

Influencers as Nodes of Problematic Content: Misinformation, Hate Speech, and Youth Responses

Rebeca Suárez-Álvarez, Antonio García-Jiménez and Beatriz Catalina-García

Universidad Rey Juan Carlos, Madrid, Spain

rebeca.suarez@urjc.es

antonio.garcia@urjc.es

beatriz.catalina@urjc.es

Abstract: This research explores the way in which young people encounter fake news through the profiles of influencers, as well as its association with their behaviour in avoiding misinformation. The study also examines how exposure to comments containing hate speech is related to either taking a stance of passive observation, or an attitude of active dissemination. This platform-based contextualization enables an integrated assessment of misinformation exposure/avoidance and hate-speech-related responses in youth audiences. Regarding hate speech, sexual orientation has been added to the mix to identify whether, in addition to gender differences, there are specific patterns of vulnerability associated with sexual diversity. A survey was conducted involving 1,800 adolescents and young adults from 14–24 years of age who reside in Spain. The findings indicate that influencers have become a key source of news for young people, which highlights the pivotal role played by these figures as information mediators within the youth ecosystem. Practices involving information verification are mostly driven by curiosity or surprise, whereas consistently stronger and active fact-checking is conducted by those who place their trust in the influencers they follow. As for exposure to hate speech, the phenomenon is widespread, with girls claiming to be the target of such aggression more often. Passivity is the dominant response across groups, whereas sharing is more frequent among boys. Sexual minorities, especially bisexual and homosexual, are more exposed to hate speech than their heterosexual peers. In terms of dissemination, people with other sexual orientations, as well as heterosexuals, engage in the highest levels of sharing, while homosexual youth display the lowest levels, making the latter group the most vulnerable.

Keywords: Youth Digital Agency, Influencer, Misinformation, Hate Speech, Social Media, Cultural Dynamics

1. Influencers as Intermediaries of Misinformation

The information consumption practices of adolescents and young people converge at the crossroads of entertainment and algorithmic logic, which is replacing the authority of traditional media. This research conceptualizes influencers as intermediaries in youth information environments, focusing on shape everyday verification and responses to problematic content. In this environment, influencers have emerged as authority figures in the information inventory of young people. These individuals combine emotion and relational dynamics through narratives that merge both opinion and rhetorical techniques in order to foster an emotional connection with the audience (Mulcahy et al., 2024). This strategy leads to a blurring of the boundary between entertainment and information, resulting in a higher risk of receiving fake news.

According to Niitsuma et al. (2025), posts disseminated by users who have a strong digital impact, such as influencers, are more likely to be reposted than publications sent by conventional users, thereby allowing the former to reach larger audiences. Along the same lines, Shin et al. (2022) confirm that during the pandemic, users believed that false content was more credible when it had been shared by their contacts, and when it came from highly popular profiles. These authors point out that influencers are able to intensify the perception of credibility toward the information they share, not only because of their fame, but also due to the genuine relationships they build with their followers. Wasike (2025) confirms that influencers' followers have significantly higher levels of trust in the content produced by these figures, yet adds that this trust also includes traditional media, which seems to confirm that the consumption of news divulged by influencers might play a complementary role in overall media trust.

The fact is that teenagers are aware of the presence and danger of fake news on social media (Selnes, 2024). To avoid it, they develop verification strategies focused mainly on checking sources, consulting traditional media, and consciously using options such as ignore, block, or report. According to Orhan (2023), increased literacy in the new media, along with enhanced critical thinking skills, will ensure that teens can more easily identify fake news. However, according to Palacios-López et al. (2023), even if adolescents possess advanced digital skills in instrumental terms, there is still a gap in their critical thinking skills. These authors have found that young people's verification strategies depend on their familiarity with the issuers, including influencers, as well as community validation, rather than a systematic processes of comparison with institutional sources.

