



Self-injury and its relationship with factors associated with social media use: a systematic review

Karen Duarte-Tánori & Daniel Fregoso-Borrego*

Universidad de Sonora, Hermosillo (Mexico)

KEYWORDS

Internet
Injuries
Internet analysis
Instagram
TikTok

ABSTRACT

Nonsuicidal self-injury is a relevant mental health issue that, in the current digital context, has acquired new forms of expression, exposure, and interaction through social media. The aim of this study was to analyze recent evidence on the relationship between non-suicidal self-injury and the use of social media. The Preferred Reporting Items for Systematic Reviews and Meta-analyses guidelines were applied for the search and selection of articles in the SCOPUS, EBSCOhost, and ProQuest databases, using inclusion criteria that considered studies in English and Spanish. A total of 13 articles were analyzed, mostly employing mixed and cross-sectional designs focused on platforms such as X, Instagram, TikTok, YouTube, and WhatsApp. Results showed that social media can function as a risk factor by promoting exposure and normalization of self-injury, but also as spaces for support and emotional containment. It is concluded that digital interaction constitutes an ambivalent environment requiring more effective moderation policies and preventive strategies focused on mental health and digital literacy, as well as studies testing models to clarify the digital space.

Autolesión y su relación con factores asociados al uso de redes sociales: revisión sistemática

PALABRAS CLAVE

Internet
Lesión
Análisis de Internet
Instagram
TikTok

RESUMEN

La autolesión no suicida constituye una problemática relevante de salud mental que, en el contexto digital actual, ha adquirido nuevas formas de expresión, exposición e interacción a través de las redes sociales. El objetivo fue analizar la evidencia reciente sobre la relación entre autolesión no suicida y el uso de redes sociales. Se aplicó el Preferred Reporting Items for Systematic Reviews and Meta-analyses para búsqueda y selección de artículos en las bases SCOPUS, EBSCOhost y ProQuest, utilizando criterios de inclusión de estudios en inglés y español. Finalmente, se analizaron 13 artículos con diseños principalmente mixtos y transversales, centrados en plataformas como X, Instagram, TikTok, YouTube y WhatsApp. Los resultados mostraron que las redes sociales pueden funcionar como factores de riesgo, al favorecer la exposición y normalización de la autolesión, pero también como espacios de apoyo y contención emocional. Se concluye que la interacción digital constituye un entorno ambivalente, que requiere políticas de moderación más eficaces y estrategias preventivas con enfoque en salud mental y alfabetización digital, así como estudios que prueben modelos para esclarecer el espacio digital.

* *Corresponding author:* Daniel Fregoso-Borrego. Faculty of social sciences, department of psychology and communication sciences, Universidad de Sonora, Blvd. Luis Encinas J, Av. Rosales & Centro, PC. 83000, Hermosillo, Sonora. daniel.fregoso@unison.mx

Cite this article as: Duarte-Tánori, K., & Fregoso-Borrego, D. (2026). Self-injury and its relationship with factors associated with social media use: a systematic review. *Psychology, Society & Education*, 18(3), 34-46. <https://doi.org/10.21071/pse.v18i3.19026>

Received: 16 February 2026. *First review:* 27 March 2026. *Accepted:* 1 September 2026.

Psychology, Society & Education is published under Creative Commons License ([CC BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/)).

ISSN 1989-709X | © 2026. Psy, Soc & Educ.



Nonsuicidal self-injury (NSSI) is a repetitive behavior intended to regulate emotional situations perceived as difficult to cope with (Chartrand et al., 2015; Nock, 2010; Santos, 2011). It is important to distinguish the concept of NSSI from the term self-harm. NSSI refers to a phenomenon that differs from situations involving, for example, completed suicide or suicidal ideation, since in NSSI it is possible to prevent irreversible consequences, whereas suicidal ideation often involves specific factors of a different nature (Gillies et al., 2018). In addition, the latter term is used heterogeneously and may include behaviors involving suicidal intent or ideation (Gillies et al., 2018). The interchangeable use of these terms may create ambiguity in the interpretation of findings. Thus, the terminology used depends on the specific phenomenon that a study seeks to address. Accordingly, the present study focused on NSSI, as it is particularly relevant to further examine how self-injury manifests and which associated factors may prevent or promote it.

NSSI is currently recognized as a public health concern due to its high prevalence in the general population and, particularly, among adolescents and young adults. According to Muehlenkamp and Gutiérrez (2007), the global prevalence of NSSI among adolescents ranges from 13% to 32%. These figures suggest that self-injury is not an isolated phenomenon but rather a growing behavioral problem with significant clinical and social implications. NSSI has been found to be associated with personal factors such as depression, anxiety, emotion regulation, and disorders such as borderline personality disorder, as well as with social factors including attachment to mothers and fathers and the use of the Internet and social media (Duarte & Fregoso, 2024; Duarte et al., 2020; Schmidt et al., 2023).

Within this context, it is important to consider the digital environment as another contextual factor that has received relatively limited attention. Technological development and the proliferation of social media have created new spaces for socialization that play an important role in the lives of adolescents and young people (Rojas-Jara et al., 2018). Platforms such as Instagram, TikTok, X (formerly Twitter), and Facebook serve as spaces for everyday interaction and the dissemination of experiences related to mental health (Grant-Allen et al., 2025; La Sala et al., 2025; Woodward et al., 2025). Although these platforms can provide social support and reduce stigma (Robinson et al., 2016), exposure to explicit self-injury content may contribute to the normalization, imitation, and social contagion of such behaviors (Grant-Allen et al., 2025; Lewis et al., 2014). Research has shown that recommendation algorithms may amplify exposure to graphic or narrative content about self-injury, thereby increasing the vulnerability of certain users (Arendt et al., 2019; Grant-Alen et al., 2025). At the same time, participation in online communities may foster emotional support and a sense of belonging among young people, potentially exerting protective effects against psychological distress (Cliffe et al., 2024). This ambivalence makes social media a complex field of study, as it may function both as a risk factor and as a coping resource.

Self-Determination Theory (SDT) addresses relatedness as a basic psychological need. Within this framework, this need

is satisfied when individuals experience connection, acceptance, and reciprocity in their relationships. Therefore, a sense of belonging can be understood as an expression of relatedness, whereas perceived social support constitutes an interpersonal process closely associated with its satisfaction (Ryan et al., 2021). For example, a recent study conducted from an SDT perspective found that these environments may either satisfy or frustrate the need for relatedness depending on the quality of interactions and the context, as creating and maintaining social connections, sharing identities and experiences, and providing reciprocal support and care represent mechanisms of relatedness that may fit within this theoretical framework (West et al., 2024). This theoretical perspective is particularly relevant to NSSI because individuals who experience stigmatized circumstances may find others with similar experiences and opportunities for interaction within online communities that are not always available offline. Thus, from an SDT perspective, it is not social media use itself that generates support or belonging, but rather the quality, reciprocity, and meaning of the interactions enabled by these platforms. Therefore, NSSI should be examined from a perspective that integrates both personal factors and the digital contexts in which adolescents and young people participate in their everyday lives. Although the literature has documented associations with emotional, family-related, and psychopathological variables, further research is still needed to understand the role of social media as spaces that may amplify exposure to content that promotes self-injury while also providing social support and fostering a sense of belonging. In this regard, the present study is relevant because it provides evidence on factors associated with NSSI while considering the interaction between individual conditions and contemporary digital environments, which may contribute to the development of prevention and intervention strategies that are better aligned with current forms of youth socialization.

The present study

Given the magnitude of NSSI and the relevance of digital environments in the lives of adolescents and young people, the present study aimed to conduct a systematic review of the literature published between 2019 and 2025 on the role of social media use as a potential risk or protective factor for NSSI, with the purpose of integrating recent findings and informing future lines of research and intervention. The review was guided by the following question: What evidence published between 2019 and 2025 indicates whether social media use is associated with increased risk or with potential protective effects regarding NSSI among adolescents and young people? Additionally, three analytical questions were proposed: What characteristics of social media use have been associated with a higher risk of NSSI? What mechanisms of influence have been described? And what conditions or patterns of use may operate as protective factors?

Method

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method was used to conduct the search and present the results. This method indicates that at least the following eight elements should be reported: 1) authors and year, 2) objectives, 3) participants and location, when reported, 4) variables involved, 5) study design, 6) analyses conducted, 7) relevant results, and 8) authors' conclusions (Page et al., 2021). The databases used to conduct the comprehensive search were SCOPUS, EBSCOhost, and ProQuest. These databases were selected because they are considered high-impact sources and provide access to publications in several languages, including Spanish and English (Packer, 2024). The search and data collection process was conducted from October 6 to November 28, 2025. Access to all articles was available because the search was conducted through an institutional account, as the university provides access to these services.

Search criteria

The terms “autolesión”, “autolesión no suicida”, “redes sociales”, “X”, “Instagram”, “Facebook”, “TikTok”, “WhatsApp”, and “Twitch” were used for searches in Spanish, whereas “non-suicidal self-injury”, “self-injury”, “non-suicidal self-harm”, “self-harm”, “social network”, “X”, “Twitter”, “Instagram”, “Facebook”, “TikTok”, “WhatsApp”, and “Twitch” were used for searches in English. Terms referring to self-injury were separated using the Boolean operator “O” or “OR,” depending on whether the search was conducted in Spanish or English, as were the terms referring to social interaction in online digital media. In turn, the Boolean operator “Y” or “AND,” depending on whether the search was conducted in Spanish or English, was used to combine terms referring to self-injury with those referring to interactions through online digital media.

The search string for SCOPUS was: TITLE-ABS-KEY ([“Self-injury” OR “non-suicidal self-injury” OR “self-harm” OR “non-suicidal self-harm”] AND [“social network” OR “X” OR “Instagram” OR “Facebook” OR “TikTok” OR “WhatsApp” OR “Twitch”]). For EBSCOhost, the search string was: TI (“self injury” OR “self harm” OR “non-suicidal self-injury” OR “non suicidal self harm” OR “autolesion” OR “Autolesion no suicida”) AND TI (“social network” OR “redes sociales” OR “instagram” OR “WhatsApp” OR “Facebook” OR “Twitch” OR “X” OR “Twitter”). Finally, for ProQuest, the search string was: (“self injury” OR “self harm” OR “non-suicidal self-injury” OR “non suicidal self harm” OR “autolesion” OR “Autolesion no suicida”) AND TI (“social network” OR “redes sociales” OR “instagram” OR “WhatsApp” OR “Facebook” OR “Twitch” OR “X” OR “Twitter”). It is important to note that searches were conducted using both “X” and “Twitter,” since the social media platform was renamed in 2023.

