Check for updates

Scandinavian Journal of Psychology

Scandinavian Journal of Psychology

EMPIRICAL ARTICLE OPEN ACCESS

What Makes Adolescents Defend Others Against Cyberhate? Insights From the Normative Social Behavior, Social Learning, and Planned Behavior Theories

Olga Jiménez-Díaz[1] | Esperanza Espino[2] | Rosario Del Rey[3] | Joaquín A. Mora-Merchán[3]

[1]Department of Developmental and Educational Psychology, Faculty of Psychology, Universidad de Sevilla, Seville, Spain | [2]Departament of Psychology, Faculty of Psychology and Education, Universidad Loyola Andalucía, Seville, Spain | [3]Department of Developmental and Educational Psychology, Faculty of Educational Science, Universidad de Sevilla, Seville, Spain

Correspondence: Rosario Del Rey (delrey@us.es)

Received: 18 February 2026 | **Revised:** 19 June 2026 | **Accepted:** 21 July 2026

Keywords: bystanders | cyberhate | defending behavior | SLT | TNSB | TPB

ABSTRACT

Most adolescents experience cyberhate primarily as bystanders. Understanding their behavior, particularly their defending behaviors, is crucial to mitigate the phenomenon and its negative effects. While prior studies have explored factors influencing these responses, further research is needed to understand how these factors interact within robust theoretical frameworks. To address this gap, this study examined adolescents' defending behaviors to cyberhate through three psychological theories: (1) the Theory of Normative Social Behavior (TNSB; social norms); (2) Social Learning Theory (SLT; social norms, self-efficacy, prior victimization and perpetration experiences, and motivators such as the need for online popularity); and (3) the Theory of Planned Behavior (TPB; attitude towards helping, perceived control, subjective norms and intention to help). A total of 2539 Spanish secondary school students (49.1% boys, 49.2% girls, and 1.7% others) aged 11–18 years ($M = 14.07$, SD = 1.39) participated filling a self-reported survey. Three structural equation models were estimated to test the specific variables of each theoretical framework while controlling for gender and age. Results revealed that all three models significantly explained defensive behaviors. Across the models, explained variance reached up to 50.8%. The most relevant factors were subjective norms, injunctive norms from friends, family, and teachers, and perceived self-efficacy. Girls and younger students were more likely to react defensively. This study represents a clear advance in the understanding of defensive behavior in the face of cyberhate, both in a theoretical and practical framework and thus contributes to optimizing the holistic approach to this emerging phenomenon in virtual contexts.

1 | Introduction

The research on cyberviolence among adolescents has shown a clear shift in focus, moving away from a deficit-oriented perspective that emphasizes risk factors and towards a more positive approach focusing protective factors. These protective factors refer to behaviors adolescents can adopt to cope with their involvement or mitigate its negative effects (e.g., Ferreira et al. 2020; Rudnicki et al. 2023). Indeed, this has been particularly evident in phenomena such as cyberbullying—understood as intentional peer violence characterized by a power imbalance and potential repetition (Domínguez-Hernández et al. 2018; Dublin City University 2024; Rudnicki et al. 2023) and, more recently, in non-consensual sexting (Barrense-Dias et al. 2024; Maes et al. 2024). However, progress in this regard remains a considerable challenge in emerging forms of cyberviolence such as cyberhate (Wachs, Valido, et al. 2023).

Cyberhate is defined as any form of online expression, typically through social media, including text messages, media content,

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2026 The Author(s). *Scandinavian Journal of Psychology* published by Scandinavian Psychological Associations and John Wiley & Sons Ltd.

Key Point

- Understanding defending behaviors is key to mitigating cyberhate.
- Analysis of Normative Social Behavior, Social Learning, and Planned Behavior theories.
- Social Learning theory model explained the most variance in defending behaviors.
- Norms (subjective and injunctive) and self-efficacy were key in defending behavior.
- Girls and younger adolescents were more likely to defend cyberhate victims.

memes, etc., that is directed towards individuals or groups with the intention of causing harm, humiliating them, or fostering hatred and hostility (Hawdon et al. 2016; Kansok-Dusche et al. 2023). This phenomenon can have in detrimental effects on adolescents' well-being and mental health (Machackova et al. 2020) and heightened prejudice (March and Marrington 2019), highlighting the need for further research and intervention. Consequently, a mounting body of scientific research has recently sought to refine the conceptualization of cyberhate (Kansok-Dusche et al. 2023) and develop its measurement (Hawdon et al. 2023). Furthermore, it has examined the factors driving adolescents to send (e.g., Wachs et al. 2022) or receive cyberhate (e.g., Obermaier and Schmuck 2022). While these efforts have advanced our understanding of the phenomenon, less attention has been given to positive actions that adolescents can take when encountering cyberhate. For example, defensive behavior (Costello et al. 2023) is an area that has received little research attention.

Certain studies have yielded valuable preliminary insights into this area, identifying key factors that may explain influence whether adolescents choose to intervene in cyberhate episodes. These factors include self-efficacy, the quality of the teacher-student relationship, and the frequency with which cyberhate is witnessed or seen (Costello et al. 2023; Obermaier 2022; Wachs, Wettstein, et al. 2024). Although these findings represent important progress, further exploration and comprehension of defensive behaviors remain imperative (Wachs, Castellanos, et al. 2023; Wachs, Krause, et al. 2023; Wachs, Valido, et al. 2023). In particular, it is crucial to explore how these variables interact with one another because defending behaviors are inherently complex. They go beyond the influence of isolated or casually attributed factors, which may even lead to contradictions. For instance, findings regarding factors such as previous cyberhate victimization remain inconsistent, as it has been linked to both higher defending due to greater empathy and lower defending due to the normalization of violence (Rudnicki et al. 2023). This suggests that such factors may not be fully understood in isolation and highlights the importance of examining their interplay with social and cognitive variables.

To address this gap in the literature, the present study adopts a multifaceted approach by examining adolescents' defending behaviors in response to cyberhate through the lenses of three prominent psychological theories: the Theory of Normative Social Behavior (Rimal and Real 2005), the Social Learning Theory (Bandura 2007), and the Theory of Planned Behavior (Ajzen 1991). These frameworks offer complementary insights, ranging from social norms to the cognitive and motivational processes linked to behavior. The integration of these perspectives is expected to provide a comprehensive understanding of the factors that shape defensive behavior, contributing to both theoretical advancement and practical strategies to effectively address cyberhate.

1.1 | Bystanders' Defending Behaviors to Cyberhate

Cross-national research indicates that between 21% and 59% of adolescents aged 11 to 17 have seen hate content online in the past 12 months (Machackova et al. 2020). Thus, compared to the 3%–13% of victims and 1%–8% of perpetrators involved in this phenomenon, bystanders represent the most prevalent form of involvement, although often indirectly (Machackova et al. 2020). However, this form of involvement – particularly the distinction between active and passive roles – remains underexplored (Costello et al. 2023).

Bystanders' responses are especially important, as they are closely linked to the prevalence, spread, and impact of cyberhate (Garland et al. 2022; Obermaier et al. 2023). Specifically, around 47.3% of adolescents report taking a defensive role when witnessing cyberhate, for instance by supporting the victim or reporting the incident (Costello et al. 2023; Obermaier 2022; Wachs, Bilz, et al. 2024), with girls and younger adolescents being more likely to intervene (Wachs, Wettstein, et al. 2024; Yu and Chen 2024). At the same time, a substantial proportion remain passive in similar situations (Schumann 2024). While these findings document clear variability in bystander responses, research has tended to focus more on describing these patterns than on explaining why adolescents adopt one response over another. As a result, existing evidence offers limited insight into why some adolescents choose to intervene whereas others remain silent. Addressing this gap is particularly relevant, given that non-intervention may contribute to the normalization of online hate speech (Wachs et al. 2019; Wright and Li 2013). Consequently, there is a need to better understand the factors associated with defending behaviors in order to account for these differences.

1.2 | Theory of Normative Social Behavior

One of the predominant theoretical frameworks that can facilitate comprehension of adolescents' defensive behavior is the Theory of Normative Social Behavior (TNSB). This theory posits that social norms influence individuals, who tend to conform to them in order to avoid social rejection and gain approval within their peer group (Rimal and Real 2005). The aforementioned social norms encompass *subjective norms*, which denote perceptions regarding the expectations of others concerning one's actions; *injunctive norms*, which represent the perceived approval of behaviors by significant others, such as friends, family, or teachers; *descriptive norms*, which signify perceptions concerning the prevalence of behaviors; and *collective norms*, which denote prevailing codes of

14679450, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/sjop.70137 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

conduct that stipulate or proscribe the behaviors that group members may adopt (Chung and Rimal 2016).

The relationship between bystanders' responses and social norms has been extensively explored in other violent phenomena, such as cyberbullying (Domínguez-Hernández et al. 2018). In this respect, bystanders' responses to cyberhate can also be strongly influenced by social norms, which play a crucial role in the decision whether or not to intervene in defense of the victim (Wachs et al. 2022). It is reasonable to hypothesize that when adolescents perceive that significant others, such as peers, disapprove of aggressive behaviors and value defensive actions, they are more likely to intervene. Thus, the perception of prosocial norms, when combined with a supportive environment that discourages negative behaviors, has been shown to encourage defensive intervention in the face of cyberhate (Wachs, Castellanos, et al. 2023). This may result in an increased sense of responsibility, which is in fact a key factor motivating adolescents to act in such situations (Domínguez-Hernández et al. 2018; Obermaier 2022).

1.3 | Social Learning Theory

In addition to social norms as contextual factors, Bandura's Social Learning Theory (SLT) emphasizes the role of individual factors such as self-efficacy, prior experience, and motivation also play a key role in shaping behavior through social learning processes (Bandura 2007). This theory has been widely used to understand various forms of cyberviolence, including cyberbullying (i.e., Xiao and Wong 2013). For instance, Xiao and Wong (2013) in a prior study involving 288 young people, primarily aged 18–25, demonstrated its explanatory value by incorporating personal variables, such as self-efficacy, motivation, prior cybervictimization experiences and socio-demographic characteristics, along with environmental variables, including social norms.

In the context of bystanders' responses to cyberhate, there is emerging evidence supporting the relevance of SLT in explaining defensive behavior, considering diverse factors. In particular, high levels of self-efficacy have been shown to be positively associated with intervention in cyberhate episodes, while low perceptions of self-efficacy have been found to inhibit such responses (Obermaier 2022; Schumann 2024). Despite the absence of clear evidence on the impact of prior cyberhate victimization on bystander intervention (Rudnicki et al. 2023), prior experiences of online hate speech, either personally or through friends, have been found to correlate positively with attitudes, perceived behavioral control, and moral norms linked to intention to intervene (Yu and Chen 2024). In contrast, prior cyberhate perpetration has been found to be negatively associated with defending others in such episodes (Wachs, Wettstein, et al. 2024), likely due to reduced empathy, a competence that is linked to defensive actions (Costello et al. 2023).

In accordance with Bandura's theory, the occurrence of socially learned behaviors is contingent upon the presence of a motivator to reproduce that behavior (Bandura 2007). Within the virtual context, it is recognized that one of the primary motivators of numerous adolescent behaviors is the pursuit of approval and social recognition on social media (Dumas et al. 2017). This need for recognition may, in turn, foster the development of prosocial behaviors (Wright 2015), which could increase the likelihood of defending others to receive genuine interactions and praise. However, the interaction of cyberhate bystanders' responses and these types of variables has not yet been explored and needs to be investigated.

1.4 | Theory of Planned Behavior

Another framework that helps explain the complexity of defensive behavior in adolescents is the Theory of Planned Behavior (TPB). TPB is a well-validated behavioral prediction model that posits that behavior is influenced by the intention to perform it and hypothesizes that the stronger the intention, the more likely the behavior is to occur (Ajzen 1991). This behavioral intention is, in turn, influenced by a combination of attitudes towards the behavior (i.e., the value placed on active bystander behavior), subjective norms, and perceived behavioral control (i.e., confidence in one's ability to engage in active bystander behaviors) (Ajzen 2020; Maes et al. 2024).

The TPB theory has been proven effective in explaining adolescent defense behavior in the face of cyberbullying (Hayashi and Tahmasbi 2022) and non-consensual sexting (Maes et al. 2024). It has also been recently incorporated to understand the intention to intervene in cyberhate (Yu and Chen 2024). In the study by Yu and Chen (2024), favorable outcomes were obtained, particularly with behavioral attitudes and perceived behavioral control, which positively predicted adolescents' propensity to intervene. However, that study concentrated exclusively on understanding behavioral intention, so it would be beneficial to investigate whether such a theory also explains behavior itself. Investigating this would enhance our comprehension of the behavior and reveal the underlying motivations behind the intention to act, which would be helpful for optimizing future intervention programs.

1.5 | The Present Study

In light of the significance of bystander intervention in the context of cyberhate, and the limited empirical evidence on this topic– particularly in the absence of a robust theoretical foundation—this study aimed to examine the factors associated with defending behavior among adolescent bystanders. To this end, we drew on three theoretical frameworks to explore different combinations of personal and contextual variables: (1) the Theory of Normative Social Behavior (TNSB), (2) Social Learning Theory (SLT), and (3) the Theory of Planned Behavior (TPB) (see Figure 1). These frameworks were examined separately to identify their relative and complementary contributions, as a necessary step towards informing future integrative or macro-theoretical approaches.

2 | Method

2.1 | Participants

This cross-sectional study involved 2539 students (49.1% boys, 49.2% girls, 1.7% others) aged 11–18 years ($M = 14.07$, SD = 1.39) from 18 middle and high schools in Andalusia, Spain. According

14679450, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/sjop.70137 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

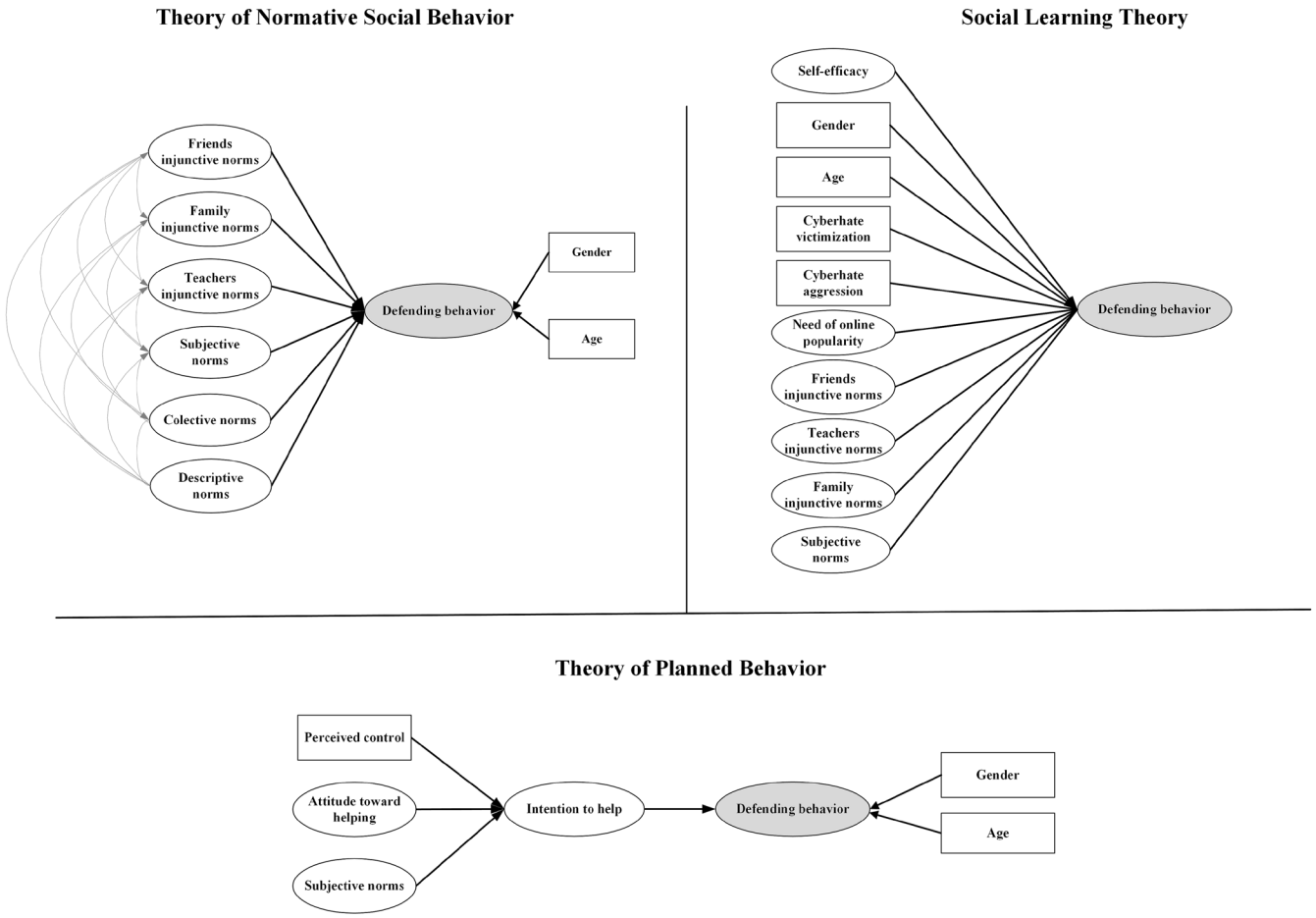


FIGURE 1 | Graphical representation of the hypothetical models of TNSB, SLT, and TPB used in this study to explain bystander defending behaviors to cyberhate.

to the publicly available information on the participating schools, it has been estimated that of the 18 schools included in the study, 16 are from middle-class areas, and two are from deprived areas. The sample was selected using an incidental design by accessibility, where schools were invited to participate.

2.2 | Procedure

This research was conducted in accordance with the ethical standards of the A.P.A. and received approval from the Andalusia Biomedical Research Ethics Coordinating Committee (2563-N-20).

The research team initially contacted the schools' management teams to discuss their potential participation in the project and provide them with relevant information, such as the study's aims and the measures to be used. The proposal was positively received, and the study was incorporated into the school's *convivencia* project, thereby securing informed consent for participation. After obtaining approval from the schools, consent was also obtained from the families. Students who had been authorized to participate were instructed on how to complete the paper-and-pencil survey in class during school hours. They were informed about the voluntary, anonymous, and confidential nature of the research, their right to withdraw at any time, and the importance of answering all questions honestly. The survey, which lasted approximately 60 min, was overseen by the teaching staff and the research team.

2.3 | Measures

Cyberhate involvement. To assess cyberhate involvement, direct items were employed about having seen, sent, received, or intentionally searched for cyberhate messages or media content directed at an individual or group within the previous 12 months. This scale was previously used and validated in Jiménez-Díaz et al. (2025a). It was assessed on a five-point Likert-type scale (from "0 = Never" to "4 = Daily"). The internal consistency was acceptable (Cronbach's $\alpha = 0.79$). As explained in data analysis, this scale was used to assess descriptive norms and collective norms, which also obtained good internal consistency, with a Cronbach's alpha of 0.81.

Defending behaviors to cyberhate. The subscale of defending behavior of *Student Bystander Behavior Scale* (*SBBS*; Thornberg and Jungert 2013), adapted by Álvarez-García et al. (2021), was used and adapted to cyberhate. It was a five-option Likert scale ranging from "0 = Never" to "4 = Daily". This scale was composed of four items that asked participants to respond or reflect on their hypothetical reactions to various forms of cyberviolence, with

14679450, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/sjop.70137 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

a particular focus on episodes of cyberhate in this study (e.g., "Talking to the perpetrator at another time to get him/her to stop"). The internal consistency was good (Cronbach's $\alpha = 0.81$).

Friends' injunctive norms. We used a set of four items from Sasson and Mesch (2016) to assess subjective norms concerning peers' approval of online risky behavior (e.g., "Most of my friends think it's OK to upload an offensive video"). We also incorporated one item from Bastiaensens et al. (2016) to gauge the injunctive norms regarding friends' approval of cyberbullying, adapted in this study to encompass cyberhate (i.e., "Most of my friends think it is OK to act violently towards someone online or on a mobile phone"). The measurement of these items was conducted using a Likert-type scale with five response options (from "0 = Never" to "4 = Daily"). This scale demonstrated good internal consistency (Cronbach's $\alpha = 0.83$).

Family's injunctive norms. We used a set of four items from Sasson and Mesch's (2016) scale on family's injunctive norms regarding online risky behavior (e.g., "My family thinks it is not worth sending an offensive message to someone online"), and we also included one item from Bastiaensens et al. (2016) to assess the injunctive norm of parents approving of cyberviolence (i.e., "My family would not approve of me acting violently online"). These items were measured using a Likert-type scale with five response options (from "0 = Never" to "4 = Daily"). This scale demonstrated an optimal internal consistency (Cronbach's $\alpha = 0.89$).

Teachers' injunctive norms. We used the previous scale referring to family's injunctive norms towards online risky behavior, albeit with adaptations made to align with the teachers' context (e.g., "My teachers think it is not OK to act violently online"). This scale also showed good internal consistency (Cronbach's $\alpha = 0.84$).

Subjective norms. We used a scale composed of three items, as proposed by Hayashi and Tahmasbi (2022), which items referred, for example, to "My important people important to me would want me to help someone who suffers violence on the Internet". These items were assessed using a seven-point Likert-type scale (from "0 = Strongly disagree" to "6 = Strongly agree"). The internal consistency was good (Cronbach's $\alpha = 0.81$).

Need for online popularity. We used four items from a scale that assessed the need for popularity (Santor et al. 2000), adapting them to virtual environments which were used in Jiménez-Díaz et al. (2025b). Items asked participants how many times they had engaged in specific social network behaviors to be popular (e.g., "Sometimes, I have ignored some people to be more popular"). These items were measured using a Likert-type scale with five response options (from "0 = Totally disagree" to "4 = Totally agree"). This scale demonstrated acceptable internal consistency (Cronbach's $\alpha = 0.70$).

Perceived control. We used an adaptation of Hayashi and Tahmasbi's (2022) item measuring perceived control over helping (i.e., "If someone is suffering online violence, it is up to me whether or not to help"). The item was Likert-type with five response options (from "0 = Totally disagree" to "4 = Totally agree").

Self-efficacy. We used the 9 items from the *Adolescent Self-efficacy Questionnaire to Solve Cyberbullying Situations* (Ferreira et al. 2020), by focusing on cyberhate (e.g., "I think I can help to resolve the situation if I see someone threatening someone else online"). These items were assessed using a five-point Likert scale (from "0 = Totally disagree" to "4 = Totally agree"). This scale demonstrated an optimal internal consistency (Cronbach's $\alpha = 0.91$).

Intention to help. We used the *Intention to Help Scale*, as proposed by Hayashi and Tahmasbi (2022), consisting of three items (e.g., "If I see someone suffering from violence on the Internet, I think about helping them"). These items were measured using a seven-point Likert scale (from "0 = Totally disagree" to "6 = Totally agree"). The internal consistency was good (Cronbach's $\alpha = 0.87$).

Attitude towards helping. We used four semantic differential items from Hayashi and Tahmasbi (2022), which items referred, for example, to "For me, helping someone who is suffering online violence is...". These items were assessed using a seven-point Likert scale (from "0 = Bad" to "6 = Good"). This scale demonstrated an optimal internal consistency (Cronbach's $\alpha = 0.91$).

2.4 | Data Analysis

After coding and re-coding the data in the statistical software SPSS—version 29, descriptive statistics were calculated for the means and standard deviations of bystanders' defending behaviors and the study variables. Then, the data were exported to EQS 6.4.

Subsequently, to achieve the first aim of the study, a structural equation model on the direct relationship between social norms and defending behavior was estimated. Following the TNSB theoretical model, social norms comprised six latent variables: subjective norms, injunctive norms from friends, family and teachers, descriptive norms, and collective norms (Chung and Rimal 2016). All latent variables were specified using their corresponding scale items as reflective indicators, except for descriptive and collective norms. The latent variable of descriptive norms comprised the four items referring to cyberhate involvement behaviors (seeing, sending, receiving, and searching). The collective norms were the non-self-mean of cyberhate involvement, which, following Rimal and Yilma (2022), was calculated by subtracting the target person's score from the school's total score on cyberhate involvement and then dividing by one less than the total number of people in the school. Defending behavior was also modeled as a latent variable indicated by its corresponding items. Gender and age were included in the model as control variables. The variable 'gender' was dichotomised into boys (0) and girls (1), excluding the option 'other' from the analysis due to its low frequency, and treating it as a missing value in the missing cases analysis.

Next, to achieve the second aim of the study, a second model was conducted following the SLT theoretical model. To this end, as environmental factors and in accordance with Xiao and Wong (2013), we included subjective norms and injunctive norms from friends, family, and teachers as latent variables.

14679450, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/sjop.70137 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

These variables were modeled using their respective items as indicators. As personal factors, the perceived self-efficacy and the need for online popularity were included as latent variables, and cyberhate victimization and perpetration experiences, gender, and age as direct variables. All of these variables were directly directed towards the latent variable of the defending behavior.

Finally, to achieve the third aim of the study, a third model was estimated following the TPB theoretical model. This model incorporated the direct variable of perceived control, along with the latent variables of helping intention and attitudes towards helping, which were directed towards the latent variable of intention to help. This intention to help, in turn, was directed towards the latent variable of defending behavior. Gender and age were included as control variables. It is also noteworthy that all latent variables in this model were calculated by the means of their items.

All models were run using the Robust Least Squares method, which was appropriate to the categorical nature of the variables under study. The fit of the models was tested with the following indices: Satorra-Bentler scaled chi-square (χ^2 S-B; Satorra and Bentler 2001); the comparative fit index (CFI) and the non-normality fit index (NNFI), where ≥ 0.90 is adequate and ≥ 0.95 is optimal; and the root mean square error of approximation (RMSEA), where ≤ 0.08 is adequate and ≤ 0.05 is optimal (Hu and Bentler 1999). In addition, the p-value < 0.05 was considered significant.

2.5 | Attrition Analysis

Missing data analyses were rigorously tested to identify potential disparities between adolescents who completed the full battery of instruments and those who had missing data in any of the study's variables or who identified their gender as "other". The results indicated no significant differences in friends' injunctive norms ($M_{\text{without missing}} = 0.76$, SD = 0.76; $M_{\text{with missing}} = 0.82$, SD = 0.81; $t(1239.317) = 1.568$, $p = 0.06$, $d = 0.07$), the need of online popularity ($M_{\text{without missing}} = 0.75$, SD = 0.80; $M_{\text{with missing}} = 0.74$, SD = 0.83; $t(1363.444) = -0.199$, $p = 0.42$, $d = -0.01$), having seen cyberhate ($M_{\text{without missing}} = 1.54$, SD = 1.20; $M_{\text{with missing}} = 1.54$, SD = 1.24; $t(2351) = 0.49$, $p = 0.48$, $d = 0.01$) and in perceived control ($M_{\text{without missing}} = 4.02$, SD = 1.99; $M_{\text{with missing}} = 4.23$, SD = 1.95; $t(2427) = 0.36$, $p = 0.36$, $d = 0.02$). However, there were significant differences in teachers' ($M_{\text{without missing}} = 3.10$, SD = 0.97; $M_{\text{with missing}} = 2.93$, SD = 1.13; $t(1138.326) = -3.509$, $p < 0.001$, $d = -0.16$) and family's injunctive norms ($M_{\text{without missing}} = 3.32$, SD = 0.86; $M_{\text{with missing}} = 3.19$, SD = 1.00; $t(1173.982) = -2861$, $p < 0.001$, $d = -0.13$), perceived self-efficacy ($M_{\text{without missing}} = 1.97$, SD = 1.32; $M_{\text{with missing}} = 1.72$, SD = 1.32; $t(890.248) = -3.800$, $p < 0.001$, $d = -0.19$), intention to help ($M_{\text{without missing}} = 4.18$, SD = 1.63; $M_{\text{with missing}} = 3.95$, SD = 1.74; $t(1058.344) = 2.881$, $p = 0.002$, $d = 0.14$), attitudes towards helping ($M_{\text{without missing}} = 5.32$, SD = 1.13; $M_{\text{with missing}} = 5.06$, SD = 1.37; $t(1027.510) = -4.558$, $p < 0.001$, $d = -0.21$), defending behavior ($M_{\text{without missing}} = 2.07$, SD = 1.17; $M_{\text{with missing}} = 1.79$, SD = 1.20; $t(2121) = -4.601$, $p < 0.001$, $d = -0.23$), and having sent ($M_{\text{without missing}} = 0.32$, SD = 0.73; $M_{\text{with missing}} = 0.48$, SD = 0.90; $t(1224.396) = 4.138$, $p < 0.001$, $d = 0.19$), received ($M_{\text{without missing}} = 0.51$, SD = 0.83; $M_{\text{with missing}} = 0.63$, SD = 0.95; $t(1307.155) = 2.836$, $p < 0.001$, $d = 0.13$) and searched for cyberhate ($M_{\text{without missing}} = 0.34$, SD = 0.79; $M_{\text{with missing}} = 0.48$, SD = 0.99; $t(1203.244) = 3.373$, $p < 0.001$, $d = 0.16$). Despite these differences, based on the Cohen's d values shown, the effect sizes were small in all cases. Consequently, these discrepancies were no longer considered in the subsequent analyses (Muñoz-Fernández et al. 2023).

3 | Results

3.1 | Descriptive Analyses

Adolescents' defending behaviors showed that the most common response was encouraging the victim to tell a trusted adult, for example, family or teachers ($M = 2.40$, SD = 1.54), followed by telling an adult or reporting it in the social network ($M = 1.99$, SD = 1.54), directly confronting the perpetrator ($M = 1.82$, SD = 1.42) and, finally, seeking social support to face the situation ($M = 1.82$, SD = 1.43). Table 1 shows the descriptive analyses of the variables included in the models.

3.2 | TNSB Model

Concerning the TNSB theoretical model, a structural equation model has been estimated with the direct relationship between social norms and bystanders' defending behavior towards cyberhate, showing good fit indices: χ^2 SB = 3418.8460; $p < 0.001$; RMSEA = 0.052; CFI = 0.95; NNFI = 0.94. The explanatory capacity of the TNSB model for the defending behavior was 28.4%.

TABLE 1 | Descriptives of the models' variables.

	n	*M*	SD
Defending behavior	2123	2.00	1.19
Seeing cyberhate	2353	1.54	1.22
Sending cyebrhate	2345	0.37	0.79
Receiving cyberhate	2346	0.55	0.88
Searching cyberhate	2343	0.38	0.86
Friends' injunctive norms	2363	0.78	0.78
Teachers' injunctive norms	2375	3.05	1.02
Family's injunctive norms	2389	3.28	0.91
Subjective norms	2226	4.51	1.52
Self-efficacy	2132	1.92	1.30
Need of online popularity	2423	0.75	0.81
Perceived control	2423	4.21	1.98
Intention to help	2214	4.12	1.67
Attitude towards helping	2249	5.25	1.22

Note: The range of values was 0–4 for defending behaviors, seeing, sending, receiving, and searching for cyberhate; friends', teachers', and family's injunctive norms; need for online popularity; perceived control; and self-efficacy. The range was 0–6 for subjective norms, intention to help, and attitudes towards helping.

14679450, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/sjop.70137 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

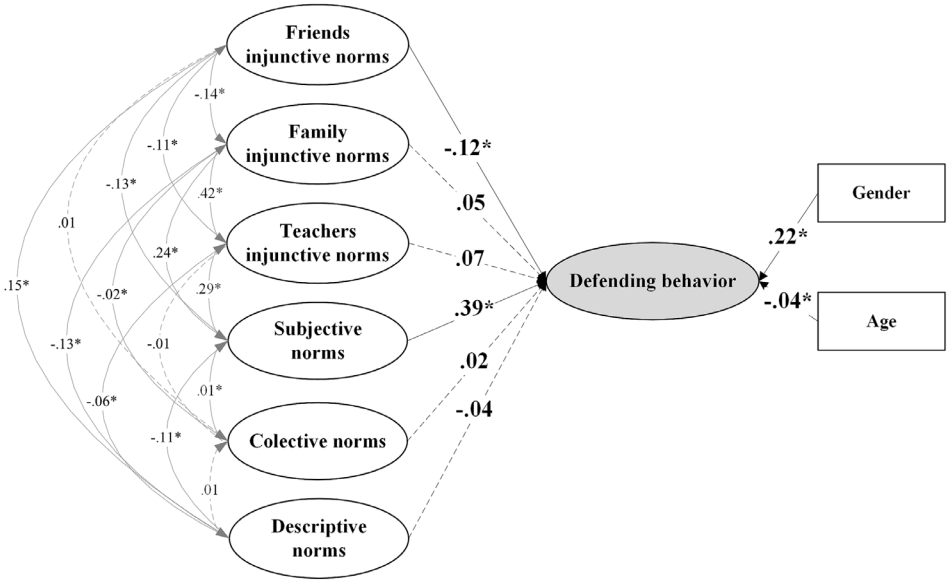


FIGURE 2 | Graphical representation of TSNB model for defending behavior towards cyberhate. *$p < 0.05$.

The relationships between variables and the β coefficients (see Figure 2) demonstrated positive significant relationships with being a girl ($\beta = 0.22$) and subjective norms ($\beta = 0.39$), and negative significant relationships with age ($\beta = -0.04$) and friends' injunctive norms ($\beta = -0.12$) with the defending behavior.

3.3 | SLT Model

The structural equation model based on SLT theoretical model also showed good fit indices: χ^2 SB = 14621.7650; $p < 0.001$; RMSEA = 0.03; CFI = 0.98; NNFI = 0.98, with an explanatory capacity of 50.8% for defending behavior.

The relationships between variables and the β coefficients (see Figure 3) demonstrated positive significant associations with being a girl ($\beta = 0.23$), subjective norms ($\beta = 0.44$), self-efficacy ($\beta = 0.25$), and teachers' ($\beta = 0.22$) and family's injunctive norms ($\beta = 0.24$). In contrast, negative significant associations were obtained with age ($\beta = -0.05$), friends' injunctive norms ($\beta = -0.24$), the need for online popularity ($\beta = -0.09$), and cyberhate victimization ($\beta = -0.05$) and perpetration ($\beta = -0.16$).

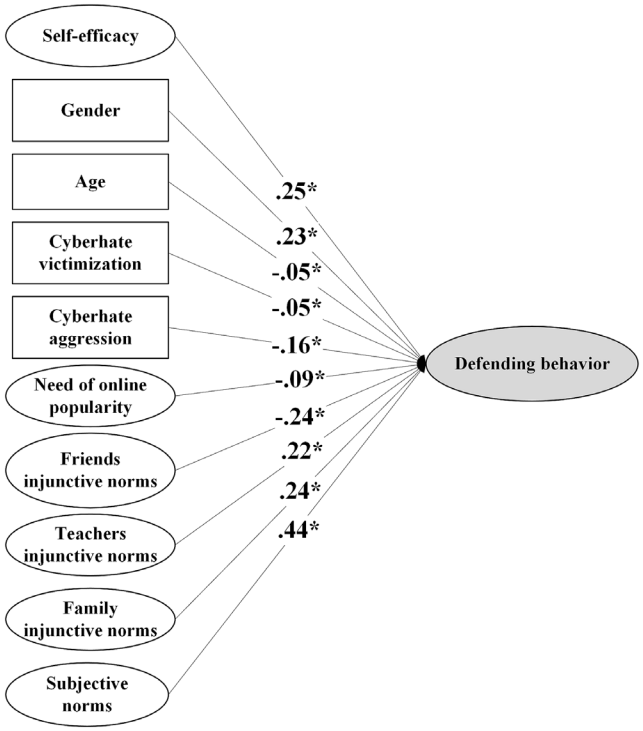


FIGURE 3 | Graphical representation of SLT model for defending behavior towards *cyberhate*. *$p < 0.05$.

3.4 | TPB Model

The structural equation model based on the TPB theoretical model showed good fit indices: χ^2 SB = 540.1609; $p < 0.001$; RMSEA = 0.02; *CFI* = 0.99; NNFI = 0.99. The explanatory capacity of the model was 60.6% for helping intention and 38.5% for defending behavior.

For helping intention (see Figure 4), positive significant associations were identified with subjective norms ($\beta = 0.63$), attitudes towards helping ($\beta = 0.17$), and perceived control ($\beta = 0.11$). Conversely, for defending behavior, positive significant associations were identified with being a girl ($\beta = 0.22$) and helping intention ($\beta = 0.58$). However, a negative association was obtained with age ($\beta = -0.06$).

4 | Discussion

The present study explored the contributions of three theoretical frameworks, that is, TNSB, SLT, and TPB, in explaining intentions and behaviors related to defending behaviors to cyberhate among adolescents. Our findings emphasize the crucial role of

14679450, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/sjop.70137 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

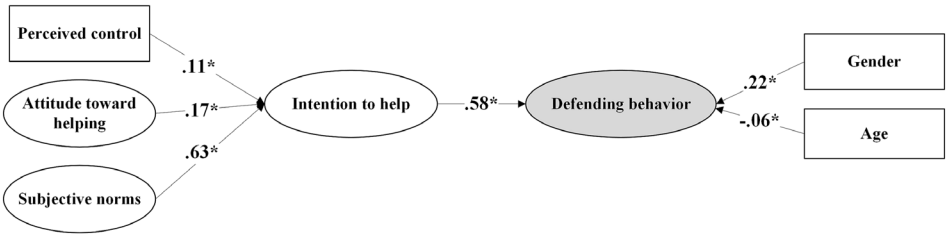


FIGURE 4 | Graphical representation of TPB model for defending behavior towards cyberhate. *$p < 0.05$.

social norms and other personal factors in shaping behavior, offering significant theoretical insights and practical guidance for developing psychoeducational programs aimed at fostering defending behaviors to cyberviolence, particularly in the context of cyberhate.

According to the first aim of our study, the results indicate that the primary contribution of the TNSB in elucidating adolescents' defending behavior to cyberhate lies in the significance of subjective norms and friends' injunctive norms. Subjective norms, which are defined as adolescents' perceptions regarding the expectations of significant people in their lives (Chung and Rimal 2016), have emerged as a key factor in the intention to adopt defending behaviors. This fact reinforces the idea that adolescents are more likely to engage in certain behaviors when they perceive them as accepted and validated by their close social environment (Rimal and Real 2005). Hence, adolescents who perceive that their social circle expects or supports defending behaviors in response to cyberhate are more inclined to act, perhaps due to an enhanced sense of responsibility (Obermaier 2022). This influence is particularly pronounced within the peer group, as evidenced by the decisive role of friends' injunctive norms regarding online violent behaviors. In this regard, adolescents who perceive that their peers disapprove of such behaviors are more likely to adopt a defending posture in cyberhate situations. This result aligns with previous research on analogous phenomena, such as cyberbullying (Domínguez-Hernández et al. 2018; Rudnicki et al. 2023). Furthermore, the perceived expectations of peers and significant others appear to carry greater weight than the descriptive and collective norms that adolescents directly observe or experience in their environment, which were not significant in our study. This finding suggests that adolescents are more motivated by the desire to align with perceived expectations than by replicating observed behaviors. This underscores the importance of interventions that cultivate a supportive peer environment, where defending behavior is widely recognized and expected within adolescents' social circles. In fact, prior research has demonstrated that fostering such an environment can significantly increase these responses (Macaulay et al. 2022; Wachs, Castellanos, et al. 2023). However, although subjective norms and friends' injunctive norms explained a significant proportion of variance, other norms were not significant; it can be concluded that there are still important aspects to explore in the complexity of this behavior, such as personal factors.

The above statement is further supported by the findings related to the second aim of our study, which examined the contribution of the SLT framework to explaining defending behavior in the context of cyberhate. Although direct comparisons in terms of explanatory power should be interpreted with caution due to differences in model structure and included variables, the SLT framework provided additional insights by incorporating individual-level factors alongside social norms. This suggests that defending behavior is influenced not only by normative processes but also by personal experiences and individual characteristics. On the one hand, the results for social norms consistently highlight the importance of subjective and friends' injunctive norms. However, when additional factors are included in the model, the importance of all social norms highlights their complex interactions, with family's and teachers' injunctive norms becoming particularly relevant in explaining such behavior. In this respect, it is known that families and teachers play a crucial role in the development of adolescents' identities by transmitting values, social norms, and acceptable behaviors (Livingstone and Byrne 2018; Verhoeven et al. 2019). Therefore, when adolescents perceive that their adult role models disapprove of violent behavior, they are more likely to recognize the importance of intervening in such situations (Domínguez-Hernández et al. 2018). This highlights the need to actively involve families and teachers in psychoeducational programs to prevent cyberhate (Wachs, Wettstein, et al. 2024; Wright et al. 2021).

A relevant contribution of our second model is the attention given to personal factors as potential correlates of the defending behaviors against cyberhate. Specifically, the inclusion of individual variables in SLT model, such as prior victimization and perpetration experiences, along with motivators such as the need for online popularity, has shed light on adolescent defending behavior, as these factors appear to inhibit such responses. Instead, only one of the individual variables included, namely self-efficacy, emerges as the most relevant factor associated with defending behavior. Specifically, adolescents who perceive themselves as capable of successfully countering cyberhate are more likely to engage in counter-speech (Wachs, Castellanos, et al. 2023; Wachs, Krause, et al. 2023), which is consistent with broader research on bystander intervention in bullying and cyberbullying (Ferreira et al. 2020; Thornberg et al. 2020). Hence, the importance of self-efficacy highlights the need to equip adolescents with the skills and confidence to respond effectively to cyberhate, possibly through programs that incorporate conflict resolution training and role-playing to empower adolescents' defending behaviors to cyberhate (Wachs, Krause, et al. 2023).

With regard to the third aim of our study, which focuses on the TPB framework, our findings provide evidence regarding the variables associated with the intention to help. In particular, subjective norms emerge as a key factor associated with defending behaviors, along with positive attitudes towards helping and perceived behavioral control. These results are consistent with

14679450, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/sjop.70137 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

a previous study conducted by Yu and Chen (2024), examining young adults' responses to hate speech on the Internet in the TPB setting. In that study, although the results regarding subjective norms were different, the results regarding attitudes towards helping and perceived control point in the same direction (Yu and Chen 2024). Thus, it appears that adolescents who perceive themselves as having control over their ability to help and who positively value such actions are more likely to express an intention to help in online hate situations as well as in other forms of cyberviolence (Hayashi and Tahmasbi 2022; Maes et al. 2024). This reinforces the notion that it is important to encourage adolescents to view defense as valuable and positive, while providing them with the tools they need to feel confident and empowered to take action.

However, as observed from the TPB approach, although the intention to defend is a strong predictor of defending behavior, it does not fully explain the final intervention in these situations. Hence, our results suggest that although adolescents intend to defend others, they do not always follow through with their actions. This points to the possibility that the decision to intervene is explained by situational or contextual factors, such as the severity of the incident (Bastiaensens et al. 2014), the relationship with the victim (Patterson et al. 2017), or the presence of other bystanders (Obermaier et al. 2016). Even if there is an initial intention to help, these variables may condition the likelihood of defending or not. This highlights two important directions for future research and prevention. First, there is a need to explore further the factors that facilitate or hinder the decision to intervene to better understand the complex dynamics of defending behavior. Second, promoting responsible decision-making in adolescents is critical. Previous research has linked responsible decision-making to seeking social support as a coping strategy in the context of cyberbullying (Espino et al. 2023), opening the possibility that similar mechanisms may explain the decision to defend in online hate situations. Consequently, the promotion of responsible decision-making in psychoeducational programs is imperative to ensure that adolescents' intentions to defend are not merely intentional but translate into final defending behavior, even in the presence of situational factors that might otherwise discourage intervention.

Finally, a salient finding common to all of our theoretical models (TNSB, SLT, and TPB) pertains to the role of gender and age in defending behavior against cyberhate. Our findings from the three models indicate that being a girl and younger is associated with a greater predisposition to adopt a defending behavior. These results are consistent with prior studies on bystanders' behavior to cyberhate. These studies suggest that girls (Wachs, Wettstein, et al. 2024) and younger individuals (Yu and Chen 2024) are more prone to adopt defending behavior when witnessing episodes of cyberhate. On the one hand, our findings on gender are in line with previous evidence that suggests that girls exhibit more prosocial behavior in social media and feel a greater sense of social responsibility, which could lead them to respond with greater empathy (Pastor et al. 2024; Van der Graaff et al. 2018). However, these findings should be interpreted with caution, as gender was examined using a binary operationalization due to sample constraints, which may not adequately capture the diversity and complexity of gender identities in cyberhate contexts. Future research should address this issue more comprehensively, particularly given the relevance of stigma in cyberhate situations and its potential influence on adolescents' willingness to intervene. On the other hand, our findings on age could be explained by the fact that as adolescents get older, they have witnessed more of these behaviors and thus may have normalized their occurrence, thereby diminishing the defending behavior to them (Machackova and Pfetsch 2016). This points to the need to design and implement psychoeducational programs that adopt a gender-based and early prevention strategy to support defending behaviors in both boys and girls from an early age.

4.1 | Practical Implications

The findings suggest three priorities for prevention in cyberhate contexts. First, interventions should strengthen supportive social norms around defending behavior, as subjective norms were consistently associated with defending across models, and adult injunctive norms were also relevant in the SLT model. This highlights the importance of involving peers, families, and teachers in communicating that defending victims of cyberhate is expected and supported (Garandeau et al. 2023; Wachs, Castellanos, et al. 2023).

Second, prevention efforts should strengthen adolescents' self-efficacy to intervene. Because self-efficacy emerged as a key correlate of defending behavior, programs should provide adolescents with clear and feasible guidance on how to identify cyberhate and respond safely and effectively in these situations (Wachs, Krause, et al. 2023; Wachs, Castellanos, et al. 2023).

Third, interventions should help adolescents translate helping intentions into action. Although helping intention was positively associated with defending behavior, it did not fully account for adolescents' responses. Prevention efforts should therefore reduce the intention–behavior gap by offering concrete decision-making guidance for ambiguous or socially risky online situations (Obermaier 2022; Wachs, Wettstein, et al. 2024).

Finally, prevention should begin early and be sensitive to groups who may be less likely to defend, while interpreting gender-related findings with caution due to the binary operationalization used in this study. Overall, the findings support prevention approaches that combine normative support, confidence to intervene, and concrete action-oriented guidance.

4.2 | Limitations

Despite the potential of this study, there are several limitations that must be assumed. First, the cross-sectional design limits our ability to determine the temporal sequence of variables and, as data were collected at a single time point, prevents the establishment of causal relationships. It is therefore essential that future longitudinal research is conducted to clarify whether the variables included in the models can predict defending behavior in the future. Second, the sophisticated character of certain forms of online hate, frequently masked by the guise of humor, engenders challenges in identification and measurement, which may also affect adolescents' recognition and reported responses

14679450, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/sjop.70137 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

to such incidents. In addition, defending behavior was assessed through hypothetical scenarios and self-report measures, which may limit the extent to which the findings reflect actual behavior and may be subject to social desirability and recall biases. This may exert an influence on adolescents' recognition and response to such incidents. Furthermore, the analysis controlled for gender only between boys and girls due to the limited number of non-binary participants, which may have reduced the sensitivity of the analyses to capture more nuanced gender identities and experiences, underscoring the need for future studies with larger and more diverse samples. Likewise, it is important to note that the study was conducted within a specific region of Spain, and caution should be exercised when generalizing the findings to broader populations. Finally, it should be noted that EQS reports parameter significance primarily through test statistics (z-values derived from the ratio of estimates to their standard errors), rather than directly providing exact p-values for each parameter (Bentler 2006). Although significance levels can be inferred from these statistics, the absence of explicitly reported exact p-values may limit the precision of statistical reporting.

5 | Conclusions

Despite these limitations, this study has provided valuable theoretical insights into understanding adolescents' defending behavior against cyberviolence, particularly cyberhate. In our study, the contributions of TNSB, SLT, and TPB frameworks highlighted social norms as key factors associated with behavior, along with individual factors such as self-efficacy, gender, and age. They have also highlighted the importance of positive attitudes towards helping and perceived control in influencing intentions to intervene, although this does not always translate into final behavior. These findings demonstrate that while psychological theories provide valuable frameworks for interpreting defending behavior, its complexity suggests the need for a macro-theory that integrates multiple factors.

Author Contributions

Olga Jiménez-Díaz: conceptualization, methodology, formal analysis, investigation, data curation, writing – original draft, visualization. **Esperanza Espino:** conceptualization, writing – review and editing, investigation, data curation, visualization. **Rosario Del Rey:** resources, writing – review and editing, investigation, supervision, project administration, funding acquisition. **Joaquín A. Mora-Merchán:** resources, writing – review and editing, investigation, supervision, project administration, funding acquisition.

Acknowledgments

The authors wish to acknowledge the Grant PID2020-115913GB-I00 and Grant PID2023-148735NB-I00 funded by MICIU/AEI/10.13039/501100011033; and the Grant FPU21/05405 by the Ministry of Universities of the Government of Spain.

Funding

This work was supported by the Grant PID2020-115913GB-I00 funded and Grant PID2023-148735NB-I00 funded by MICIU/AEI/10.13039/501100011033; and by the Ministry of Universities of the Government of Spain (FPU21/05405).

Ethics Statement

This study was approved by the Andalusian Biomedical Research Ethics Coordinating Committee (2563-N-20).

Consent

Participation in the study was voluntary. All participants, their guardians, and schools gave informed consent/assent after being fully informed about the study.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are openly available in OSF at https://osf.io/q5j8g/, reference number 10.17605/OSF.IO/9ASRT.

References

Ajzen, I. 1991. "The Theory of Planned Behavior." *Organizational Behavior and Human Decision Processes* 50, no. 2: 179–211. https://doi.org/10.1016/0749-5978(91)90020-T.

Ajzen, I. 2020. "The Theory of Planned Behavior: Frequently Asked Questions." *Human Behavior and Emerging Technologies* 2, no. 4: 314–324. https://doi.org/10.1002/hbe2.195.

Álvarez-García, D., R. Thornberg, and Z. Suárez-García. 2021. "Validation of a Scale for Assessing Bystander Responses in Bullying." *Psicothema* 33, no. 4: 623–630. https://doi.org/10.7334/psicothema2021.140.

Bandura, A. 2007. "Social Learning Theory of Aggression." In *Control of Aggression*. Routledge.

Barrense-Dias, Y., J.-C. Suris, and L. Chok. 2024. ""Maybe They Don't Necessarily Realize the Damage They Can Do...": A Qualitative Study on Bystanders to Non-Consensual Forwarding of Nudes Among Adolescents in Switzerland." *International Journal of Bullying Prevention* 7: 623–631. https://doi.org/10.1007/s42380-024-00249-2.

Bastiaensens, S., S. Pabian, H. Vandebosch, et al. 2016. "From Normative Influence to Social Pressure: How Relevant Others Affect Whether Bystanders Join in Cyberbullying." *Social Development* 25, no. 1: 193–211. https://doi.org/10.1111/sode.12134.

Bastiaensens, S., H. Vandebosch, K. Poels, K. Van Cleemput, A. DeSmet, and I. De Bourdeaudhuij. 2014. "Cyberbullying on Social Network Sites. An Experimental Study Into Bystanders' Behavioral Intentions to Help the Victim or Reinforce the Bully." *Computers in Human Behavior* 31: 259–271. https://doi.org/10.1016/j.chb.2013.10.036.

Bentler, P. M. 2006. *EQS 6 Structural Equations Program Manual*. Multivariate Software, INC.

Chung, A., and R. N. Rimal. 2016. "Social Norms: A Review." *Review of Communication Research* 4: 1–28. https://doi.org/10.12840/issn.2255-4165.2016.04.01.008.

Costello, M., J. Hawdon, A. V. Reichelmann, et al. 2023. "Defending Others Online: The Influence of Observing Formal and Informal Social Control on One's Willingness to Defend Cyberhate Victims." *International Journal of Environmental Research and Public Health* 20, no. 15: 6506. https://doi.org/10.3390/ijerph20156506.

Domínguez-Hernández, F., L. Bonell, and A. Martínez-González. 2018. "A Systematic Literature Review of Factors That Moderate Bystanders' Actions in Cyberbullying." *Cyberpsychology: Journal of Psychosocial Research on Cyberspace* 12, no. 4: 4. https://doi.org/10.5817/CP2018-4-1.

14679450, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/sjop.70137 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Dublin City University. 2024. "School Bullying, an Inclusive Definition." UNESCO Chair. Bullying-An-Inclusive-Definition-UNESCO-WABF.pdf.

Dumas, T. M., M. Maxwell-Smith, J. P. Davis, and P. A. Giulietti. 2017. "Lying or Longing for Likes? Narcissism, Peer Belonging, Loneliness and Normative Versus Deceptive Like-Seeking on Instagram in Emerging Adulthood." *Computers in Human Behavior* 71: 1–10. https://doi.org/10.1016/j.chb.2017.01.037.

Espino, E., A. Guarini, and R. Del Rey. 2023. "Effective Coping With Cyberbullying in Boys and Girls: The Mediating Role of Self-Awareness, Responsible Decision-Making, and Social Support." *Current Psychology* 42, no. 36: 32134–32146. https://doi.org/10.1007/s12144-022-04213-5.

Ferreira, P. C., A. M. V. Simão, A. Paiva, and A. Ferreira. 2020. "Responsive Bystander Behavior in Cyberbullying: A Path Through Self-Efficacy." *Behavior & Information Technology* 39, no. 5: 511–524. https://doi.org/10.1080/0144929X.2019.1602671.

Garandeau, C. F., T. Turunen, S. Saarento-Zaprudin, and C. Salmivalli. 2023. "Effects of the KiVa Anti-Bullying Program on Defending Behavior: Investigating Individual-Level Mechanisms of Change." *Journal of School Psychology* 99: 101226. https://doi.org/10.1016/j.jsp.2023.101226.

Garland, J., K. Ghazi-Zahedi, J.-G. Young, L. Hébert-Dufresne, and M. Galesic. 2022. "Impact and Dynamics of Hate and Counter Speech Online." *EPJ Data Science* 11, no. 1: 1. https://doi.org/10.1140/epjds/s13688-021-00314-6.

Hawdon, J., A. Oksanen, and P. Räsänen. 2016. "Exposure to Online Hate in Four Nations: A Cross-National Consideration." *Deviant Behavior* 38: 254–266. https://doi.org/10.1080/01639625.2016.1196985.

Hawdon, J., A. Reichelmann, M. Costello, et al. 2023. "Measuring Hate: Does a Definition Affect Self-Reported Levels of Perpetration and Exposure to Online Hate in Surveys?" *Social Science Computer Review* 42: 08944393231211270. https://doi.org/10.1177/08944393231211270.

Hayashi, Y., and N. Tahmasbi. 2022. "Psychological Predictors of Bystanders' Intention to Help Cyberbullying Victims Among College Students: An Application of Theory of Planned Behavior." *Journal of Interpersonal Violence* 37, no. 13–14: NP11333–NP11357. https://doi.org/10.1177/0886260521992158.

Hu, L. T., and P. M. Bentler. 1999. "Cutoff Criteria for fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus new Alternatives." *Structural Equation Modeling* 6, no. 1: 1–55. https://doi.org/10.1080/10705519909540118.

Jiménez-Díaz, O., R. Del Rey, E. Espino, and J. A. Casas. 2025b. "Understanding Cyberhate: The Importance of Like-Seeking Behaviours." *Journal for the Study of Education and Development* 48, no. 3: 601–629. https://doi.org/10.1177/02103702251356207.

Jiménez-Díaz, O., S. Wachs, R. Del Rey, and J. A. Mora-Merchán. 2025a. "Associations Between Searching and Sending Cyberhate: The Moderating Role of the Need of Online Popularity and Toxic Online Disinhibition." *Cyberpsychology, Behavior, and Social Networking* 28, no. 1: 37–43. https://doi.org/10.1089/cyber.2024.0305.

Kansok-Dusche, J., C. Ballaschk, N. Krause, et al. 2023. "A Systematic Review on Hate Speech Among Children and Adolescents: Definitions, Prevalence, and Overlap With Related Phenomena." *Trauma, Violence & Abuse* 24, no. 4: 2598–2615. https://doi.org/10.1177/15248380221108070.

Livingstone, S., and J. Byrne. 2018. "Parenting in the Digital Age: The Challenges of Parental Responsibility in Comparative Perspective (19–30)." Nordicom, University of Gothenburg. https://urn.kb.se/resolve?urn=urn:nbn:se:norden:org:diva-12015.

Macaulay, P. J. R., L. R. Betts, J. Stiller, and B. Kellezi. 2022. "Bystander Responses to Cyberbullying: The Role of Perceived Severity, Publicity, Anonymity, Type of Cyberbullying, and Victim Response." *Computers in Human Behavior* 131: 107238. https://doi.org/10.1016/j.chb.2022.107238.

Machackova, H., C. Blaya, M. Bedrosova, D. Smahel, and E. Staksrud. 2020. "Children's Experiences With Cyberhate. EU Kids Online." https://doi.org/10.21953/lse.zenkg9xw6pua.

Machackova, H., and J. Pfetsch. 2016. "Bystanders' Responses to Offline Bullying and Cyberbullying: The Role of Empathy and Normative Beliefs About Aggression." *Scandinavian Journal of Psychology* 57, no. 2: 169–176. https://doi.org/10.1111/sjop.12277.

Maes, C., J. Van Ouytsel, and L. Vandenbosch. 2024. "Active Bystanders in the Forwarding of Sexting Messages: Applying a Theory of Planned Behavior in Youth." *New Media & Society* 28: 14614448241287729. https://doi.org/10.1177/14614448241287729.

March, E., and J. Marrington. 2019. "A Qualitative Analysis of Internet Trolling." *Cyberpsychology, Behavior, and Social Networking* 22, no. 3: 192–197. https://doi.org/10.1089/cyber.2018.0210.

Muñoz-Fernández, N., V. Sánchez-Jiménez, M.-L. Rodríguez-deArriba, L. Nacimiento-Rodríguez, P. Elipe, and R. Del Rey. 2023. "Traditional and Cyber Dating Violence Among Adolescents: Profiles, Prevalence, and Short-Term Associations With Peer Violence." *Aggressive Behavior* 49, no. 3: 261–273. https://doi.org/10.1002/ab.22069.

Obermaier, M. 2022. "Youth on Standby? Explaining Adolescent and Young Adult Bystanders' Intervention Against Online Hate Speech." *New Media & Society* 26, no. 8: 4785–4807. https://doi.org/10.1177/14614448221125417.

Obermaier, M., N. Fawzi, and T. Koch. 2016. "Bystanding or Standing by? How the Number of Bystanders Affects the Intention to Intervene in Cyberbullying." *New Media & Society* 18, no. 8: 1491–1507. https://doi.org/10.1177/1461444814563519.

Obermaier, M., and D. Schmuck. 2022. "Youths as Targets: Factors of Online Hate Speech Victimization Among Adolescents and Young Adults." *Journal of Computer-Mediated Communication* 27, no. 4: zmac012. https://doi.org/10.1093/jcmc/zmac012.

Obermaier, M., D. Schmuck, and M. Saleem. 2023. "I'll Be There for You? Effects of Islamophobic Online Hate Speech and Counter Speech on Muslim In-Group Bystanders' Intention to Intervene." *New Media & Society* 25, no. 9: 2339–2358. https://doi.org/10.1177/14614448211017527.

Pastor, Y., V. Pérez-Torres, H. Thomas-Currás, L. L. Lobato-Rincón, M. Á. López-Sáez, and A. García. 2024. "A Study of the Influence of Altruism, Social Responsibility, Reciprocity, and the Subjective Norm on Online Prosocial Behavior in Adolescence." *Computers in Human Behavior* 154: 108156. https://doi.org/10.1016/j.chb.2024.108156.

Patterson, L. J., A. Allan, and D. Cross. 2017. "Adolescent Perceptions of Bystanders' Responses to Cyberbullying." *New Media & Society* 19, no. 3: 366–383. https://doi.org/10.1177/1461444815606369.

Rimal, R. N., and K. Real. 2005. "How Behaviors Are Influenced by Perceived Norms: A Test of the Theory of Normative Social Behavior." *Communication Research* 32, no. 3: 389–414. https://doi.org/10.1177/0093650205275385.

Rimal, R. N., and H. Yilma. 2022. "Descriptive, Injunctive, and Collective Norms: An Expansion of the Theory of Normative Social Behavior (TNSB)." *Health Communication* 37, no. 13: 1573–1580. https://doi.org/10.1080/10410236.2021.1902108.

Rudnicki, K., H. Vandebosch, P. Voué, and K. Poels. 2023. "Systematic Review of Determinants and Consequences of Bystander Interventions in Online Hate and Cyberbullying Among Adults." *Behavior & Information Technology* 42, no. 5: 527–544. https://doi.org/10.1080/0144929X.2022.2027013.

Santor, D. A., D. Messervey, and V. Kusumakar. 2000. "Measuring Peer Pressure, Popularity, and Conformity in Adolescent Boys and Girls: Predicting School Performance, Sexual Attitudes, and Substance

14679450, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/sjop.70137 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Abuse." *Journal of Youth and Adolescence* 29, no. 2: 163–182. https://doi.org/10.1023/A:1005152515264.

Sasson, H., and G. Mesch. 2016. "Gender Differences in the Factors Explaining Risky Behavior Online." *Journal of Youth and Adolescence* 45, no. 5: 973–985. https://doi.org/10.1007/s10964-016-0465-7.

Satorra, A., and P. M. Bentler. 2001. "A Scaled Difference chi-Square Test Statistic for Moment Structure Analysis." *Psychometrika* 66, no. 4: 507–514. https://doi.org/10.1007/BF02296192.

Schumann, S. 2024. "Examining Bystanders' Decisions to Not Intervene in Incidents of Hate Speech Online and Offline." https://doi.org/10.31234/osf.io/e7q5n.

Thornberg, R., and T. Jungert. 2013. "Bystander Behavior in Bullying Situations: Basic Moral Sensitivity, Moral Disengagement and Defender Self-Efficacy." *Journal of Adolescence* 36, no. 3: 475–483. https://doi.org/10.1016/j.adolescence.2013.02.003.

Thornberg, R., L. Wänström, R. Elmelid, A. Johansson, and E. Mellander. 2020. "Standing Up for the Victim or Supporting the Bully? Bystander Responses and Their Associations With Moral Disengagement, Defender Self-Efficacy, and Collective Efficacy." *Social Psychology of Education* 23, no. 3: 563–581. https://doi.org/10.1007/s11218-020-09549-z.

Van der Graaff, J., G. Carlo, E. Crocetti, H. M. Koot, and S. Branje. 2018. "Prosocial Behavior in Adolescence: Gender Differences in Development and Links With Empathy." *Journal of Youth and Adolescence* 47, no. 5: 1086–1099. https://doi.org/10.1007/s10964-017-0786-1.

Verhoeven, M., A. M. G. Poorthuis, and M. Volman. 2019. "The Role of School in Adolescents' Identity Development. A Literature Review." *Educational Psychology Review* 31, no. 1: 35–63. https://doi.org/10.1007/s10648-018-9457-3.

Wachs, S., L. Bilz, A. Wettstein, and D. L. Espelage. 2024. "Validation of the Multidimensional Bystander Responses to Racist Hate Speech Scale and Its Association With Empathy and Moral Disengagement Among Adolescents." *Aggressive Behavior* 50, no. 1: e22105. https://doi.org/10.1002/ab.22105.

Wachs, S., M. Castellanos, A. Wettstein, L. Bilz, and M. Gámez-Guadix. 2023. "Associations Between Classroom Climate, Empathy, Self-Efficacy, and Countering Hate Speech Among Adolescents: A Multilevel Mediation Analysis." *Journal of Interpersonal Violence* 38, no. 5–6: 5067–5091. https://doi.org/10.1177/08862605221120905.

Wachs, S., N. Krause, M. F. Wright, and M. Gámez-Guadix. 2023. "Effects of the Prevention Program "HateLess. Together Against Hatred" on Adolescents' Empathy, Self-Efficacy, and Countering Hate Speech." *Journal of Youth and Adolescence* 52, no. 6: 1115–1128. https://doi.org/10.1007/s10964-023-01753-2.

Wachs, S., A. Valido, D. L. Espelage, M. Castellanos, A. Wettstein, and L. Bilz. 2023. "The Relation of Classroom Climate to Adolescents' Countering Hate Speech via Social Skills: A Positive Youth Development Perspective." *Journal of Adolescence* 95, no. 6: 1127–1139. https://doi.org/10.1002/jad.12180.

Wachs, S., A. Wettstein, L. Bilz, D. L. Espelage, M. F. Wright, and M. Gámez-Guadix. 2024. "Individual and Contextual Correlates of Latent Bystander Profiles Toward Racist Hate Speech: A Multilevel Person-Centered Approach." *Journal of Youth and Adolescence* 53, no. 6: 1271–1286. https://doi.org/10.1007/s10964-024-01968-x.

Wachs, S., A. Wettstein, L. Bilz, and M. Gámez-Guadix. 2022. "Adolescents' Motivations to Perpetrate Hate Speech and Links With Social Norms." *Comunicar* 30, no. 71: 9–20. https://doi.org/10.3916/C71-2022-01.

Wachs, S., M. F. Wright, R. Sittichai, et al. 2019. "Associations Between Witnessing and Perpetrating Online Hate in Eight Countries: The Buffering Effects of Problem-Focused Coping." *International Journal of Environmental Research and Public Health* 16, no. 20: 3992. https://doi.org/10.3390/ijerph16203992.

Wright, M. F. 2015. "Perceptions of Popularity-Related Behaviors in the Cyber Context: Relations to Cyber Social Behaviors." *Child Development Research* 2015, no. 1: 636929. https://doi.org/10.1155/2015/636929.

Wright, M. F., and Y. Li. 2013. "Normative Beliefs About Aggression and Cyber Aggression Among Young Adults: A Longitudinal Investigation." *Aggressive Behavior* 39, no. 3: 161–170. https://doi.org/10.1002/ab.21470.

Wright, M. F., S. Wachs, and M. Gámez-Guadix. 2021. "Youths' Coping With Cyberhate: Roles of Parental Mediation and Family Support." *Comunicar* 29, no. 67: 21–33. https://doi.org/10.3916/C67-2021-02.

Xiao, B. S., and Y. M. Wong. 2013. "Cyber-Bullying Among University Students: An Empirical Investigation From the Social Cognitive Perspective." *International Journal of Business and Information* 8, no. 1: 34–69.

Yu, C., and Y. Chen. 2024. "Online Hate Speech Detection and Management From Bystander Intervention Perspective Based on ETPB Model." *Journal of Electrical Systems* 20, no. 10s: 659–669.

14679450, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/sjop.70137 by Spanish Cochrane National Provision (Ministerio de Sanidad), Wiley Online Library on [11/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License