2. Influencers and Hate Speech

The architecture of social media and algorithmic configurations aim to obstruct toxic content and the circulation of hateful messages and reactions from followers. Ghaffari (2020) and Weißmann et al. (2025) agree that influencers are especially vulnerable to digital hate due to their hypervisibility and the ambivalent social perception they arouse, which ranges from admiration, to envy, and to moral discrediting. Their lifestyles are often interpreted as excessive, privileged, or artificial, which foments narratives that legitimise hostility (Valenzuela-García et al., 2023).

Among the reasons behind the dissemination of hate speech, Santos Nieto (2024) argues that the heterogeneity of audiences, anonymity, the lack of online inhibition, and the presence of emotion create a field of interaction where polarisation and aggression are integrated into the routines of digital participation. In these hostile environments, female influencers are more vulnerable to these hate messages, especially when they display non-normative standards of femininity. This type of non-mainstream behaviour and discourse elicit misogynistic responses involving physical ridicule, contempt, and sexualisation, as well as the vilification of divergent female performativity (Ghaffari, 2020; Valenzuela-García et al. 2023). According to Weißmann et al. (2025), levels of hate also vary depending on the topics addressed by influencers. Nevertheless, Thomas (2022) indicates that these figures deal with hostility through coping strategies focused mainly on moderating mechanisms using previously established filters, keyword lists, blocks, and automated systems. Other management techniques include self-protection practices based on self-censorship and voluntary withdrawal from platforms, which implies a significant restriction of their participation in digital communities.

Regarding the response of followers to this content, Obermaier (2024) shows that frequent exposure to digital hate provokes ambivalent reactions that range from sensitisation to desensitisation. While some young people develop “digital civil courage” and challenge hate speech, others see hostile discourse as normal, and they simply avoid intervening. According to Cáceres-Zapatero et al. (2023), although adolescents and young people are able to identify the seriousness and consequences of the harm caused by hate speech, passivity is the most common attitude toward this type of content. However, they also assert that people who suffer from malicious discourse are more likely to report or confront such messages. According to Madriaza et al. (2025), an increase in negative attitudes toward the most vulnerable groups, whether due to race, ethnicity, religion, or stereotypes can result in a decline in intergroup trust, which foments social polarisation.

Regarding young people who belong to LGBTI+ groups, they look to influencers as role models in order to reinforce their sense of belonging, explore their identity, and escape situations of discrimination (Sabala, 2020). At the same time, however, these spaces are capable of increasing the vulnerability of these groups, due to the fact that although they are supposedly inclusive, these spaces often perpetuate gay norms, which could generate feelings of exclusion within the community itself (Xu, 2024). According to Fisher et al. (2024) and Sánchez-Sánchez et al. (2024), social media are spaces where LGTBI groups face discourse that is hostile, discriminatory, and watchful. The tactics used combine insults, the discrediting of one’s identity, and explicit types of dehumanisation, which have psychological consequences for users who belong to these gender minorities.

3. Objectives and Methodology

To analyse the informational role of influencers, two analytical approaches have been used. The first focuses on the information factors that integrate the frequency of exposure to current affairs content published by influencers (Wasike, 2025), as well as behaviour related to the verification or avoidance of misinformation (Selnes, 2024). The second approach focuses on the analysis of hate speech, which includes the study of exposure to hostile messages directed at influencers (Weißmann et al. 2025), as well as an examination of the attitude they display toward such content, either active or passive (Obermaier, 2024) (Table 1). In the case of hate speech, the analysis also incorporates gender and sexual orientation, taking into account the self-declared proclivity of the participants, including the following: heterosexual, homosexual (gay or lesbian), bisexual, and other (asexual, not sure yet).

The study addresses Spanish people from 14-24 years of age (mean = 18.5. SD = 2.93). The total sample comprises 1.800 surveys. Weighting techniques were applied to adjust for potential imbalances arising from the sampling process with regard to gender and age. The methodology used was web-assisted interviewing through online panels (Web Panel). Of the sample, 50.2% were male (n = 904), and 49.8% female, with a difference of 0.4% between the two groups, confirming that the sample is balanced. The distribution of the sample shows that 78.3% are heterosexual, 10.7% bisexual, and 4.5% homosexual, while 3.3% were classified as having other

sexual orientations. Those who preferred not to state their orientation stood at 3.1%. Associations between variables were tested using Pearson’s chi-square test of independence. When it was significant, Cramér’s V was reported as an effect-size index to quantify the strength of the association. Spearman’s rank correlation was used to assess associations between ordinal variables without assuming normality.