Search filters applied

The filters applied included the publication period (2019-2025), peer-reviewed journal articles, publications in English

and Spanish, and relevant disciplines (psychology, sociology, medicine, multidisciplinary fields, or related areas). The search was conducted using titles, abstracts, and keywords. Specifically, for SCOPUS, the following filters were applied: publication years from 2019 to 2025; subject areas including psychology, social sciences, computer science, and multidisciplinary studies; document type restricted to “article”; and language restricted to Spanish and English. For EBSCOhost, the filters applied were “last 6 years” (2019 to November 2025), academic publications, and peer-reviewed academic publications. For ProQuest, the “last 6 years” filter was applied.

Exclusion criteria

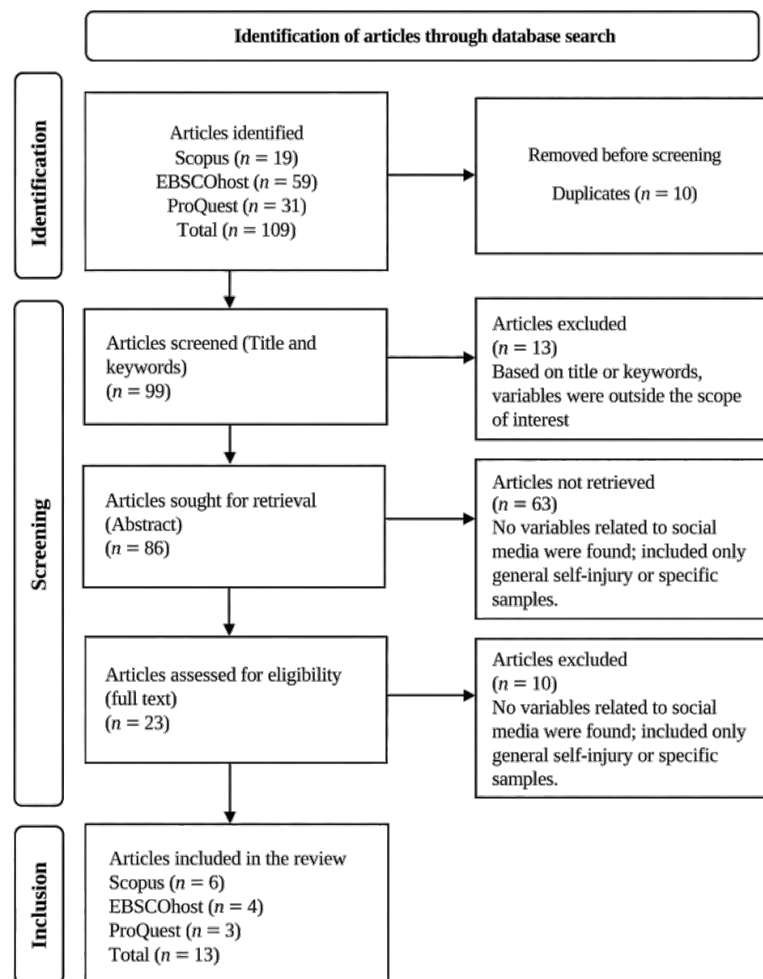
Articles with highly specific samples were excluded, for example, those focused on particular conditions such as autism spectrum disorder, single-parent families, membership in a minority group, or the presence of a disease or medical condition. This decision was based on the conceptual delimitation of the phenomenon under study. Including studies focused exclusively on clinical subgroups could lead to an overrepresentation of contexts involving specific vulnerabilities and bias the interpretation of the findings regarding social media use and how the problem manifests in the general population. Duplicate articles were also excluded. Duplicates were identified by organizing the records in Excel and entering the title, authorship, and year of publication of each retrieved article; duplicates were then detected using the “Remove duplicates” tool. Likewise, studies were excluded if the concept of self-injury did not correspond to the definition presented in the introduction or if they did not include social media use. Publications consisting of documentary reviews, such as other systematic and bibliometric reviews, were also excluded. This screening process was conducted independently by two reviewers. At each stage of the flow diagram, both reviewers identified the articles to be excluded and reached 100% agreement; therefore, no discrepancies arose (Figure 1).

Methodological quality assessment

To assess the methodological quality of the included studies, the Mixed Methods Appraisal Tool (MMAT; Hong et al., 2018) was applied, which allows qualitative, quantitative, and mixed-methods studies to be evaluated within systematic reviews. The assessment was conducted using the five specific criteria proposed according to the type of methodological design reported in each study. These criteria examine aspects such as the appropriateness of the design, sampling quality, suitability of the measurement instruments, consistency between data and analyses, as well as the validity of the results and findings. Each article was independently assessed by two reviewers and broadly classified as being of high, moderate, or low quality, with the aim of contextualizing the strength of the synthesized evidence.

Figure 1

Flow diagram showing the article retrieval process for analysis based on PRISMA

**Table 1**

Number of articles retrieved from the databases consulted

Base	Initial frequency	Final frequency	Percentage (initial-final)
SCOPUS	19	6	32%
EBSCOhost	59	4	7%
ProQuest	31	3	10%
Total	109	13	12%

Results

Table 1 shows the number of articles ultimately retrieved, as well as the percentage relative to the initial search. SCOPUS yielded the highest percentage, followed by ProQuest and, lastly, EBSCOhost. Although EBSCOhost produced the highest initial number of records ($n = 59$), it was the database from which the fewest articles were ultimately retrieved in percentage terms.

The most relevant findings from each study were then synthesized and presented, considering both the associated factors and their characteristics. Table 2 provides an overview of several key elements used to describe each article, including

authors, year, location, type of study, among others. The most frequently analyzed social media platforms were X, Instagram, TikTok, YouTube, and WhatsApp, where interactions related to self-injury were concentrated. The most commonly studied samples consisted of adolescents and young adults (11 to 49 years of age), although some studies analyzed digital content directly rather than human participants (Kobilke & Markiewitz, 2022; Lavis & Winter, 2020; Lim et al., 2022; Martínez-Pastor et al., 2023; Scherr et al., 2019). Mixed-methods and cross-sectional designs predominated, with descriptive, exploratory, or correlational approaches aimed at identifying risk factors, interaction dynamics, and manifestations of self-injury (Table 2).

Table 2*Basic description of each article*

Database	Author (year)	Location	Sample characteristics	Design	Variables	Tools
SCOPUS						
	Alsobai et al. (2024)	United States	173 adolescents (13–21 years; $M = 17$; $SD = 2.15$); 67% women, 23% men, 10% nonbinary. 33,469 Instagram conversations	Mixed-methods, cross-sectional, correlational. MMAT quality: High	Offline and online risk experiences	Risk Behavior Questionnaire; Inventory of Statements About Self-Injury; Cyber-Aggression Victimization and Cyber-Aggression Perpetration; Unwanted Sexual Solicitations and Approaches and Youth-Produced Sexual Images (Sexting). Instagram conversation records
	Atauri-Mezquida et al. (2025)	Spain	11,749 young X users	Mixed-methods, cross-sectional, descriptive. MMAT quality: Moderate	Content, sensitivity, self-injury in content, interaction	Tweet records and automated X classification
	Kobilke & Markiewicz (2022)	Not specified	209 YouTube videos. Content and audience of 105 videos	Mixed-methods, explanatory, cross-sectional. MMAT quality: Moderate	Harmful representations, helpful representations, interaction with videos, video genre	Content records
	Lookingbill & Le (2024)	United States	8 TikTok users, 150 videos	Mixed-methods, cross-sectional, descriptive. MMAT quality: High	Experience with the TikTok algorithm; avoidance strategies; information practices; perceptions of moderation; representation of NSSI	Semi-structured interview. Content records using the hashtags #sh and #scars
	Seong et al. (2021)	South Korea	902 adolescents. 11–15 years old with a history of NSSI. 55.1% men, 44.9% women	Quantitative, cross-sectional, explanatory. MMAT quality: High	Self-injury, social circumstances, suicidal ideation and history	Self-injury assessment; suicidal history and suicidal ideation; behavioral characteristics of self-injury; social circumstances
	Szlyk et al. (2023)	Not specified	93 adolescents (15–17 years; $M = 16.2$); 73% women, 45% non-White, 57% sexual minorities, 80% had engaged in NSSI at some point in their lives, 65% had suicidal ideation, and 58% had attempted suicide	Mixed-methods, cross-sectional, descriptive-comparative. MMAT quality: Moderate	Suicidal behaviors; NSSI; interaction with depressive content; perceived advantages and disadvantages; emotional impact on Instagram, Facebook, X, Reddit, and Tumblr	Open-ended questions. Patient Health Questionnaire-9

Table 2 (continued)

Database	Author (year)	Location	Sample characteristics	Design	Variables	Tools
EBSCO-host	Lavis & Winter (2020)	United Kingdom	10,169 posts. 36,934 comments, Reddit. 10 women aged 10 to 24 years	Qualitative, digital ethnography, descriptive, cross-sectional. MMAT quality: High	Motivations for seeking self-injury content, peer support, emotional impact, use of graphic images, protective mechanisms	Semi-structured interview, ethnographic observation
	Martínez-Pastor et al. (2023)	Not specified	1,951 messages on X with the hashtags #selfharm, #shtwt, ouchietwt. Including likes, retweets, replies, and profiles	Quantitative, descriptive, cross-sectional. MMAT quality: High	Interaction on X, type of user, content	Data mining
	Record et al. (2020)	United States	417 university students with Instagram. 77% women and 23% men, aged 18 to 49 ($M = 20.7$; $SD = 3.2$)	Quantitative, explanatory, cross-sectional. MMAT quality: Moderate	Reporting knowledge, attitude toward reporting, subjective norm, behavioral control, intention	Instagram use; reporting knowledge; attitudes, subjective norms, behavioral control, and intention
	Scherr et al. (2019)	Not specified	13,132 images on Instagram with the hashtags #suicide, #cutting, #selbstmord and #ritzen, #	Quantitative, comparative, cross-sectional. MMAT quality: High	Presence of NSSI, hashtag language, relative risk	AlexNet algorithm; scraping
ProQuest						
	Latina et al. (2023)	Germany	35 adolescents with a mean age of 15.16 ($SD = 1.48$). 24 women and 11 men. 23 with a clinical condition	Quantitative, experimental, explanatory, cross-sectional. MMAT quality: Moderate	Heart rate; heart rate variability index; experience of ostracism; negative and positive affect	Self-report of positive affect and negative affect. WhatsApp tasks using Simulated On-Line Ostracism (SOLO) and Cyberball
	Lim et al. (2022)	Singapore	35,904 tweets about suicide, self-injury, and COVID-19	Qualitative, descriptive, cross-sectional, observational. MMAT quality: High	Themes related to suicide, self-injury, and COVID-19. Date, continent, number of Instagram followers	Topic modeling using STM
	Peart et al. (2023)	New Zealand	35 transcripts of conversations between counselors and young people aged 13 to 25 years ($M = 16.7$)	Qualitative, exploratory, cross-sectional. MMAT quality: High	Counseling, suicide risk, emojis, validation, confidentiality, empowerment	Transcription; interview

Note. MMAT = Methodological quality assessment according to the Mixed Methods Appraisal Tool (rating: low, moderate, or high).

Table 3*Main findings for each article organized by thematic categories*

Thematic category / Authorship	Objectives	Analyses conducted	Results	Findings
1. Algorithmic amplification				
'Atauri-Mezquida et al. (2025)	Analyze hashtags related to NSSI on X	SQL queries; MySQL; frequencies; manual review of sensitivity and blocking; comparison with X policies	8.3% of tweets were blocked; 35.25% were classified as “sensitive”; inconsistency between X classification and effective blocking	Content moderation on X is inconsistent and ineffective; self-injury is normalized and romanticized; content can be accessed and posted without restrictions; X does not consider text to be sensitive
'Kobilke & Merkwitz (2022)	Assess the extent to which YouTube videos about the “Momo challenge” present harmful or helpful representations	Descriptive statistics; bivariate correlations; multivariate regressions; modularity analysis and network visualization (Gephi)	68% of videos contain harmful representations; the danger index was correlated with likes ($r = .4, p < .001$), comments ($r = .37, p < .001$), dislikes ($r = .36, p < .001$), and views ($r = .36, p < .05$). Only the index significantly predicts comments, likes, and dislikes ($R^2 = .14; R^2 = .15; R^2 = .12$, respectively). Seven communities were identified	Harmful representations are common and generate greater interaction; YouTube moderation was delayed and inconsistent; the most active communities prefer more sensationalist and risky content
Lookingbill & Le (2024)	Explore how TikTok users evade algorithmic barriers	Thematic and content analysis	“Algospeak” (algorithm), implicit signaling, and self-censorship are used to avoid moderation; TikTok moderates inconsistently; the NSSI community self-regulates to reduce harm	Algorithmic moderation reinforces stigmatization; “folk” theories about the algorithm are developed; the community creates internal norms of care and signaling
Scherr et al. (2019)	Not specified	Algorithm training. Relative risk analysis	87% accuracy in detecting NSSI in images. The risk of viewing images containing NSSI is higher with German hashtags than with English hashtags (+38.83%)	The algorithm detects cutting with high accuracy; hashtag language influences the risk of exposure; German hashtags (#ritzen) show a higher prevalence of graphic content
2. Social reinforcement and normalization of self-injury on social media				
Martínez-Pastor et al. (2023)	Analyze communicative interactions between users who post or view explicit images of self-injury on X	Social Network Analysis, nodes, manual review of messages	Likes are the most frequent interaction with messages under #shtwt; like networks are more cohesive for #self-harm, #shtwt, and ouchietwt (.0018, .0004, .0056, respectively) than retweets or replies	The hashtags #shtwt and #ouchietwt reinforce self-injurious behaviors; likes act as reinforcement; the most highly valued messages show explicit injuries. X acts as a space for reinforcing self-injurious behaviors

Table 3 (continued)

Thematic category / Authorship	Objectives	Analyses conducted	Results	Findings
3. Exposure to self-injury content and psychosocial effects				
³ Lavis & Winter (2020)	Explore the benefits and risks of peer support on social media surrounding self-injury. Question the contagion model	Thematic analysis of textual and visual content; triangulation across platforms and interviews; categorization by age and context	Peer support is central; young people were already engaging in self-injury before seeking content; graphic images function as signals for help; content can be protective or harmful depending on the context 7,503 (20.9%) tweets were related to self-injury. Six main themes were identified: 1) public policy responses; 2) biopsychosocial impact on self-injury; 3) comparison of mortality rates; 4) mental health support; 5) reported cases and public reactions; and 6) figurative use of the word “suicide”	The contagion model is insufficient; self-injury content is linked to prior experiences and the search for understanding; support can save lives, but it can also generate emotional burden Increase in self-injury during the lockdown period. Relationship between self-injury and substance abuse; tweets requesting help from people considering suicide or self-injury. More cases among adolescents and young people than among adults
Lim et al. (2022)	Examine how people relate to suicide and self-injury during the COVID-19 pandemic from a global perspective through public discourse on X	Topic modeling by content, date, and followers. Discourse analysis and thematic content analysis. Categories	Adolescents with NSSI report more negative repercussions when posting; those who attempted suicide mention more supportive actions in response to concerning content; those with suicidal ideation report more unhelpful social interactions	Interaction with depressive content varies according to the history of suicidal behavior; risks and opportunities for social media outreach were identified; differences were found in perceptions of support, emotional impact, and willingness to participate in research
² Szlyk et al. (2023)	Examine how interaction with depressive content on social media varies among adolescents with and without suicidal behaviors and NSSI. Explore opinions regarding social media outreach	Inductive/deductive thematic coding; reliability calculation ($\kappa = 0.52-0.75$); bivariate analysis (χ^2)		
4. Pre-existing vulnerabilities and associated social factors				
Alsobai et al. (2024)	Identify risk profiles according to offline and online experiences. Correlate profiles with Instagram information. Analyze linguistic differences in conversations	Mixed Factor Analysis, comparative analysis, X ² , SAGE, content analysis	Five profiles were identified: low risk (51%), medium risk (29%), increased sexting (8%), increased self-injury (8%), and high-risk perpetration (4%). Significant correlations were found between the identified profiles and risk flags in conversations	Offline experiences manifest differently online; sexting and self-injury are significantly interrelated; high-risk profiles show promotion of illegal activities and violence

Table 3 (continued)

Thematic category / Authorship	Objectives	Analyses conducted	Results	Findings
Latina et al. (2023)	Evaluate SOLO through WhatsApp to examine ostracism among adolescents and its effect on physiological and emotional reactions	Multilevel analysis with interaction assessment	The clinical group showed a higher heart rate during SOLO ($b = 4.62$; $p < .05$, $d = .35$), and greater negative affect ($\beta = -0.46$; $p < .001$). Lower positive affect was observed (clinical: $\beta = -0.71$; $p < .001$; control: $\beta = -1.33$; $p < .001$)	The SOLO task design generated greater emotional distress and physiological activation in the clinical group; it was more ecologically valid than Cyberball
Seong et al. (2021)	Explore how social and behavioral characteristics explain suicidal behavior among adolescents with self-injury, with emphasis on the role of social media	Multivariate logistic regression	Having friends who engage in self-injury ($\beta = 1.92$; $p < .05$), school bullying ($\beta = 2.08$; $p < .01$), and posting on social media ($\beta = 3.15$; $p < .001$) predict self-injury and suicidal ideation	Self-injury begins before the age of 11; posting content on social media is associated with self-injury and suicidal ideation; social media may amplify vulnerabilities
5. Social media as spaces for support and intervention				
Peart et al. (2023)	Explore how counselors engage with young people in a proactive digital intervention and how they conduct suicide risk assessment	Reflexive thematic analysis; triangulation with user feedback; iterative coding and thematic categorization	Seven themes emerged: use of microskills, trust-building, permission to discuss suicide, structured approach, confidentiality, validation, and identification of interventions	Counseling skills can be adapted to the digital environment; the proactive approach enables meaningful conversations about suicide; emojis and anonymity facilitate connection
Record et al. (2020)	Evaluate predictors of knowledge and intention to use the reporting tool	Descriptive statistics, χ^2 , multiple linear regression	19.7% are aware of the tool. Reporting intention is predicted by attitude ($\beta = 0.31$; $p < .05$), subjective norms ($\beta = 0.27$; $p < .05$), and behavioral control ($\beta = 0.3$; $p < .05$). An R^2 of .5 was reached	The tool has not been widely disseminated; once informed, users show a high intention to use it; women and long-term users are more aware of it

Note. ¹Authors whose studies also contributed to thematic category 2; ²authors whose studies also contributed to thematic category 4; ³authors whose studies also contributed to thematic category 5.

Table 3 presents the main findings of the included studies, organized according to their objectives, analyses conducted, and results. To examine these findings in greater depth, thematic categories were established based on the analysis of the findings and classified as algorithmic amplification, social reinforcement and normalization of self-injury on social media, exposure to self-injury content and psychosocial effects, pre-existing vulnerabilities and associated social factors, and social media as spaces for support and intervention. It should be noted that some of the studies analyzed represented more than one thematic category due to their findings, objectives, or approach; however, it was possible to identify the category to which each study contributed the most information. Accordingly, these studies were marked with a superscript number to indicate the additional thematic category to which they contributed.

Overall, the studies consistently highlighted the ambivalent role of social media. The most frequently used analyses were thematic, correlational, and regression analyses, revealing exposure to sensitive content, the ineffectiveness of algorithmic moderation, and the importance of peer support.

Discussion

The objective of the present study was to analyze the literature published between 2019 and 2025 on the relationship between social media and self-injury in order to identify the main findings regarding risks, mechanisms of influence, and potential protective factors. The findings provided a broader understanding of the complex relationship among these elements. Overall, the reviewed studies agreed that digital plat-

forms constitute an ambivalent environment because, as mentioned in the introduction, they can function both as a risk factor by facilitating exposure to, imitation of, and normalization of self-injury, and as a protective factor by enabling social support, emotional validation, and online help-seeking (Grant-Allen et al., 2025; Lavis & Winter, 2020).

Overall, most of the reviewed articles defined self-injury as the deliberate infliction of physical harm to one's own body without suicidal intent, consistent with the classical literature (Chartrand et al., 2015; Nock, 2010). Nevertheless, some recent studies recognize that online manifestations, such as sharing images of scars, using related hashtags, or narrating experiences of emotional suffering, may represent symbolic forms of self-injury (Atauri-Mezquida et al., 2025; Martínez-Pastor et al., 2023). This suggests that self-injurious behavior has been transformed within the digital environment, extending into new forms of technology-mediated expression.

Regarding the research questions, the studies identified certain risk factors as particularly relevant to the use of digital media. These included exposure to explicit or sensitive content, which may foster imitation and normalization of physical harm (Atauri-Mezquida et al., 2025; Kobilke & Markiewitz, 2022); online social interactions that reinforce self-injurious behaviors through mechanisms of approval or recognition, such as retweets, likes, validating comments, and shares (Martínez-Pastor et al., 2023); as well as pre-existing personal vulnerabilities, such as depression, anxiety, or a prior history of suicide (Seong et al., 2021; Szlyk et al., 2023). The findings of the review also highlight that the presence of digital communities may promote self-injurious practices, as group belonging can reinforce aspects of an individual's identity associated with suffering (Lookingbill & Le, 2024). In addition, recommendation algorithms on platforms such as TikTok, X, or Instagram may amplify exposure to self-injury content, appearing to create risk bubbles that increase emotional vulnerability (Grant-Allen et al., 2025). These findings are consistent with the assumptions of the social contagion model (Arendt et al., 2019), although some studies suggest that this model is insufficient to explain the phenomenon in its entirety because it does not consider the subjective or relational functions that posts may have for users (Lavis & Winter, 2020). The contagion model primarily emphasizes processes of imitation or the dissemination of behaviors through exposure to self-injury-related content; however, it does not fully explain the psychological and social motivations underlying adolescents' and young people's participation in these types of digital interactions (Arendt et al., 2019). In this regard, other theoretical frameworks could complement this interpretation. For example, emotion regulation perspectives propose that self-injury may function as a strategy for managing intense emotions or psychological distress (Yang et al., 2025), whereas digital social identity approaches suggest that social media may facilitate the development of online communities in which certain experiences or behaviors are shared and normalized (Dyson et al., 2016). Likewise, ecological models of behavior make it possible to understand self-injury in digital environments

as the result of interactions among individual, relational, and contextual factors (Duarte et al., 2020).

Similarly, the findings of this study emphasize that Internet use may be beneficial rather than exclusively harmful, as several protective factors associated with social media use were identified despite the aforementioned risks. The reviewed literature showed that digital environments can serve as spaces for peer support, where young people find understanding, companionship, and strategies for coping with emotional distress (Lavis & Winter, 2020; Peart et al., 2023). This support may reduce isolation and encourage professional help-seeking, particularly when interactions are empathetic and mediated by counselors or supervised communities (Peart et al., 2023). Likewise, the community self-regulation observed on TikTok and other platforms reflected users' own attempts to establish internal norms of care and support (Lookingbill & Le, 2024). These behaviors indicate that, in certain contexts, social media can reproduce spaces of collective resilience.

In contrast, the sense of belonging and community that may develop on social media also entails certain risks when it occurs in spaces where self-injury-related content circulates. In this regard, limitations in algorithmic moderation allow graphic material or content that romanticizes self-injury to remain available (Atauri-Mezquida et al., 2025; Kobilke & Markiewitz, 2022), while interactions within these communities may provide such content with visibility and recognition, reinforcing its circulation and normalization. This may increase opportunities for exposure and contribute to certain representations of self-injury being perceived as socially acceptable. At the same time, online communities fulfill communicative and belonging-related functions for young people with prior self-injurious behaviors (Lavis & Winter, 2020; Szlyk et al., 2023). Identification with users who share similar experiences may provide support, but it may also encourage continued participation in spaces where self-injury becomes integrated into a shared identity (West et al., 2024). Thus, belonging to these communities may simultaneously represent a protective and a risk factor, depending on the characteristics of the interactions and the content circulating within them. However, exposure to such content should not necessarily be interpreted as a direct cause of NSSI, since some young people were already engaging in self-injurious behaviors before turning to these digital spaces in search of understanding, support, or shared experiences (Lavis & Winter, 2020).

Likewise, mechanisms that facilitate exposure to self-injury content were identified. Specific hashtags and communities, such as #shtwt and #selfharm, group posts together and facilitate their identification. This is compounded by the ineffectiveness of some moderation systems in blocking dangerous content, allowing certain posts to remain accessible (Atauri-Mezquida et al., 2025; Lookingbill & Le, 2024). Algorithms may increase the visibility of content similar to that with which users have previously interacted. In addition, through "algospeak," expressions are modified to circumvent filters and maintain the circulation of content (Grant-Allen et al., 2025). These mechanisms make it more difficult for users

to disengage from self-injury content and encourage continued participation in certain communities. Therefore, digital literacy should be strengthened and policies should be developed to identify these dynamics and promote safer interactions on these platforms (La Sala et al., 2025).

The foregoing suggests that platforms tend to function as risk factors when users have pre-existing vulnerabilities, such as depression, suicidal ideation, or a history of self-injury, and are repeatedly exposed to content that normalizes or increases the visibility of these behaviors (Seong et al., 2021; Szlyk et al., 2023). In these contexts, interaction dynamics inherent to social media, such as “likes,” comments, or shares, may function as forms of social reinforcement that legitimize self-injurious behavior (Martínez-Pastor et al., 2023), while recommendation algorithms may amplify exposure to such content or facilitate its circulation (Lookingbill & Le, 2024). In contrast, social media may function as a protective resource when it facilitates spaces for peer support, emotional validation, or help-seeking (Lavis & Winter, 2020; Peart et al., 2023). This effect is more likely when communities develop internal norms of care or strategies to reduce harm within digital spaces themselves (Lookingbill & Le, 2024). In this regard, the impact of social media depends on the interaction among users’ pre-existing vulnerability, the type of online participation, and the regulatory dynamics present within each digital environment.

Limitations and future directions

Although the PRISMA guidelines were rigorously followed, several limitations should be considered. First, the search period (2019-2025) may have excluded earlier relevant studies providing longitudinal evidence. Second, only articles published in English and Spanish were included, which limits cultural representativeness. Finally, the conceptual and methodological heterogeneity across studies makes quantitative comparison difficult; therefore, the findings should be interpreted with caution and primarily from a descriptive perspective, particularly because cross-sectional studies predominated, limiting the ability to infer causality. Likewise, the cultural context of the regions in which knowledge on self-injury in relation to social media has been generated may represent an important source of bias. Accordingly, and considering the issues discussed above, future research should focus on conducting explanatory studies that test models capable of generalizing the findings and progressively approximating the complex reality surrounding the phenomenon under study.

Practical implications

These findings may contribute to the development of preventive strategies and digital interventions aimed at adolescents and young people, taking into account the role of social media in NSSI. It is important to strengthen policies for moderating sensitive content and to promote digital literacy and mental health programs aimed at identifying risks and fostering safer online interactions. Likewise, future research could

develop explanatory models, early detection tools, and intervention protocols focused on promoting emotional well-being among young people.

Conclusion

The synthesized evidence showed that the relationship between NSSI and social media is complex and ambivalent. Social media emerged as an ambivalent environment that may either mitigate or promote self-injury. The findings indicated that recommendation algorithms, the lack of effective moderation, and the persistence of coded language (“algospeak”) increase exposure to sensitive content, underscoring the urgency of more transparent digital policies and media literacy initiatives with a mental health focus. Likewise, social media should be understood as social contexts in which difficult emotions are expressed and negotiated, rather than solely as harmful spaces. Overall, the findings suggest the need to design trauma-informed, preventive, and participatory digital interventions that promote safe and resilient online communities. They also highlight the need to clarify the mechanisms underlying protective and risk factors, as the effects of social media depend on specific conditions such as users’ pre-existing vulnerability, the type of digital participation, and the degree of moderation present within online communities.

Authors contribution

Conceptualization: K.G.D.-T.
Data curation: D.F.B.
Formal analysis: K.G.D.-T., D.F.B.
Investigation: K.G.D.-T., D.F.B.
Methodology: D.F.B.
Project administration: K.G.D.-T.
Resources: K.G.D.-T.
Software: K.G.D.-T., D.F.B.
Supervision: K.G.D.-T.
Validation: K.G.D.-T., D.F.B.
Visualization: D.F.B.
Writing – Original draft: K.G.D.-T., D.F.B.
Writing – Review & editing: K.G.D.-T., D.F.B.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of interests

The authors declare that there is no conflict of interest.

Data availability statement

Data pooling is not appropriate in this article, since the analyses were not based on new datasets but rather on previously published studies included in the systematic review.

References

**References marked with an asterisk indicate studies included in the review.

- *Alsoubai, A., Razi, A., Agha, Z., Ali, S., Stringhini, G., De Choudhury, M., & Wisniewski, P. (2024). Profiling the offline and online risk experiences of youth to develop targeted interventions for online safety. *Proceedings of the ACM on Human-Computer Interaction*, 8(114), 1-37. <https://doi.org/10.1145/3637391>
- Arendt, F., Scherr, S., & Romer, D. (2019). Effects of exposure to self-harm on social media: Evidence from a two-wave panel study among young adults. *New Media & Society*, 21(11-12), 2422-2442. <https://doi.org/10.1177/1461444819850106>
- *Atauri-Mezquida, D., Nogales-González, C., & Martínez-Pastor, E. (2025). Exploring self-harm on Twitter (X): Content moderation and its psychological effects on adolescents. *Online Journal of Communication and Media Technologies*, 15(1), 1-12. <https://doi.org/10.30935/ojcm/15867>
- Chartrand, H., Bhaskaran, J., Sareen, J., Katz, I., & Bolton, J. (2015). Correlates of nonsuicidal self-injury and suicide attempts among tertiary care, emergency department patients. *Canadian Journal of Psychiatry*, 60(6), 276-283. <https://doi.org/10.1177/070674371506000606>
- Dyson, M., Hartling, L., Shulhan, J., Chisholm, A., Milne, A., Sundar, P., Scott, S., & Newton, A. (2016). A systematic review of social media use to discuss and view deliberate self-harm acts. *PLoS ONE* 11(5), Article e0155813. <https://doi.org/10.1371/journal.pone.0155813>
- Duarte, K., & Fregoso, D. (2024). Relación entre autolesión no suicida y cibervictimización en adolescentes en adolescentes mexicanos. In C. Cruz, A. Ruiz-Celis, M. Pacheco y N. Ruvalcaba (eds.), *Investigación actual en psicología social* (pp. 425-438). Universidad de Guadalajara.
- Duarte, K., Vera, J., & Fregoso, D. (2020). Factores contextuales relacionados a las conductas de autolesión no suicidas: Una revisión sistemática. *Psicología y Salud*, 31(1), 51-59. <https://doi.org/10.25009/pys.v31i1.2675>
- Gillies, D., Christou, M., Dixon, A., Featherston, O., Rapti, I., Garcia-Anguita, A., Villasis-Keever, M., Reebye, P., Christou, E., Al Kabir, N., & Christou, P. A. (2018). Prevalence and characteristics of self-harm in adolescents: Meta-analyses of community-based studies 1990-2015. *Journal of the American Academy of Child and Adolescent Psychiatry*, 57(10), 733-741. <https://doi.org/10.1016/j.jaac.2018.06.018>
- Grant-Allen, G., Wang, L., Amini, J., Dhaliwal, S., Sinyor, M., & Mitchell, R. (2025). Self-harm and suicide-related content on TikTok: Thematic analysis. *Journal of Medical Internet Research*, 27, 1-11. <http://dx.doi.org/10.2196/77828>
- Hong, Q., Fàbregues, S., Bertlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M., