual environment described earlier in the article. The ODS consisted of 23 items divided into six dimensions: (a) *dissociative anonymity* (four items, e.g., "I feel that I can hide my identity online"), (b) *invisibility* (five items, e.g., "I feel like I am invisible in the online environment"), (c) *asynchronicity* (three items, e.g., "I can postpone replying to others in the online environment"), (d) *solipsistic introjections* (three items, e.g., "I perceive how that person's intended to talk about in the online environment"); (e) *dissociative imagination* (four items, e.g., "The online environment has no connection to reality"), and (f) *minimization of authority* (four items, e.g., "I am away from real life authorities in the online environment"). Participants used a 7-point Likert-type response scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). We computed an average global online disinhibition score, with higher scores indicating more online disinhibition ($\alpha = .85$).

Moral Disengagement

We administered the Moral Disengagement scale (MMDS-S; Bandura et al., 1996; Rubio-Garay et al., 2017) to assess the extent to which people justify and rationalize moral transgressions. Specifically, the MMDS-S is composed of 32 Likert-type items measuring eight different types of moral disengagement mechanisms: (a) *moral justification* (four items, e.g., "It is all right to fight to protect your friends"), (b) *euphemistic labeling* (four items, e.g., "To hit obnoxious classmates is just giving them 'a lesson'"), (c) *advantageous comparison* (four items, e.g., "Damaging some property is no big deal when you consider that others do worse"), (d) *displacement of responsibility* (four items, e.g., "If people are living in poor conditions, they are not responsible for attacking"), (e) *diffusion of responsibility* (four items, e.g., "If a group decides together to do something harmful, the responsibility lies with the entire group"), (f) *distortion of consequences* (four items, e.g., "Telling small lies doesn't really do any harm"), (g) *dehumanization* (four items, e.g., "Some people are like animals"), and (h) *attribution of blame* (four items, e.g., "Some people deserve to be mistreated because of their actions"). The response format was a 5-point Likert-type scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). For the purposes of our study, we calculated an average global moral disengagement score, where higher scores indicate greater moral disengagement ($\alpha = .83$).

Direct Cyberaggression and Cybercontrol Perpetration

We used the 20-item perpetration subscale of the Cyber Dating Abuse Questionnaire (CDAQ; Borrajo et al., 2015b) to assess the frequency of two types of behaviors against one's

partner: (a) *direct cyberaggression* (11 items; e.g., "I wrote a comment on the wall of SNSs to insult or humiliate my partner or ex-partner"; $\alpha = .89$) and (b) *cybercontrol* (nine items; e.g., "I have controlled the time at which my partner or ex-partner could last connect to mobile applications"; $\alpha = .83$). The response format was a 6-point Likert-type scale with the following anchors: 1 (*never*), 2 (*not in the last year, but it occurred before*), 3 (*rarely: 1 or 2 times*), 4 (*sometimes: between 3 and 10 times*), 5 (*often: between 10 and 20 times*), and 6 (*always: more than 20 times*). Average scores were calculated for each behavior: Higher scores indicate a higher frequency of perpetration.

Past Experiences of Cyberdating Abuse Victimization

We administered the 20-item victimization subscale of the CDAQ (Borrajo et al., 2015b) to measure the frequency of cyberdating abuse victimization in the past. This measure assessed two dimensions: (a) *direct cyberaggression* victimization (11 items; e.g., "My partner or ex-partner created a fake profile about me on a social network to cause me problems") and (b) *cybercontrol* victimization (nine items; e.g., "My partner or ex-partner used my passwords [phone, social networks, email] to snoop on my messages and/or contacts without my permission"). The response format was the same as for the measure described above. We calculated an average global victimization score, where higher scores indicate higher frequency of cyberdating abuse victimization ($\alpha = .92$).

Hostile and Benevolent Sexism

We used the Ambivalent Sexism Inventory (ASI; Expósito et al., 1998) to assess the participants' sexist attitudes. It is composed of 22 Likert-type items subdivided into two dimensions: (a) *hostile sexism* (11 items, e.g., "Women exaggerate the problems they have at work"; $\alpha = .93$) and (b) *benevolent sexism* (11 items, e.g., "Women should be loved and protected by men"; $\alpha = .84$). Participants responded using a 6-point Likert-type response format ranging from 0 (*strongly disagree*) to 5 (*strongly agree*). Average scores were calculated for each dimension of sexism; higher scores indicate stronger sexist attitudes.

Sociodemographic Information

Data about participants' gender ("What is your gender? Man/Woman/Other") and age ("What is your age?"), and whether they were in a relationship at the time of data collection ("Are you currently in a relationship? Yes/No") were collected.

Analytic Strategy

We analyzed the data using SPSS (Version 25). Before conducting the primary analyses, we assessed the assumptions of normality and multicollinearity, and calculated the main descriptive statistics and correlations among the study variables. Additionally, we performed independent samples t-tests to examine potential gender differences (1 = man; 2 = woman). Thereafter, to analyze the indirect effects of online disinhibition on direct cyberaggression perpetration via moral disengagement (Hypothesis 1), we performed a mediation analysis using Model 4 of the PROCESS program (Version 4.1; Hayes, 2018). We included online disinhibition as the predictor (X), direct cyberaggression perpetration as the criterion variable (Y), and moral disengagement as a mediating variable (M1). We used the nonparametric bootstrapping procedure with 10,000 resamples to estimate the 95% confidence interval (CI) for all analyses in PROCESS. In accordance with Schoemann et al. (2017), we also conducted a post hoc power analysis for simple mediation using Monte Carlo simulations (5,000 replications and 20,000 draws) to test indirect effects. Our sample ($N = 362$) had the ability to detect an indirect effect with 82% power for direct cyberaggression perpetration.

Next, to test whether gender (Hypothesis 2) moderates the indirect effect of online disinhibition on the perpetration of direct cyberaggression via moral disengagement, we conducted a moderated mediation analysis using Model 7 of the PROCESS program. The online disinhibition was included as a predictor variable (X), direct cyberaggression perpetration as a criterion variable (Y), moral disengagement as a mediating variable, and gender as moderating variable (W). Similarly, we conducted two moderated-moderated mediation analyses using Model 11 (PROCESS program) to test whether sexism (both hostile and benevolent) interacts with gender to moderate the aforementioned relationship (Hypothesis 3). The online disinhibition was included as a predictor variable (X), direct cyberaggression perpetration as a criterion variable (Y), moral disengagement as a mediating variable, and gender (W) and sexism (hostile and benevolent, separately; Z) as moderating variables. Finally, we performed a moderated mediation analysis using Model 15 (PROCESS program) to test whether past experiences of cyberdating abuse victimization acts as a moderating variable exacerbating both the indirect effect of online disinhibition on the perpetration of direct cyberaggression via moral disengagement and its direct effect (Hypothesis 4). Online disinhibition was included as a predictor variable (X), direct cyberaggression as a criterion variable (Y), moral disengagement as a mediating variable, and past experiences of cyber victimization as moderating variable (W).

We firstly conducted the analyses without covariates, and then assessed each model with covariates to compare results.

Table 1 Descriptive Statistics, Matrix Correlations, and Tests of Mean Differences Based on Gender Among Study Variables

	1	2	3	4	5	6	7	8	9
1. Online disinhibition	–								
2. Moral disengagement	.31**	–							
3. DCAP	.06	.16**	–						
4. CCP	.06	.10	.35***	–					
5. Hostile sexism	.14**	.22***	.20***	.12*	–				
6. Benevolent sexism	.18***	.34***	.21***	.17***	.63***	–			
7. CDAV	.11*	.12*	.26***	.61***	.15**	.16**	–		
8. Gender[a]	−.03	−.24***	−.11*	.06	−.28***	−.27***	−.07	–	
9. Current relationship[b]	.07	.08	−.02	.17***	.04	.06	.30***	−.07	–
Potential range	(1–7)	(1–5)	(1–6)	(1–6)	(0–5)	(0–5)	(1–6)	–	–
$M_{overall}$ (SD)	3.33 (0.78)	1.99 (0.40)	1.06 (0.27)	1.54 (0.60)	0.85 (0.94)	0.85 (1.74)	1.38 (0.53)	–	–
M_{men} (SD)	3.35 (0.78)	2.09 (0.43)	1.09 (0.37)	1.50 (0.59)	1.14 (0.90)	1.35 (1.84)	1.42 (0.53)	–	–
M_{women} (SD)	3.30 (0.78)	1.90 (0.34)	1.03 (1.12)	1.57 (0.61)	0.61 (0.91)	0.42 (1.52)	1.35 (0.53)	–	–
Gender difference t	0.62	4.55***	1.99*	− 1.14	5.52***	5.20***	1.34	–	–
Cohen's d	0.06	0.49	0.07	− 0.11	0.59	0.55	0.11	–	–
Skewness/Kurtosis	−0.02/0.19	0.47/0.09	1.80/1.31	0.48/−0.37	0.50/−0.13	0.30/−0.37	0.59/−0.37	–	–

Note. $N_{overall} = 362$; $N_{men} = 166$, $N_{women} = 196$.

DCAP direct cyberaggression perpetration, *CCP* cybercontrol perpetration, *CDAV* cyberdating abuse victimization.

[a]1 = man, 2 = woman; [b]1 = yes, 2 = no.

*p < .05; **p < .01; ***p < .001.

Specifically, we controlled for gender, sexism (hostile and/or benevolent), and past experiences of cyberdating abuse victimization when the moderating role of these variables was not assessed. Moreover, in all models, we controlled for whether participants were in a romantic relationship at the time of data collection (1 = *yes*; 2 = *no*) because individuals often show a bias towards recognizing cyberabuse behaviors within their romantic relationships, normalizing and justifying this type of violence (e.g., Sánchez-Hernández et al., 2020, 2023, 2024).

Results

Preliminary Analyses

There were no missing values in our dataset. The kurtosis and skewness values of online disinhibition and moral disengagement fell within acceptable limits of ± 2, indicating a normal distribution (Gravetter & Wallnau, 2014). However, the skewness and kurtosis values of the remaining variables (i.e., direct cyberaggression perpetration, cybercontrol perpetration, hostile and benevolent sexism, and experiences of cyberdating abuse victimization) were not normally distributed. Therefore, we applied a two-step transformation to these variables to achieve a more normal distribution and enhance the robustness of the main effects (Templeton, 2011). After transforming the variables, there was no evidence of multicollinearity, as the correlations among

variables were all below .80 (Shrestha, 2020). Descriptive statistics, correlations, and gender differences in study variables are provided in Table 1.

Moral Disengagement as Mediator of the Online Disinhibition–Direct Cyberaggression Perpetration Link

As shown in Table 2, online disinhibition positively predicted moral disengagement. Similarly, moral disengagement positively predicted direct cyberaggression perpetration. Moreover, the results highlighted a statistically significant indirect effect of online disinhibition on direct cyberaggression perpetration via moral disengagement ($b = 0.01$, $SE = 0.01$, 95% CI [0.01, 0.03]). Specifically, high online disinhibition was associated with higher levels of moral disengagement, which, in turn, was related to more frequent direct cyberaggression toward a partner. The variables included in the model predicted 7.75% of the variance ($R^2 = .08$) in direct cyberaggression toward a partner. The total effect of online disinhibition on direct cyberaggression perpetration was also significant ($b = 0.04$, $SE = 0.01$, 95% CI [0.02, 0.07]). As expected, the indirect effect of online disinhibition on cybercontrol via moral disengagement was not statistically significant ($b = 0.03$, $SE = 0.01$, 95% CI [−0.003, 0.05]).

When we tested this model controlling the effects of the covariates (gender, hostile and benevolent sexism, past experiences of cyberdating abuse victimization, and

**Table 2** Unstandardized Regression Coefficients, Standard Error, and Summary Information for the Mediation Model 4 ("Direct Cyberaggression" and "Cybercontrol")

	Moral Disengagement (MD)			Direct Cyberaggression		
	Coeff.	SE	95% CI	Coeff.	SE	95% CI
Constant	2.02***	0.19	[1.64, 2.39]	−0.42***	0.08	[−0.57, −0.27]
Online disinhibition	0.13***	0.02	[0.08, 0.18]	0.01	0.01	[−0.01, 0.03]
MD				0.06**	0.02	[0.03, 0.10]
	$R^2 = .21$ $F(7, 354) = 13.5, p < .001$			$R^2 = .78$ $F(8, 353) = 154.66, p < .001$		
	Effects	SE			SE	95% CI
Total effect	0.02				0.01	[0.003, 0.04]
Indirect effect	0.01				0.003	[0.002, 0.02]

	MD			Cybercontrol		
	Coeff.	SE	95% CI	Coeff.	SE	95% CI
Constant	1.46***	0.09	[1.29, 1.63]	1.18***	0.18	[0.82, 1.55]
Online disinhibition	0.16***	0.03	[0.11, 0.21]	0.02	0.04	[−0.06, 0.11]
MD				0.14	0.08	[−0.03, 0.30]
	$R^2 = .10$ $F(1, 360) = 38.93, p < .001$			$R^2 = .01$ $F(2, 359) = 2.00, p = .137$		
	Effects	SE			SE	95% CI
Total effect	0.04				0.04	[−0.04, 0.12]
Indirect effect	0.02				0.01	[−0.003, 0.05]

Note. N = 362.

SE standard error, *CI* confidence interval.

**p < .01, ***p < .001.

Table 3 Unstandardized Regression Coefficients, Standard Error, and Summary Information for the Moderated Mediation Model 7 ("Direct Cyberaggression")

	Moral Disengagement (MD)			Direct Cyberaggression		
	Coeff.	*SE*	95% CI	*Coeff.*	*SE*	95% CI
Constant	1.21***	0.28	[0.67, 1.76]	0.20***	0.06	[0.09, 0.31]
Online disinhibition (OD)	0.31*	0.08	[0.16, 0.47]	0.02	0.01	[−0.002, 0.05]
Gender[a]	0.16	0.17	[−0.17, 0.50]			
OD X Gender	−0.10*	0.05	[−0.20, −0.01]			
MD				0.11***	0.03	[0.06, 0.16]
	$R^2 = .16$ $F(3, 358) = 22.66, p < .001$			$R^2 = .08$ $F(2, 359) = 15.09, p < .001$		

	Effect	*SE*	95% CI
Conditional indirect effects			
Man	0.02	0.01	[0.01, 0.04]
Woman	0.01	0.005	[0.004, 0.02]
Moderated−mediation index	−0.01	0.01	[−0.03, 0.001]

Note. N = 362.

[a] 1 = man, 2 = woman.

* $p < .05$; *** $p < .001$.
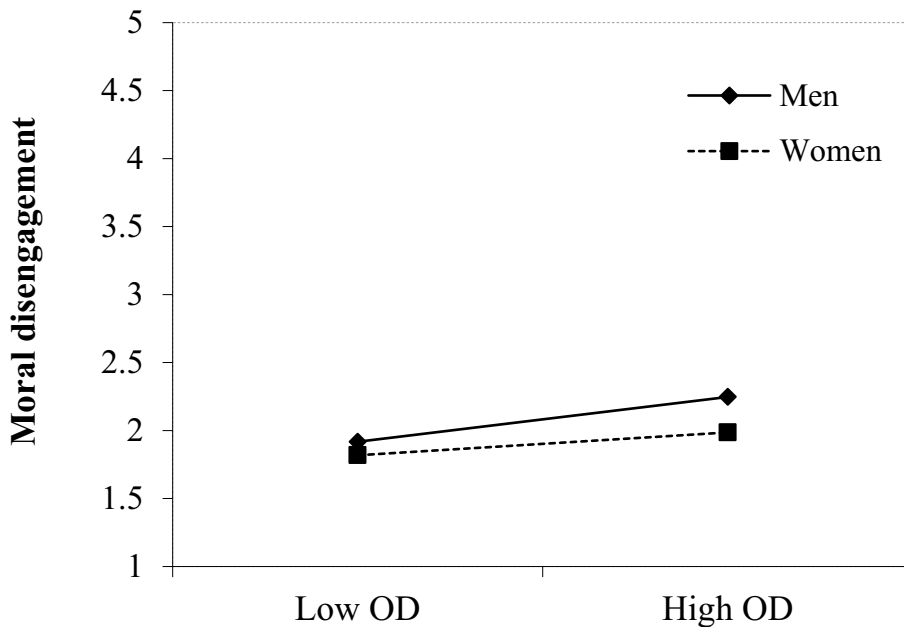
Note. OD = online disinhibition. OD is graphed at −1 *SD* (low) and + 1 *SD* (high).

Fig. 2 Two-Way Interaction Between Online Disinhibition and Gender in Moral Disengagement

current relationship), we observed that the indirect effect of online disinhibition on direct cyberaggression perpetration via moral disengagement was statistically significant ($b = 0.01$, $SE = 0.004$, 95% CI [0.004, 0.02]) and the explained variance increased ($R^2 = .24$). The indirect effect of online disinhibition on cybercontrol perpetration via moral disengagement was not statistically significant ($b = 0.01$, $SE = 0.01$, 95% CI [−0.01, 0.03]). See Table S1 in the online supplementary material (OSF). These findings provide support for Hypothesis 1.

According to Haidt's (2001, 2012) arguments, moral disengagement may be motivated by the desire to justify the engagement in immoral behaviors and maintain a coherent self-concept as a rational and moral agent. Thus, moral disengagement may also function as a post hoc rationalization to eliminate negative emotions after one commits a violent act (e.g., Haidt, 2001, 2012; Shu et al., 2011), such as cyberdating abuse. For this reason, we explored alternative mediation models with moral disengagement as a criterion variable. First, we conducted a mediation analysis to test the indirect effect of online disinhibition on moral disengagement through cybercontrol perpetration; however, it was not significant ($b = -0.001$, $SE = 0.002$, 95% CI [−0.005, 0.004], $R^2 = .21$). In contrast, we found that the mediation model examining the influence of online disinhibition on moral disengagement through direct cyberaggression perpetration was

statistically significant ($b = 0.01$, $SE = 0.01$, 95% CI [0.0003, 0.02], $R^2 = .23$; see Table S2 [OSF]). That is, greater online disinhibition leads to a higher frequency of direct cyberaggression perpetration, which is related to higher levels of moral disengagement. Therefore, our results also supported this alternative theoretical perspective.

Moderation by Gender

As shown in Table 3, we found a statistically significant interaction effect between online disinhibition and gender on moral disengagement. Specifically, high online disinhibition predicted increases in moral disengagement in both men ($b = 0.20$, $SE = 0.04$, $t = 5.64$, $p < .001$, 95% CI [0. 31, 0.27]) and women ($b = 0.07$, $SE = 0.03$, $t = 2.16$, $p = .032$, 95% CI [0.01, 0.14]); however, the effects were stronger among men (see Fig. 2).

Additionally, high moral disengagement predicted a higher frequency of direct cyberaggression perpetration toward a partner. The moderated mediation index was not statistically significant ($b = -0.01$, $SE = 0.01$, 95% CI [−0.03, 0.001]; $R^2 = .80$). In contrast, when we controlled for the effects of covariates (hostile and benevolent sexism, experiences of cyberdating abuse victimization, and current relationship) in the model, the results revealed a significant indirect effect of online disinhibition on the perpetration of

Table 4 Unstandardized Regression Coefficients, Standard Error, and Summary Information for the Moderated–Moderated Mediation Model 11 ("Direct Cyberaggression")

	Moral Disengagement (MD)			Direct Cyberaggression		
	Coeff.	SE	95% CI	Coeff.	SE	95% CI
Constant	1.90***	0.41	[1.09, 2.71]	0.20***	0.06	[0.09, 0.31]
Online disinhibition (OD)	0.02	0.13	[−0.23, 0.27]	0.02	0.01	[−0.002, 0.05]
Gender [a]	−0.19	0.23	[−0.65, 0.27]			
OD x Gender	0.05	0.07	[−0.09, 0.19]			
Hostile sexism (HS)	−0.34	0.24	[−0.84, 0.12]			
OD X HS	0.18*	0.07	[0.03, 0.33]			
Gender X HS	0.21	0.16	[−0.11, 0.52]			
OD X Gender X HS	−0.10*	0.05	[−0.19, −0.002]			
MD				0.11***	0.03	[0.06, 0.16]
	$R^2 = .20$ $F_{(7, 354)} = 12.69, p < .001$			$R^2 = .08$ $F_{(2, 359)} = 15.09, p < .001$		

Conditional indirect effects

	Effect	SE	95% CI
MenLow HS	0.01	0.01	[−0.01, 0.03]
Men – High HS	0.02	0.01	[0.01, 0.04]
Women – Low HS	0.01	0.01	[0.004, 0.02]
Women – High HS	0.01	0.01	[−0.004, 0.02]
Moderated – moderated mediation index	−0.01	0.01	[−0.02, 0.002]

	MD			Direct Cyberaggression		
	Coeff.	SE	95% CI	Coeff.	SE	95% CI
Constant	1.41***	0.33	[0.76, 2.05]	0.20***	0.06	[0.09, 0.31]
Online disinhibition (OD)	0.21*	0.10	[0.02, 0.41]	0.02	0.01	[−0.002, −0.05]
Gender	0.11	0.19	[−0.27, 0.48]			
OD X Gender	−0.07	0.06	[−0.18, 0.04]			
Benevolent sexism (BS)	−0.17	0.13	[−0.43, 0.09]			
OD X BS	0.07	0.04	[−0.01, 0.31]			
Gender X BS	0.13	0.09	[−0.06, 0.31]			
OD X Gender X BS	−0.04	0.03	[−0.09, 0.02]			
MD				0.11***	0.03	[0.06, 0.16]

Table 4 (continued)

	$R^2 = .23$ $F(7, 354) = 14.82, p < .001$		$R^2 = .08$ $F(2, 359) = 15.09, p < .001$	
	Effect	SE	SE	95% CI
Conditional indirect effects				
Men – Low BS	0.01	0.01	0.01	[−0.002, 0.04]
Men – High BS	0.02	0.01	0.01	[0.01, 0.04]
Women – Low BS	0.01	0.04	0.04	[0.001, 0.02]
Women – High BS	0.01	0.01	0.01	[−0.01, 0.02]
Moderated – moderated mediation index	−0.004	0.004	0.004	[−0.01, 0.003]

Note. N = 362.

a 1 = man, 2 = woman.

*$p < .05$; ***$p < .001$.

direct cyberaggression by moral disengagement, moderated by gender. That is, online disinhibition was associated with greater moral disengagement, particularly stronger among men (vs. women), which, in turn, was linked to a higher frequency of direct cyberagression toward a partner. The variables included in the model accounted for 23.86% of the variance in the perpetration of direct cyberaggression. The moderated mediation index was statistically significant ($b = −0.01$, $SE = 0.01$, 95% CI [−0.02 −0.002]; see Table S3), thereby supporting Hypothesis 2.

Moderation by Gender and Sexism

Regarding the moderating role of hostile sexism, as depicted in Table 4, the interaction effects between online disinhibition and gender and between gender and hostile sexism were not significant. In contrast, the results revealed a statistically significant interaction effect between online disinhibition and hostile sexism on moral disengagement. Simple slope analyses indicated that high online disinhibition predicted increases in moral disengagement in both participants with high sexism (+1 SD; $b = 0.19$, $SE = 0.03$, $t = 6.03$, $p < .001$, 95% CI [0.13, 0.25]) and participants with low sexism (−1SD; $b = 0.09$, $SE = 0.03$, $t = 2.58$, $p = .010$, 95% CI [0.02, 0.16]). However, the effects were stronger among participants with high hostile sexism (see Fig. 3, Panel A).
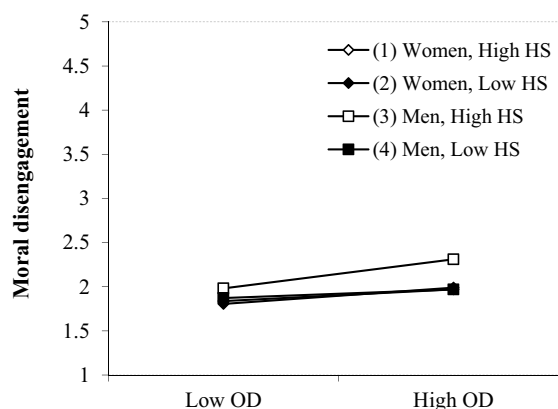
Additionally, the results showed a statistically significant three-way interaction effect between online disinhibition, gender, and hostile sexism. Specifically, high online disinhibition predicted greater moral disengagement in men with high sexism scores ($b = 0.20$, $SE = 0.04$, $t = 5.66$, $p < .001$, 95% CI [0.13, 0.27]), but this was not true in men with low scores ($b = 0.7$, $SE = 0.05$, $t = 1.19$, $p = .23$, 95% CI [−0.04, 0.17]). Among women, online disinhibition predicted higher moral disengagement in those with low levels of sexism ($b = 0.9$, $SE = 0.04$, $t = 2.40$, $p = .02$, 95% CI [0.02, 0.17]), but not in those with high levels ($b = 0.06$, $SE = 0.06$, $t = 0.98$, $p = .33$, 95% CI [−0.06, 0.18]; see Fig. 3, Panel B). Nevertheless, the effect was stronger among men with high hostile sexism. The variables included in the model accounted for 7.75% of the variance in the perpetration of direct cyberaggression. Contrary to our predictions, the moderated-moderated mediation index was not statistically significant ($b = −0.01$, $SE = 0.01$, 95% CI [−0.02, 0.002]).

Regarding the moderating role of benevolent sexism, the results showed no statistically significant interaction effects between the predictor variables. The moderated mediation index was not statistically significant ($b = −0.004$, $SE = 0.004$, 95% CI [−0.01, 0.003]; $R^2 = .08$; see Table 4 for more detail). When we included the covariates (hostile or benevolent sexism, past experiences of cyberdating abuse victimization, and current relationship) in the model, we observed that the moderated-moderated

Panel A
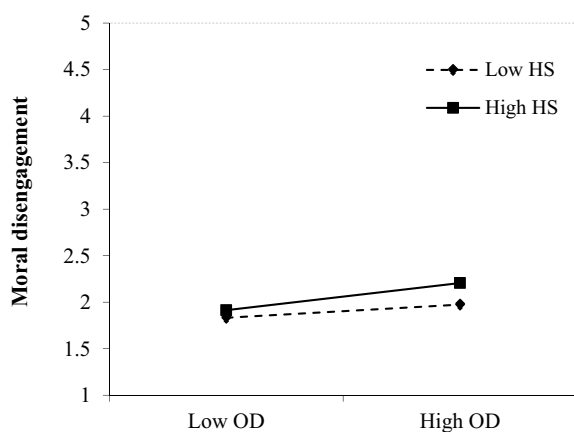
Panel B

Note. OD = online disinhibition; HS = hostile sexism. OD and HS graphed at −1 *SD* (low) and + 1 *SD* (high).

Fig. 3 Two-Way Interaction Between Online Disinhibition and Hostile Sexism (Panel **A**) and Three-Way Interaction Between Online Disinhibition, Gender, and Hostile Sexism (Panel **B**) in Moral Disengagement

mediation index was not statistically significant for both hostile ($b = -0.01$, $SE = 0.01$, 95% CI [−0.02, 0.001], $R^2 = .24$) and sexism ($b = -0.003$, $SE = 0.003$, 95% CI [−0.01, 0.002], $R^2 = .24$; see Table S4). These findings do not support Hypothesis 3.

Moderation by Past Experiences of Cyberdating Abuse Victimization

As shown in Table 5, online disinhibition positively predicted moral disengagement. Furthermore, the results indicated a statistically significant interaction effect between moral disengagement and past experiences of cyberdating abuse victimization in direct cyberaggression perpetration against partners: high moral disengagement was associated with an increase in direct cyberaggression perpetration among participants highly victimized (+ 1 *SD;* $b = 0.17$, $SE = 0.03$, $t = 5.46$, $p < .001$, 95% CI [0.11, 0.23]), but it was not significant among participants less victimized (−1 *SD*; $b = 0.02$, $SE = 0.03$, $t = 0.48$, $p = .629$, 95% CI [−0.05, 0.08]; see Fig. 4, Panel A).

In addition, the results showed a significant indirect effect of online disinhibition on the perpetration of direct cyberaggression by moral disengagement, moderated by past experiences of cyberdating abuse victimization. Specifically, high levels of online disinhibition were associated with greater moral disengagement, which, in turn, was related to increases in direct cyberaggression perpetration against one's partner in participants highly victimized but this was not significant in those with less experiences of victimization.

Similarly, the analysis revealed a statistically significant interaction effect between online disinhibition and past experiences of cyberdating abuse victimization in direct cyberaggression perpetration. As illustrated in Fig. 4 (Panel B), conditional direct effects indicated that, for participants highly victimized ($b = 0.05$, $SE = 0.02$, $t = 2.92$, $p = .004$, 95% CI [0.02, 0.08]), high online disinhibition was associated with more frequent direct cyberaggression perpetration against one's partner, but this association was not significant in participants less victimized ($b = -0.005$, $SE = 0.02$, $t = -0.29$, $p = .770$, 95% CI [−0.04, 0.03]).

The variables included in the model predicted 27.93% of the variance in the perpetration of direct cyberaggression against a partner. The moderated mediation index was statistically significant ($b = 0.02$, $SE = 0.01$, 95% CI [0.01, 0.04]). When we included the covariates in the model (gender, hostile and benevolent sexism, and current relationship), we found that this explained 28.30% of the variance, and the moderated mediation index remained statistically significant ($b = 0.02$, $SE = 0.01$, 95% CI [0.01, 0.03]; see Table S5). These findings supported Hypothesis 4.

Discussion

Although existing literature has shown that online disinhibition and moral disengagement are significant factors contributing to cyberbullying perpetration, there is a gap in the research regarding how these psychological mechanisms may influence cyberdating abuse. In the present study, we aimed to examine the association between online

Table 5 Unstandardized Regression Coefficients, Standard Error, and Summary Information for the Moderated Mediation Model 15 ("Direct Cyberaggression")
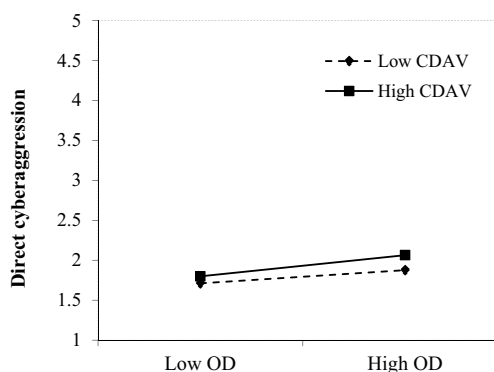
	Moral Disengagement (MD)			Direct Cyberaggression		
	Coeff.	SE	95% CI	Coeff.	SE	95% CI
Constant	1.46***	0.09	[1.29, 1.63]	0.68***	0.14	[0.40, 0.95]
Online disinhibition (OD)	0.16***	0.03	[0.11, 0.21]	−0.05	0.03	[−0.11, 0.02]
MD				−0.11	0.06	[−0.23, 0.02]
CDAV				−0.32***	0.10	[−0.51, −0.13]
OD X CDAV				0.05*	0.02	[0.01, 0.09]
MD X CDAV				0.14***	0.04	[0.06, 0.23]
	$R^2 = .10$			$R^2 = .28$		
	$F(1, 360) = 38.93, p < .001$			$F(5, 356) = 27.59, p < .001$		

	Effect	SE	95% CI
Conditional direct effect			
Low CDAV	−0.005	0.02	[−0.04, 0.03]
High CDAV	0.05	0.02	[0.02, 0.08]
Conditional indirect effect			
Low CDAV	0.003	0.003	[−0.003, 0.01]
High CDAV	0.03	0.01	[0.01, 0.04]
Moderated mediation index	0.02	0.01	[0.01, 0.04]

Note. N = 362.

CDAV Cyberdating abuse victimization, *SE* standard error, *CI* confidence interval.

* *p* < .05; *** *p* < .001.

Panel A
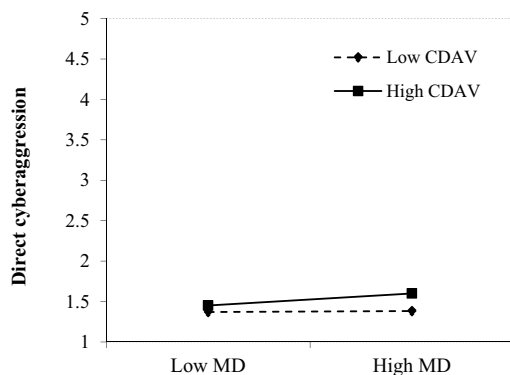
Panel B

Note. MD = moral disengagement; OD = online disinhibition; CDAV = direct cyberaggression victimization. MD, OD, and CDAV are graphed

at −1 *SD* (low) and + 1 *SD* (high).

Fig. 4 Two-Way Interactions Between Moral Disengagement and Past Experiences of Cybervictimization (Panel **A**) and Between Online Disinhibition and Past Experiences of Cybervictimization (Panel **B**) in Direct Cyberaggression Perpetration

disinhibition and the perpetration of direct cyberaggression and cybercontrol against one's partner through moral disengagement, while also exploring the moderating role of gender, sexism, and past experiences of cyberdating abuse victimization on these associations.

The Mediating Role of Moral Disengagement

In support of Hypothesis 1, the results revealed that online disinhibition affected direct cyberaggression perpetration through moral disengagement. Similar to what Wang and Ngai (2020) observed in cyberbullying behavior, high online disinhibition was related to greater moral disengagement, which, in turn, was associated with a more frequent perpetration of direct cyberaggression against partners, but this was not true for the cybercontrol dimension. Cybercontrolling behaviors are often considered manifestations of love and concern towards the partner (Nardi-Rodríguez et al., 2018) and/or a consequence of the use of digital media rather than as an expression of IPV (Belotti et al., 2022). This social perception about cybercontrol may help to understand why moral disengagement mechanisms are not related to its perpetration: Cybercontrolling behaviors might not require the activation of these disinhibitory mechanisms to take place because they seem not to imply a clear violation of moral standards. Therefore, cybercontrol perpetration could be triggered by other psychological processes.

In contrast, direct cyberaggression behaviors against a partner tend to be more easily perceived as a manifestation of violence within a couple's relationship and occur less frequently than cybercontrol behaviors (Caridade et al., 2019; Villorra et al., 2021). In this respect, Suler (2004) stated

that online disinhibition increases the likelihood of engaging in behaviors that would be less likely in a traditional context, such as deliberate direct aggression. Factors such as invisibility and dissociative anonymity may reduce empathy for the victim and the perceived harm that aggressive behavior causes them (Heirman & Walrave, 2008), which, in turn, seem to be related to moral disengagement through detaching from self-sanctions around direct cyberaggression behavior (Runions & Bak, 2015). This is in line with previous findings suggesting that the perpetrator may self-regulate and release feelings of guilt to maintain and justify violence against others (Bandura, 1986, 1999).

Gender and Sexism as Moderators

Building on Bandura's (1986, 1989) social cognitive theory, we testified some intrapersonal factors related to modulating the influence of online disinhibition and moral disengagement in direct cyberaggression perpetration against a partner. In support of Hypothesis 2, our results showed that gender moderated the indirect effect of online disinhibition on direct cyberaggression via moral disengagement. Specifically, higher online disinhibition was primarily associated with increased moral disengagement in men (vs. women), which, in turn, was related to a higher frequency of direct cyberaggression perpetration against a partner. This result is consistent with previous findings indicating that young men experience greater toxic online disinhibition (e.g., Wang & Ngai, 2020; Wang et al., 2021) and greater moral disengagement (e.g.,Navas et al., 2021; Sánchez-Jiménez & Muñoz-Fernández, 2021) than young women. Although online disinhibition and moral disengagement seem to associate

with the perpetration of direct cyberaggression against a partner in a similar way as in the field of cyberbullying, it is important to note that cyberdating abuse is also influenced by gender social norms. In heterosexual romantic relationships, there is often a complex power dynamic derived from the patriarchal structure that is relevant when examining cyberdating abuse. Our findings therefore are in line with empirical work supporting the assertion that cyberdating abuse is gender asymmetric: Direct cyberaggression behaviors used as a tool of control and power in the relationship occur to a greater extent in men (Reed et al., 2021), and not because they are biologically men but because of gender socialization and their identity with the meaning of masculinity (Stosny, 1995). In a patriarchal society, women have a lower status than men and, consequently, some men may consider that women deserve violent treatment in certain situations (Expósito et al., 1998), for example, when women question men's power in the relationship. In this sense, men could easily conform and adapt their judgment to their moral norms cognitively to justify direct cyberaggression against a partner, and even more so when the transgressive behavior is encouraged by online disinhibiting factors blurring moral boundaries (Bandura, 2002).

Regarding the moderating role of sexism, the results indicated that gender and hostile sexism jointly moderat the effect of online disinhibition on moral disengagement. Specifically, higher online disinhibition was primarily associated with increased moral disengagement among men with high hostile sexism (vs. low hostile sexism). This result reinforces the previous literature by demonstrating the existence of a positive association between hostile sexist attitudes and moral disengagement (Rollero & De Piccoli, 2020; Sánchez-Jiménez & Muñoz-Fernández, 2021). Men, and individuals higher in hostile sexism, seem to more readily adjust their moral norms to justify various transgressive and immoral behaviors, such as IPV against women, and mitigate potential feelings of responsibility and guilt (Navas et al., 2021; Page & Pina, 2015). However, contrary to our predictions, the results revealed that hostile sexism did not interact with gender to moderate the indirect effect of online disinhibition on direct cyberaggression via moral disengagement. This could be due to the fact that the observed third-order interaction between the predictors did not show a sufficient effect size, and larger samples are needed to test this model. More research is needed to clarify this issue.

On the other hand, our results indicated that benevolent sexism did not interact with online disinhibition and gender to moderate their effects on moral disengagement and, consequently, on direct cyberaggression. This finding could be explained by the nature of benevolent sexism: whereas hostile sexism reflects antagonism and hostility toward women as a discriminated group, benevolent sexism adopts a positive view of respect, affection, and protection towards them

(Glick & Fiske, 1996). In this sense, research has shown that hostile sexism is consistently associated with the justification and perpetration of explicit manifestations of IPV, whereas benevolent sexism may often act as a mitigator of this link (Glick & Fiske, 1996; Moya, 2003). Specific to cyberdating abuse, some studies have found that hostile sexism, but not benevolent sexism, in men predicts the perpetration of direct cyberaggression against one's partner (e.g., insults, threats, or humiliation through digital media; Martínez-Pecino & Durán, 2019; Rodríguez-Domínguez et al., 2018). Thus, it makes sense that high online disinhibition leads to greater activation of moral disengagement mechanisms primarily in people with strong hostile sexist beliefs, but not in people with high benevolent sexism. These individuals can more readily adjust their cognitive moral norms to justify immoral behaviors such as violence. However, more research is needed to support these assumptions.

Past Experiences of Cyberdating Abuse Victimization as Moderator

In support of Hypothesis 4, our data showed that past experiences of cyberdating abuse moderated the indirect effect of online disinhibition on direct cyberaggression through moral disengagement. That is, high levels of online disinhibition were associated with greater moral disengagement, which, in turn, was related to increases in direct cyberaggression perpetration against one's partner in participants highly victimized in the past, but not in those less victimized. These findings are in line with Cuadrado-Gordillo and Fernández-Antelo (2019) work showing that being highly victimized is related to greater levels of moral disengagement. More specifically, these results suggest the use of moral disengagement could lead to the internalization of aggression as an appropriate strategy to resolve conflicts among cyberdating abuse victims, similar to what has been observed in other contexts of online violence (e.g., online hate speech; Wachs et al., 2022). Moral disengagement's self-regulatory mechanisms may play a relevant role in the people's acceptance and justification of direct cyberaggression perpetration when one has also been a victim of cyberdating abuse. That is, they may approve of such violence against themselves and perceive it as a problem-solving tool within their romantic relationships, which, in turn, could favor a culture of shared cyberabuse between partners (Wong-Lo & Bullock, 2014).

Additionally, our results revealed that past experiences of cyberdating abuse victimization also moderated the direct effect of online disinhibition on direct cyberaggression. In particular, elevated levels of online disinhibition predicted more direct cyberaggression against partners among participants who had more victimization experiences, but not in those less victimized. These findings are consistent with Wong-Lo and Bullock (2014), who observed that online

disinhibition increases the likelihood that cybervictims will assume the role of cyberaggressors, regardless of gender. Moreover, these findings support Moore (2015), who suggested that online disinhibition does not homogeneously lead to higher levels of cyberbullying, but may under specific circumstances. In the online context, individuals have tools at their disposal that increase the likelihood of engaging in reactive IPV. For example, the feeling of being protected behind a screen or the physical distance between the offender and the victim may encourage victims to adopt maladaptive coping strategies such as the perpetration of reactive cyberaggression (Stonard et al., 2017). In this way, technologies may be creating a false sense of empowerment for victims that, rather than empowering them to seek solutions to the violent situation, could lead to unhealthy strategies that further perpetuate and normalize cyberdating abuse among partners (Alvarez, 2012).

Limitations and Future Directions

Although our research expands knowledge in the cyberdating abuse field, some limitations should be pointed out. First, we conducted a correlational study with a cross-sectional design in our research, so we cannot establish causal relationships between the study variables. Future researchers could implement longitudinal methods or experimental design to achieve more controllability over the results and obtain interpretations of causality. Likewise, we encourage future researchers to use other innovative approaches, such as dyadic research designs involving both couple partners. Second, we assessed all constructs using self-report measures. As is typically the case in surveys on sensitive topics and undesirable behaviors, responses may have been subject to social desirability and recall bias (Deans & Bhogal, 2019). Third, we did not use a measure of moral disengagement adapted to the context and the behavior under study (i.e., cyberdating abuse); therefore, we did not assess online disinhibition and moral disengagement in the same behavioral setting. Our results should be replicated using a context-appropriate moral disengagement scale to remedy this limitation. Nevertheless, our results confirmed that online disinhibition positively influences the direct cyberaggression perpetration via moral disengagement, which suggests that the moral disengagement concept may transcend the context to which it is applied.

Fourth, we selected the participants by nonprobability snowball sampling and we established several inclusion criteria (i.e., having Spanish nationality, being between 18 and 35 years of age, having a heterosexual orientation, and having been in a past or current romantic relationship); therefore, our results are not generalizable to the entire population. Other researchers should corroborate our findings using

random sampling and collecting heterogeneous samples to examine the potential influence of variables such as age, nationality, cultural values, sexual orientation, or gender identity (cisgender vs. non-cisgender). Fifth, respondents could think about all their romantic relationships when we addressed cyberdating abuse; we were unable to test whether the cyberdating abuse took place with the same partner or a different one. Therefore, conclusions derived from our work need to be taken with caution. Similarly, we did not consider the potential effect of daily internet use in our analyses. Given that previous studies suggest that extensive internet exposure time increases the probability of suffering and committing cyberdating abuse (e.g., Caridade & Braga, 2020), researchers should replicate our results by controlling for this variable.

Practice Implications

The findings of the present research have some practice implications. First, our findings could encourage clinical psychologists to detect and dismantle the cognitive and behavioral processes that contribute to the justification and normalization of cyberdating abuse, considering the peculiarities of this type of violence, as well as the personal characteristics and relationship dynamics. Second, psychology professionals could use our work to develop psychoeducational programs of cyberdating abuse prevention and intervention aimed at respectively preventing and mitigating the negative effects of online disinhibition and moral disengagement mechanisms in romantic relationships. Third, our results constitute a warning regarding the potential detriments of misuse of digital media and online disinhibition among young adults, stressing the need to lead practical efforts toward the responsible use of digital media and the promotion of healthy and quality relationships. Fourth, this work supports the gender inequality-persistence framing and the need to consider gender and sexist attitudes in the intervention and clinical practice of cyberdating abuse. The social perception that cyberdating abuse is gender symmetrical has been increasing in recent years, partially because empirical research has noted that this violence often occurs bidirectionally (e.g., Villora et al., 2021). However, the power dynamics that are established within violent relationships are not outside gendered social norms. Therefore, professionals working in different cyberdating abuse fields (intervention, clinical, research, etc.) need to join forces, approaching the problem from a gender perspective. Overall, we hope that this study can account, at least minimally, for the need to continue researching the variables associated with the cyberdating abuse perpetration to promote social policies and specific resources for the prevention and effective intervention of this problem.

Conclusion

Cyberdating abuse is increasingly observed in romantic relationships, and there is a risk of such abusive behaviors becoming normalized. A key contribution of this research is to highlight the psychological mechanisms of online disinhibition and moral disengagement in the perpetration of direct cyberaggression against partners within young heterosexual couples. These findings suggest that gender and past experiences of victimization may be essential factors for understanding who may be most vulnerable to online disinhibition and moral disengagement activation in the perpetration of direct cyberaggression against partners. Our research may help researchers develop specific interventions in violent heterosexual relationships that consider the influence of these factors and explore new avenues of work that delve deeper into the psychosocial mechanisms that arise from violent relationships and the coping strategies that could constructively confront cyberdating abuse.

Data Availability This study' design and its analyses were not pre-registered. Supplementary materials, data, and scripts are publicly available and can be accessed at [OSF].

Compliance with Ethical Standards

Ethics Approval The study received ethical approval from the University of Granada (approval number: 1050/CEIH/2020).

Patient Consent Statement Informed consent was obtained from all participants at the start of the survey.

Conflict of Interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Alvarez, A. R. (2012). "IH8U": Confronting cyberbullying and exploring the use of cybertools in teen dating relationships. *Journal of Clinical Psychology, 68*(11), 1205–1215. https://doi.org/10.1002/jclp.21920

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist, 55*(5), 469–480. https://doi.org/10.1037/0003-066X.55.5.469

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice Hall.

Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist, 44*(9), 1175–1184. https://doi.org/10.1037/0003-066X.44.9.1175

Bandura, A. (1999). Moral disengagement in the preparation of inhumanities. *Personal and Social Psychology Review, 3*(3), 193–209.

Bandura, A. (2002). Selective moral disengagement in the exercise of moral agency. *Journal of Moral Education, 31*(2), 101–119. https://doi.org/10.1080/0305724022014322

Bandura, A. (2016). *Moral disengagement: How people do harm and live with themselves*. Worth Publishers.

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology, 71*(2), 364–374. https://doi.org/10.1037/00223514.71.2.364

Barlett, C. P., & Helmstetter, K. M. (2018). Longitudinal relations between early online disinhibition and anonymity perceptions on later cyberbullying perpetration: A theoretical test on youth. *Psychology of Popular Media Culture, 7*(4), 561–571. https://doi.org/10.1037/ppm0000149

Belotti, F., Ieracitano, F., Donato, S., & Comunello, F. (2022). Towards 'romantic media ideologies': Digital dating abuse seen through the lens of social media and/or dating in teenage narratives. *The Communication Review, 25*(1), 1–24. https://doi.org/10.1080/10714421.2022.2033576

Borrajo, E., & Gámez-Guadix, M. (2016). Abuso" online" en el noviazgo: Relación con depresión, ansiedad y ajuste diádico ["Online" dating abuse: Relationship with depression, anxiety and dyadic adjustment]. *Behavioral Psychology, 24*(2), 221–235. http://hdl.handle.net/10486/679217

Borrajo, E., Gámez-Guadix, M., & Calvete, E. (2015a). Cyber dating abuse: Prevalence, context, and relationship with offline dating aggression. *Psychological Reports, 116*(2), 565–585. https://doi.org/10.2466/21.16.PR0.116k22w4

Borrajo, E., Gámez-Guadix, M., Pereda, N., & Calvete, E. (2015b). The development and validation of the cyber dating abuse

questionnaire among young couples. *Computers in Human Behavior, 48*, 358–365. https://doi.org/10.1016/j.chb.2015.01.063

Branson, M., & March, E. (2021). Dangerous dating in the digital age: Jealousy, hostility, narcissism, and psychopathy as predictors of cyber dating abuse. *Computers in Human Behavior, 119*, 106711. https://doi.org/10.1016/j.chb.2021.106711

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.

Butler, L. C., Fissel, E. R., Gildea, B., & Fisher, B. S. (2023). Understanding intimate partner cyber abuse across partnership categories based on gender identity and sexual orientation. *Victims & Offenders, 18*(1), 77–100. https://doi.org/10.1080/15564886.2022.2139032

Caprara, G. V., Tisak, M. S., Alessandri, G., Fontaine, R. G., Fida, R., & Paciello, M. (2014). The contribution of moral disengagement in mediating individual tendencies toward aggression and violence. *Developmental Psychology, 50*(1), 71–85. https://doi.org/10.1037/a0034488

Caridade, S., & Braga, T. (2020). Youth cyber dating abuse: A meta-analysis of risk and protective factors. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace, 14*(3), Article 2. https://doi.org/10.5817/CP20203-2

Caridade, S., Braga, T., & Borrajo, E. (2019). Cyber dating abuse (CDA): Evidence from a systematic review. *Aggression and Violent Behavior, 48*, 152–168. https://doi.org/10.1016/j.avb.2019.08.018

Cheung, C. M. K., Wong, R. Y. M., & Chan, T. K. H. (2020). Online disinhibition: Conceptualization, measurement, and implications for online deviant behavior. *Industrial Management & Data Systems, 121*(1), 48–64. https://doi.org/10.1108/IMDS-08-2020-0509

Cuadrado-Gordillo, I., & Fernández-Antelo, I. (2019). Analysis of moral disengagement as a modulating factor in adolescents' perception of cyberbullying. *Frontiers in Psychology, 10*, 12–22. https://doi.org/10.3389/fpsyg.2019.01222

D'Errico, F., & Paciello, M. (2018). Online moral disengagement and hostile emotions in discussions on hosting immigrants. *Internet Research, 28*(5), 1313–1335. https://doi.org/10.1108/IntR-03-2017-0119

Deans, H., & Bhogal, M. S. (2019). Perpetrating cyber dating abuse: A brief report on the role of aggression, romantic jealousy, and gender. *Current Psychology, 38*, 1077–1082. https://doi.org/10.1007/s12144-017-9715-4

Donoso-Vázquez, T., Hurtado, M. J. R., & Baños, R. V. (2018). La adolescencia ante la violencia de género 2.0: Concepciones, conductas y experiencias. *Educación xx1, 21*(1), 109–133. http://www.redalyc.org/articulo.oa?id=70653466006

Expósito, F., Moya, M. C., & Glick, P. (1998). Sexismo ambivalente: Medición y correlatos. *Revista De Psicología Social, 55*(2), 893–905. https://doi.org/10.1174/021347498760350641

Falla, D., Ortega-Ruiz, R., Runions, K., & Romera, E. M. (2021). Moral disengagement strategies in online and offline bullying. *Psychosocial Intervention, 30*(2), 85–93. https://doi.org/10.5093/pi2020a21

Fernández-González, L., Calvete, E., & Orue, I. (2020). Adolescent dating violence stability and mutuality: A 4-year longitudinal study. *Journal of Interpersonal Violence, 35*(4), 2012–2032. https://doi.org/10.1177/0886260517699953

Gámez-Guadix, M., Borrajo, E., & Almendros, C. (2016). Risky online behaviors among adolescents: Longitudinal relations among problematic Internet use, cyberbullying perpetration, and meeting strangers online. *Journal of Behavioral Addictions, 5*(1), 100–107. https://doi.org/10.1556/2006.5.2016.013

Gámez-Guadix, M., Borrajo, E., and Calvete, E. (2018). Partner abuse, control and violence through internet and smartphones:

Characteristics, evaluation and prevention. Papeles del Psicólogo, 39(3), 218–227. https://doi.org/10.23923/pap.psicol2018.2874.

Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology, 70*(4), 491–512. https://doi.org/10.1037/0022-3514.70.3.491

Glick, P., & Fiske, S. T. (2001). An ambivalent alliance: Hostile and benevolent sexism as complementary justifications for gender inequality. *American Psychologist, 56*(2), 109–118. https://doi.org/10.1037/0003-066X.56.2.109

Gravetter, F. J., & Wallnau, L. B. (2014). Introduction to the t statistic. *Essentials of statistics for the behavioral sciences* (8th ed., p. 252). Wadsworth.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*(4), 814–834. https://doi.org/10.1037/0033-295X.108.4.814

Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. Pantheon.

Hancock, K., Keast, H., & Ellis, W. (2017). The impact of cyber dating abuse on self-esteem: The mediating role of emotional distress. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 11(2). https://doi.org/10.5817/CP2017-2-2

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. The Guilford Press.

Heirman, W., & Walrave, M. (2008). Assessing concerns and issues about the mediation of technology in cyberbullying. *Cyberpsychology, 2*(2), 1–12.

Hinduja, S., & Patchin, J. W. (2021). Digital dating abuse among a national sample of US youth. *Journal of Interpersonal Violence, 36*(23–24), 11088–11108. https://doi.org/10.1177/08862605198973

Karasavva, V., Swanek, J., Smodis, A., & Forth, A. (2022). From myth to reality: Sexual image abuse myth acceptance, the dark tetrad, and non-consensual intimate image dissemination proclivity. *Journal of Sexual Aggression, 29*(1), 1–17. https://doi.org/10.1080/13552600.2022.2032430

Klement, K. R., Sagarin, B. J., & Skowronski, J. J. (2019). Accusers lie and other myths: Rape myth acceptance predicts judgments made about accusers and accused perpetrators in a rape case. *Sex Roles, 81*(1–2), 16–33. https://doi.org/10.1007/s11199-018-0950-4

Linares, R., Aranda, M., García-Domingo, M., Amezcua, T., Fuentes, V., & Moreno-Padilla, M. (2021). Cyber-dating abuse in young adult couples: Relations with sexist attitudes and violence justification, smartphone usage and impulsivity. *Plosone, 16*(6), e0253180. https://doi.org/10.1371/journal.pone.0253180

Longden, L. (2014). Online disinhibition: Heightened potential for professional conduct issues among clergy in the Church of England? *Journal of Beliefs & Values, 35*(1), 1–9. https://doi.org/10.1080/13617672.2014.884844

Lowry, P. B., Zhang, J., & Wu, T. (2017). Nature or nurture? A meta-analysis of the factors that maximize the prediction of digital piracy by using social cognitive theory as a framework. *Computers in Human Behavior, 68*, 104–120. https://doi.org/10.1016/j.chb.2016.11.015

Lu, Y., Van Ouytsel, J. V., Walrave, M., Ponnet, K., & Temple, J. R. (2018). Cross-sectional and temporal associations between cyber dating abuse victimization and mental health and substance use outcomes. *Journal of Adolescence, 65*, 1–5. https://doi.org/10.1016/j.adolescence.2018.02.009

Mahoney, J. C., Farrell, D. M., & Murphy, C. M. (2022). Prevalence and predictors of cyber psychological abuse among adults. *Journal of Family Violence, 37*(1), 151–163. https://doi.org/10.1007/s10896-021-00317-x

Martín-Fernández, M., Gracia, E., Marco, M., Vargas, V., Santirso, F. A., & Lila, M. (2018). La medida de la aceptabilidad de la

violencia de pareja contra la mujer: Desarrollo y validación de la escala A-IPVAW. *The European Journal of Psychology Applied to Legal Context, 10*(1), 26–34. https://doi.org/10.5093/ejpalc2018a3

Martínez-Pecino, R., & Durán, M. (2019). I love you but I cyberbully you: The role of hostile sexism. *Journal of Interpersonal Violence, 34*(4), 812–825. https://doi.org/10.1177/0886260516645817

Moore, C. (2015). Moral disengagement. *Current Opinion in Psychology, 6*, 199–204. https://doi.org/10.1016/j.copsyc.2015.07.018

Moya, M. (2003). El análisis psicosocial de género. In J. F. Morales y C. Huici (Eds.), *Estudios de psicología social* (pp.175–222). UNED.

Naquin, C. E., Kurtzberg, T. R., & Belkin, L. Y. (2010). The finer points of lying online: E-mail versus pen and paper. *Journal of Applied Psychology, 95*(2), 387–394. https://doi.org/10.1037/a0018627

Nardi-Rodríguez, A., Pastor-Mira, M. A., López-Roig, S., & Ferrer-Pérez, V. (2018). Identifying beliefs behind boys' use of mobile phones to monitor girlfriends and girls' acceptance: A reasoned-action approach. *Journal of Youth Studies, 21*(7), 922–939. https://doi.org/10.1080/13676261.2017.1422600

Navas, M. P., Maneiro, L., Cutrín, O., Gómez-Fraguela, J. A., & Sobral, J. (2021). Sexism, moral disengagement, and dark triad traits on perpetrators of sexual violence against women and community men. *Sexual Abuse, 34*(7), 857–884. https://doi.org/10.1177/10790632211051689

Oleszkowicz, A., & Misztela, A. (2015). How do young poles perceive their adulthood? *Journal of Adolescent Research, 30*(6), 683–709. https://doi.org/10.1177/0743558415569727

Paciello, M., Fida, R., Cerniglia, L., Tramontano, C., & Cole, E. (2013). High cost helping scenario: The role of empathy, prosocial reasoning and moral disengagement on helping behavior. *Personality and Individual Differences, 55*(1), 3–7. https://doi.org/10.1016/j.paid.2012.11.004

Paciello, M., Tramontano, C., Nocentini, A., Fida, R., & Menesini, E. (2020). The role of traditional and online moral disengagement on cyberbullying: Do externalising problems make any difference? *Computers in Human Behavior, 103*, 190–198. https://doi.org/10.1016/j.chb.2019.09.024

Page, T. E., & Pina, A. (2015). Moral disengagement as a self-regulatory process in sexual harassment perpetration at work: A preliminary conceptualization. *Aggression and Violent Behavior, 21*, 73–84. https://doi.org/10.1016/j.avb.2015.01.004

Reed, L. A., Ward, L. M., Tolman, R. M., Lippman, J. R., & Seabrook, R. C. (2021). The association between stereotypical gender and dating beliefs and digital dating abuse perpetration in adolescent dating relationships. *Journal of Interpersonal Violence, 36*(9–10), NP5561–NP5585. https://doi.org/10.1177/0886260518801933

Rodríguez-Domínguez, C., Durán Segura, M., & Martínez Pecino, R. (2018). Ciberagresores en el noviazgo adolescente y su relación con la violencia psicológica, el sexismo y los celos. *Health and Addictions, 18*(1), 17–27. https://doi.org/10.21134/haaj.v18i1.329

Rollero, C., & De Piccoli, N. (2020). Myths about intimate partner violence and moral disengagement: An analysis of sociocultural dimensions sustaining violence against women. *International Journal of Environmental Research and Public Health, 17*(21), 8139. https://doi.org/10.3390/ijerph17218139

Rubio-Garay, F., Amor, P. J., & Carrasco, M. A. (2017). Dimensionality and psychometric properties of the Spanish version of the Mechanisms of Moral Disengagement Scale (MMDS-S). *Revista De Psicopatología y Psicología Clínica, 22*(1), 43–54. https://doi.org/10.5944/rppc.vol.22.num.1.2017.16014

Runions, K. C., & Bak, M. (2015). Online moral disengagement, cyberbullying, and cyber-aggression. *Cyberpsychology, Behavior, and Social Networking, 18*(7), 400–405. https://doi.org/10.1089/cyber.2014.0670

Runions, K., Shapka, J. D., Dooley, J., & Modecki, K. (2013). Cyber-aggression and victimization and social information processing: Integrating the medium and the message. *Psychology of Violence, 3*(1), 9–26. https://doi.org/10.1037/a0030511

Sánchez-Hernández, M. D., Herrera, M. C., & Expósito, F. (2024). Perception of Cyberdating Abuse from the Victims' Perspective: Effect of the Type of Suffered Behavior and Gender. *Current Psychology,* 1–14. https://doi.org/10.1007/s12144-024-05985-8

Sánchez-Hernández, M. D., Herrera, M. C., & Expósito, F. (2023). Does the digital environment evoke anxiety cycles in romantic relationships? The Roles of social–interpersonal and individual factors in cyberdating abuse perpetration. *International Journal of Human-Computer Interaction,* 1–18. https://doi.org/10.1080/10447318.2023.2204275

Sánchez-Hernández, M. D., Herrera-Enríquez, M. C., & Expósito, F. (2020). Controlling behaviors in couple relationships in the digital age: Acceptability of gender violence, sexism, and myths about romantic love. *Psychosocial Intervention, 29*(2), 67–81. https://doi.org/10.5093/pi2020a1

Sánchez-Jiménez, V., & Muñoz-Fernández, N. (2021). When are sexist attitudes risk factors for dating aggression? The role of moral disengagement in Spanish adolescents. *International Journal of Environmental Research and Public Health, 18*(4), 1947. https://doi.org/10.3390/ijerph18041947

Sanfilippo, M., Yang, S., & Fichman, P. (2017). Trolling here, there, and everywhere: Perceptions of trolling behaviors in context. *Journal of the Association for Information Science and Technology, 68*(10), 2313–2327. https://doi.org/10.1002/asi.23902

Schoemann, A. M., Boulton, A. J., & Short, S. D. (2017). Determining power and sample size for simple and complex mediation models. *Social Psychological and Personality Science, 8*(2), 379–386. https://doi.org/10.1177/1948550617750734

Shrestha, N. (2020). Detecting multicollinearity in regression analysis. *American Journal of Applied Mathematics and Statistics, 8*, 39–42. https://doi.org/10.12691/ajams-8-2-1

Shu, L. L., Gino, F., & Bazerman, M. H. (2011). Dishonest deed, clear conscience: When cheating leads to moral disengagement and motivated forgetting. *Personality and Social Psychology Bulletin, 37*(3), 330–349. https://doi.org/10.1177/0146167211398138

Smith, K., Cénat, J. M., Lapierre, A., Dion, J., Hébert, M., & Côté, K. (2018). Cyber dating violence: Prevalence and correlates among high schools students from small urban areas in Quebec. *Journal of Affective Disorders, 234*, 220–223. https://doi.org/10.1016/j.jad.2018.02.043

Smith, P. K. (2019). Research on cyberbullying: Strengths and limitations. In H. Vandebosch & L. Green (Eds.), *Narratives in Research and Interventions on Cyberbullying among Young People* (pp. 9–27). Springer.

Soriano-Ayala, E., Cala, V. C., & Orpinas, P. (2023). Prevalence and predictors of perpetration of cyberviolence against a dating partner: A cross-cultural study with Moroccan and Spanish youth. *Journal of Interpersonal Violence, 38*(3–4), 4366–4389. https://doi.org/10.1177/08862605221115111

Soto, A. M., & Ibabe, I. (2022). Recommended instruments for analyzing cyber dating violence: A systematic review. *The Spanish Journal of Psychology, 25*, e4. https://doi.org/10.1017/SJP.2021.50

Statista. (2022). *Redes sociales: porcentaje de usuarios por edad en España en 2022 [Social networks: Percentage of users by age in Spain in 2022].* https://es.statista.com/estadisticas/1260093/redes-sociales-porcentaje-de-usuarios-por-edad-en-espana/

Stonard, K. E. (2021). The prevalence and overlap of technology-assisted and offline adolescent dating violence. *Current Psychology, 40*(3), 1056–1070. https://doi.org/10.1007/s12144-018-0023-4

Stonard, K. E., Bowen, E., Walker, K., & Price, S. A. (2017). "They'll always find a way to get to you": Technology use in adolescent

romantic relationships and its role in dating violence and abuse. *Journal of Interpersonal Violence, 32*(14), 2083–2117. https://doi.org/10.1177/0886260515590787

Stosny, S. (1995). *Treating attachment abuse: A compassionate approach*. Springer.

Suler, J. (2004). The online disinhibition effect. *Cyberpsychology & Behavior, 7*(3), 321–326. https://doi.org/10.1089/1094931041291295

Templeton, G.F. (2011) A two-step approach for transforming continuous variables to normal: Implications and recommendations for IS research. *Communications of the Association for Information Systems, 28*(4), 41–58. https://doi.org/10.17705/1CAIS.02804

Van Ouytsel, J., Ponnet, K., & Walrave, M. (2017). The associations of adolescents' dating violence victimization, well-being and engagement in risk behaviors. *Journal of Adolescence, 55*, 66–71. https://doi.org/10.1016/j.adolescence.2016.12.005

Villora, B., Navarro, R., & Yubero, S. (2021). The role of social-interpersonal and cognitive-individual factors in cyber dating victimization and perpetration: Comparing the direct, control, and combined forms of abuse. *Journal of Interpersonal Violence, 36*(17–18), 8559–8584. https://doi.org/10.1177/0886260519851172

Villora, B., Yubero, S., & Navarro, R. (2019a). Associations between feminine gender norms and cyber dating abuse in female adults. *Behavioral Sciences, 9*(4), 35. https://doi.org/10.3390/bs9040035

Villora, B., Yubero, S., & Navarro, R. (2019b). Cyber dating abuse and masculine gender norms in a sample of male adults. *Future Internet, 11*(4), 84. https://doi.org/10.3390/fi11040084

Wachs, S., Bilz, L., Wettstein, A., Wright, M. F., Krause, N., Ballaschk, C., & Kansok-Dusche, J. (2022). The online hate speech cycle of violence: Moderating effects of moral disengagement and empathy in the victim-to-perpetrator relationship. *Cyberpsychology, Behavior, and Social Networking, 25*(4), 223–229. https://doi.org/10.1089/cyber.2021.0159

Wang, L., & Ngai, S. S. Y. (2020). The effects of anonymity, invisibility, asynchrony, and moral disengagement on cyberbullying perpetration among school-aged children in China. *Children and Youth Services Review, 119*, 105–613. https://doi.org/10.1016/j.childyouth.2020.105613

Wang, X., Qiao, Y., Li, W., & Dong, W. (2021). How is online disinhibition related to adolescents' cyberbullying perpetration? Empathy and gender as moderators. *The Journal of Early Adolescence, 42*(5), 704–732. https://doi.org/10.1177/02724316211064515

Wang, X., Wang, W., Qiao, Y., Gao, L., Yang, J., & Wang, P. (2022). Parental phubbing and adolescents' cyberbullying perpetration: A moderated mediation model of moral disengagement and online disinhibition. *Journal of Interpersonal Violence, 37*(7–8), NP5344–NP5366.

Wong-Lo, M., & Bullock, L. M. (2014). Digital metamorphosis: Examination of the bystander culture in cyberbullying. *Aggression and Violent Behavior, 19*(4), 418–422. https://doi.org/10.1016/j.avb.2014.06.007

Wright, M. F. (2016). Cyber victimization on college campuses: Longitudinal associations with suicidal ideation, depression, and anxiety. *Criminal Justice Review, 41*(2), 190–203. https://doi.org/10.1177/0734016816634785

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.