Table 1: Aspects and variables of analysis

Dimension	Variables	Response options
Information practices: consumption and verification strategies	How often do you keep up with current events through information published by the influencers you follow?	High frequency: 2 or 3 times a day, several times a day. Moderate frequency: once a day. Low frequency: I hardly ever keep up with the news, at least not every day. No frequency: never.
	What behaviour do you use to avoid misinformation?	Item 2.1 = I always verify all the information. Item 2.2 = I consult several sources only when I'm interested in a certain news item. Item 2.3 = I only verify the information when I find it surprising or believe it can't be true. Item 2.4 = I don't verify the news received from the profiles I follow, because I trust them.
Hate speech: reception and behaviour	Through the profiles of influencers, to what extent have you encountered unpleasant, threatening, or hateful comments toward the influencers themselves because of their gender identity, sexual orientation, place of origin, ethnic group, disability, or neurodiversity such as autism, dyslexia, or a similar condition?	High frequency: frequently, continuously. Moderate frequency: occasionally. Low frequency: almost never. No frequency: never.
	When you see unpleasant, threatening or hateful videos or comments on social media for the same reasons, how often do you react in the following ways?	Passive attitude: I observe and do nothing. Active attitude: I share the comment. Each sub-item is measured on a five-point Likert scale grouped by frequency: High frequency: frequently, very frequently. Moderate frequency: occasionally. Low frequency: rarely. No frequency: never.

4. Results

4.1 Information Practices: Consumption and Verification Strategies

The consumption of information disseminated by influencers is a daily and widespread practice among adolescents and young people ($M = 3.53$). Nine out of ten young people (90%) say they consult influencers with various levels of frequency in order to stay informed. Of these, 24.8% consult them frequently, 28% moderately, and 37.3% infrequently. Of the total sample, the proportion of young people who obtain information frequently through these profiles is 11.9% for boys and 12.8% for girls, with a difference of 0.9 percentage points. The internal distribution within each group reveals that 25.7% of girls and 23.8% of boys consult influencer profiles frequently to stay informed. Girls have a slightly higher rate of high-frequency consumption (25.7% vs. 23.8%), while boys report higher average frequency (29.3% vs. 26.7%). In both genders, low frequency is the most common pattern (boys 37.4%, girls 37.2%), and zero frequency stands at around 10% (boys 9.5%, girls 10.5%).

Regarding verification strategies, the most common practice among teenagers and young people is to verify information by consulting several sources, but only if the news arouses their interest (item 2.2). However, this statement was included in slightly less than half of the responses (45.5%, $M = 0.45$). Only one in four always checks the news information (item 2.1; 27.6%, $M = 0.28$), and a similar percentage indicate that they only verify it when they find it surprising or implausible (item 2.3: 28.3%; $M = 0.28$). This suggests that most of the respondents do not use precautionary filters and instead activate verification strategies when warning signs appear. Some 13.4% (item 2.4; $M = 0.13$) do not check the content they receive through the influencers they follow because they trust them, revealing a core of users with high vulnerability and a low level of discernment.

By gender, the habit of consuming information from influencers coexists with verification protocols that are activated mainly due to interest or surprise. In item 2.1 (I always check all the information), the percentage is similar between the genders (boys 28.4%; girls 26.8%). In item 2.2, girls show a stronger tendency to verify news when it arouses their interest (girls 48.9%; boys 42.1%). When they find it surprising or believe it cannot be true (item 2.3), boys slightly exceed girls (29.5% vs 27.1%), reflecting enhanced verification when the content seems

dubious. In item 2.4, 15.1% of the boys and 11.6% of the girls do not verify the information because they trust the influencers to give them accurate news (Table 2).

Regarding high-frequency consumption, among those who consume the most news on influencers' profiles, verification based on interest in the news dominates (item 2.2), with an average of 45.9%, although girls are the ones who verify the most for this reason (51.9% for girls, 39.5% for boys). For item 2.1 (I always verify all information), similar percentages are observed between the genders (32.8% boys; 27.6% girls; the average is 30.1%). This is also true when the information seems surprising or implausible (item 2.3), with an average of 32.1% (33.3% boys; 31.0% girls). Although the percentages are low for not verifying the information received by influencers because the young people trust them (item 2.4), this number is not insignificant, reaching an average of 16.8%, slightly more frequent among boys (17.9%) than girls (15.6%).

Table 2: Consumption frequency and verification strategies

Gender	Frequency	Item 2.1	Item 2.2	Item 2.3	Item 2.4
Boys	High frequency	14.2%	10.3%	18.9%	16%
	Moderate frequency	15.5%	14.4%	12.8%	15.1%
	Low frequency	18.4%	16.9%	18.9%	21%
	No frequency	3.5%	4.7%	4.3%	4.6%
Girls	High frequency	12.8%	14.6%	13.2%	15.1%
	Moderate frequency	12.4%	14.4%	12.4%	11.9%
	Low frequency	18.8%	20.4%	15.9%	14.2%
	No frequency	4.4%	4.3%	3.7%	2.3%

Note: The percentages in Table 2 correspond to the proportion of affirmative responses within the group (gender/frequency), which represent non-exclusive behaviours.

The Chi-square and Cramer's V analyses confirm that the relationship between the frequency of information consumption through influencers, on the one hand, and verification strategies on the other, is weak among all the items. Constant verification (Item 2.1; $\chi^2 = 3.911$, $p = 0.271$), and with a very small effect (Cramer's $V = 0.049$), indicates that incessant verification almost never varies based on the frequency of consumption. Regarding the consultation of multiple sources only when the subject is interested in the news story (Item 2.2), the association is lower ($\chi^2 = 1.736$, $p = 0.629$ and $V = 0.049$), which is consistent with the stability of this practice with regard to frequency ($\approx 42\text{--}49\%$). Item 2.3, regarding the verification of information only when it is surprising or implausible, shows a slightly stronger effect ($\chi^2 = 4.886$, $p = 0.180$ and $V = 0.055$), which indicates that this behaviour increases slightly, although with a weak relationship. Item 2.4, related to not verifying news due to the trust placed in the influencer, shows the highest association ($\chi^2 = 7.000$, $p = 0.072$ and $V = 0.065$), which reveals that complete trust varies more according to frequency, although the difference is not statistically significant.

4.2 Hate Speech: How it is Received on Influencer Profiles and Related Behaviour

Nearly nine out of ten respondents (89.6%) have seen hate speech at least one time on influencer profiles aimed at these digital players, with a high frequency of 37.7%. Girls are exposed to offensive or discriminatory comments more often (42.4%), compared to 33.2% for boys. Boys display higher exposure at moderate frequency (36% vs. 32.4%) and low frequency (20.7% vs. 14.7%). The No Frequency category is similar for both groups (10.2% boys; 10.5% girls), confirming the high prevalence of this phenomenon as a shared experience. The overall mean of exposure to hate speech was $M = 3.12$, with girls at $M = 3.21$, and boys at $M = 3.03$.

The active attitude, or "sharing the comment", has a low incidence among young people. The figure of 52.1% say that they have never done this (no frequency), while 14.8% have done so infrequently, and 18.1% moderately. Only 14.5% say they have shared hate messages frequently. By gender, girls have never shared hate messages (55.1% vs. 48.8%). Boys show higher percentages in the other frequencies: low frequency (18.4% vs. 13.2%), moderate frequency (21.2% vs. 17.6%) and high frequency (14.5% vs. 11.5%), which indicates that although the active dissemination of hate speech is a very limited practice, boys are more involved in it than girls (Figure 1).

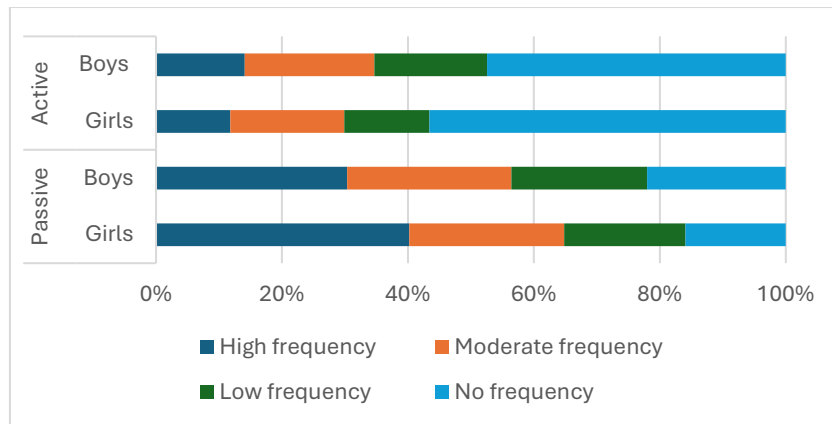


Figure 1: Attitudes toward hate speech by gender

Gender differences in interaction frequency are statistically significant ($\chi^2 = 80.893$; $V = 0.180$), indicating a moderate association, which shows that frequency differs according to gender, with an active attitude being less frequent at low levels, while a passive attitude is more prevalent at higher frequencies, especially among boys. The analysis of adolescents and young people who report having seen hate messages on influencer profiles at high frequencies reveals that in both genders, passive attitudes increase when exposure is high. A total of 50.3% of girls and 28.4% of boys who view them at a high frequency observe them without doing anything, revealing that repeated exposure to hate speech reinforces the normalisation of inaction. On the other hand, active attitudes also increase with the frequency of exposure, but to a lesser extent. 14.5% of boys who view such content frequently share it, compared to 11.5% of girls. These results show that the most prevalent response to hate speech continues to be passivity.

4.2.1 Attitudes Toward Hate Speech According to Sexual Orientation

If we look at sexual orientation, bisexuals (50.8%) and homosexuals (48.6%) are the adolescents and young people who have seen hate speech the most on social media. Although heterosexuals account for 78.3% of the total sample, they report a lower frequency (35.8%), followed by other sexual orientations (32.4%). This indicates that sexual minorities (bisexuals and homosexuals) are more vulnerable, more exposed to hate speech, and more likely to face a hostile environment. Based on sexual orientation, the mean exposure was $M = 3.07$ among heterosexual respondents, $M = 3.55$ for homosexuals, and $M = 3.42$ for bisexuals.

The passive attitude of “watch and do nothing” is the most prevalent in all the groups ($M = 2.93$). With similar percentages, young people of other sexual orientations (39.3%) and homosexuals (39.2%) are the group with the highest rate of passivity, followed by heterosexuals (36.7%) and bisexuals (28%). The active attitude of “sharing the comment” is in the minority ($M = 1.99$), compared to the prevalence of passivity (52.1%), indicating that the majority do not resend hate speech. As for those who never share such content, the rates are 50.7% for heterosexuals, 50.3% for bisexuals, 49% for other orientations, and 45.9% for homosexuals (Figure 2).

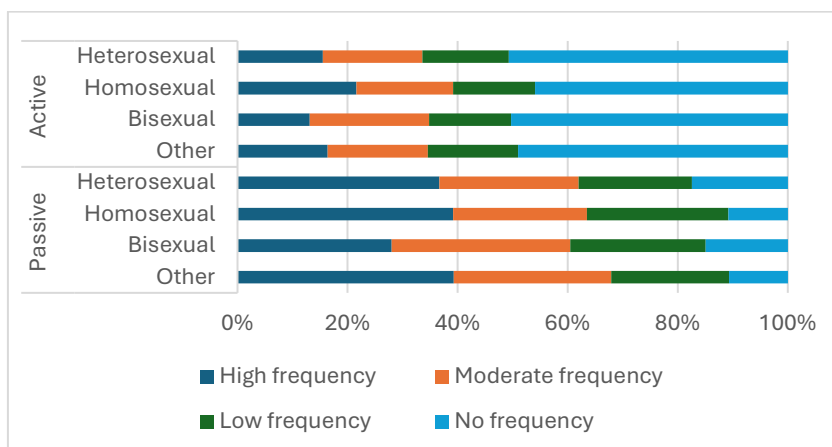


Figure 2: Attitudes toward hate speech by sexual orientation

Based on the results of the inferential analysis of the frequency of hate speech observations ($\chi^2 = 60.226$, $p < 0.001$; $V = 0.092$), the findings indicate that the differences between sexual orientations are minimal, and the effect size is very low. Similarly, a passive attitude ($\chi^2 = 19.028$, $p = 0.390$, $V = 0.062$) shows that inaction is distributed evenly among all groups. The active attitude ($\chi^2 = 12.566$, $p = 0.817$, $V = 0.051$) confirms that the propensity to share hate messages does not vary significantly according to sexual orientation. Moreover, due to the smaller sample size of the bisexual and homosexual groups, these findings reduce the relative weight of the overall contrast. Their estimates display more uncertainty due to their differences being diluted, which causes the aggregate effect to remain small, despite the high percentages of hate speech viewed on influencers' profiles.

Repeated exposure to hate speech does not lead to an active reaction. Instead, it coexists with patterns of inaction, especially among the most exposed groups. Thus, although bisexuals and homosexuals are the groups most frequently exposed to hate speech ($M = 3.55$ and $M = 3.42$, respectively), compared to heterosexuals ($M = 3.07$), the prevalent response is still not to share them (percentages for other orientations: 39.3%, homosexuals 39.2%, heterosexuals 36.7% and bisexuals 28%). Regarding the active attitude of sharing content, those who show the highest frequency are young people of other sexual orientations (20.1%), followed by heterosexuals (15.5%), bisexuals (15.2%), and finally, homosexuals (9.6%).

5. Discussion and Conclusions

This research confirms that influencers have become the preferred source of news, especially among adolescents and young people. Of this group, nine out of ten ($M = 3.53$) say they consume information from influencers with some level of frequency, and nearly five out of ten do so regularly. However, this consumption is not accompanied by preventive protocols for verifying misinformation. The efforts of these young people to verify information depend mostly on whether they are interested in the news item in question, and whether they find a report either surprising or implausible. As they do not use many verification strategies, the role of influencers as nodes of authority is being consolidated, and young people's criteria for assessing and substantiating content is being redefined (Niitsuma et al., 2025; Shin et al. 2022). These selective practices increase the risk of consuming misinformation, which is also mediated by the use of algorithms. Gender differences are modest, as confirmed by inferential tests using Cramer's V values, which have yielded rates from 0.049–0.087. However, asymmetries in exposure to misinformation have also been observed. Girls are more likely to verify information when they are interested in a particular news item, while boys tend not to check the information when they trust the influencer who has sent it, and they mainly verify news when they believe the content is dubious.

Regarding hate speech, its level of exposure ($M = 3.12$) reflects a media environment in which contact with hostile messages is common. Although girls observe hate speech more often than their male counterparts, in terms of agency, boys share them more frequently (14.5% vs. 11.5%). This indicates that males are more willing to reproduce and spread hostile content, which might have an impact on tolerance the propagation of hate speech. It has been observed that repeated exposure to hostile messages does not result in an active response. Instead, such exposure consolidates a pattern of passivity that legitimises the phenomenon of malicious comments and perpetuates the dissemination of hate speech (Cáceres-Zapatero et al., 2023). Thus, the *online bystander effect* has been confirmed, in which algorithms are redefining the agency of young audiences, displacing them toward passive forms of participation when they experience insidious content (Obermaier, 2024).

A pattern of inequality is evident in the digital experience of adolescents and young people based on their sexual orientation. Although bisexual and homosexual individuals comprise a minority of the sample (14.9%), they have the highest rates of exposure to violent messages and hate speech on influencer profiles, which confirms the persistence of dynamics of vulnerability associated with non-normative identities (Sadaba, 2020; Fisher et al., 2024). Sexual minorities are an especially vulnerable audience, as they must confront communicative scenarios that consolidate a hostile ecosystem, which affects their well-being in the cyber sphere (Sánchez-Sánchez et al., 2024). Although the response to such hostility is mostly passive, it is not homogeneous. Young people of other sexual orientations are the ones who share malignant messages the most, followed by heterosexuals. By contrast, homosexuals are the least likely to share vicious comments, which indicates that this group is exceptionally vulnerable to hate speech, and they act as a risk marker. However, in the most exposed groups, the dominant response continues to be observation without action, which bolsters the idea that repeated exposure to hate speech does not lead to defiance; instead, it consolidates inaction as a common practice when faced with digital violence.

The limitations of this research are related to the quantitative methodological design, which does not allow causal relationships to communicative behaviour of adolescents and young people on the other. Moreover, although the sample is large and representative of adolescent and young users, as its sociocultural origin is limited to one specific nation, extrapolation of the results to other countries and cultures is impeded. For future research, regarding the degree of agency of young people, it would be useful to develop longitudinal studies to explore the evolution regarding the consumption of content created by influencers, with a focus on addressing these personalities as epicentres of problematic content that might encourage the viralisation of fake news and hate speech.

Ethics Declaration

This study complies with the European General Data Protection Regulation (GDPR; Regulation (EU) 2016/679), which was enacted in 2016 and has been in force since 25 May 2018. It also adheres to Spain's Organic Law 3/2018 of 5 December on the Protection of Personal Data and the Guarantee of Digital Rights. Ethical approval of this investigation was obtained from the Research Ethics Committee at Rey Juan Carlos University, registration number 270520243382024. All procedures related to this research have complied with the principles of the Declaration of Helsinki and the applicable EU data-protection framework.

AI Declaration

During the preparation of this work, the authors used ChatGPT Plus to enhance the clarity of the writing. After using this tool, the authors reviewed and edited the content as needed. The authors take full responsibility for the content of the manuscript.

Funding

This work was supported by the Media repertoires and practices in adolescence and youth: uses, cyber-wellbeing and digital vulnerabilities in social networks [PID2022-138281NB-C21]; The Ministry of Science, Innovation and Universities [MCIN/AEI/10.13039/501100011033]; ERDF A way of making Europe.

References

- Cáceres-Zapatero, M.D., Brändle, G., Paz-Rebollo, M.A. (2023). "Stances on hate speech: Population opinions and attitudes". *Profesional de la información*, 32(4), e320410. <https://doi.org/10.3145/epi.2023.jul.10>
- Fisher, C.B., Tao, X., Ford, M. (2024). "Social media: A double-edged sword for LGBTQ+ youth". *Computers in Human Behavior*, 156, 108194. <https://doi.org/10.1016/j.chb.2024.108194>
- Ghaffari, S. (2023). "Discourses of celebrities on Instagram: digital femininity, self-representation and hate speech". In: KhosraviNik, M. (ed.) *Social Media Critical Discourse Studies*. Routledge, London, pp. 43–60.
- Madriaza, P., Hassan, G., Brouillette-Alarie, S., Mounchingam, A.N., Durocher-Corfa, L., Borokhovski, E., Pickup, D., Paillé, S. (2025). "Exposure to hate in online and traditional media: A systematic review and meta-analysis of the impact of this exposure on individuals and communities". *Campbell Systematic Reviews*, 21(1), e70018. <https://doi.org/10.1002/cl2.70018>
- Mulcahy, R., Barnes, R., de Villiers Scheepers, R., Kay, S., List, E. (2024). "Going viral: Sharing of misinformation by social media influencers". *Australasian Marketing Journal*, 33(3), pp. 296–309. <https://doi.org/10.1177/14413582241273987>
- Niitsuma, T., Yoshida, M., Tamori, H., Nakawake, Y. (2025). "Prestige bias drives the viral spread of content reposted by influencers in online communities". *Scientific Reports*, 15(1), p. 15282. <https://doi.org/10.1038/s41598-025-98955-4>
- Obermaier, M. (2024). "Youth on standby? Explaining adolescent and young adult bystanders' intervention against online hate speech". *New Media & Society*, 26(8), pp. 4785–4807. <https://doi.org/10.1177/14614448221125417>
- Orhan, A. (2023). "Fake news detection on social media: the predictive role of university students' critical thinking dispositions and new media literacy". *Smart Learning Environments*, 10(29), p. 1-14. <https://doi.org/10.1186/s40561-023-00248-8>
- Palacios López, M., Bonete Vizcaíno, F., Gelado Marcos, R. (2023). "New agents of mass disinformation. Analysis of the publications produced by Spanish political influencers on Instagram". *Estudios sobre el Mensaje Periodístico*, 29(4). <https://doi.org/10.5209/esmp.88576>
- Sabala, S. (2020). *The perception of LGBTQ influencers on social media*. Rochester Institute of Technology, ProQuest Dissertations & Theses.
- Sánchez-Sánchez, A.M., Ruiz-Muñoz, D., Sánchez-Sánchez, F.J. (2024). "Mapping homophobia and transphobia on social media". *Sexuality Research and Social Policy*, 21(1), pp. 210–226. <https://doi.org/10.1007/s13178-023-00879-z>

- Santos Neto, V.S. dos (2024). "Criadores de Conteúdo e a Ascensão dos Discursos de Ódio entre Brasileiros e Portugueses em Plataformas Digitais: Análise de um Experimento em Rede Social". *Revista de Comunicação, Cultura e Artes*, 4(1), pp. 80–97. <https://doi.org/10.34623/z2g4-rj46>
- Selnes, F.N. (2024). "Adolescents' experiences and (re) action towards fake news on social media: perspectives from Norway". *Humanities and Social Sciences Communications*, 11(1), pp. 1–11. <https://doi.org/10.1057/s41599-024-04237-1>
- Shin, I., Wang, L., Lu, Y.T. (2022). "Twitter and endorsed (fake) news: The influence of endorsement by strong ties, celebrities, and a user majority on credibility of fake news during the COVID-19 pandemic". *International Journal of Communication*, 16, pp. 2573–2595.
- Thomas, K., Kelley, P.G., Consolvo, S., Samermit, P., Bursztein, E. (2022). "It's common and a part of being a content creator: Understanding how creators experience and cope with hate and harassment online". In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 121, pp. 1–15. <https://doi.org/10.1145/3491102.350187>
- Valenzuela-García, N., Maldonado-Guzmán, D.J., García-Pérez, A., Del-Real, C. (2023). "Too lucky to be a victim? An exploratory study of online harassment and hate messages faced by social media influencers". *European Journal on Criminal Policy and Research*, 29(3), pp. 397–421. <https://doi.org/10.1007/s10610-023-09542-0>
- Wasike, B. (2025). "Influencing the news: how social media influencers affect news media trust". *Information, Communication & Society*, pp. 1–18. <https://doi.org/10.1080/1369118X.2025.2496907>
- Weißmann, S., Philipp, A., Verwiebe, R., Krauter, C.O., Fritsch, N.S., Buder, C. (2025). "Clicks, comments, consequences: Are content creators' socio-structural and platform characteristics shaping the exposure to negative sentiment, offensive language, and hate speech on YouTube?". *arXiv preprint*, arXiv:2504.07676. <https://doi.org/10.48550/arXiv.2504.07676>
- Xu, Y. (2024). "A study of LGBTQ+ empowerment and normative dialogue in Chinese social media platforms". *Transactions on Social Science, Education and Humanities Research*, 11, pp. 1–10. <https://doi.org/10.62051/952bn252>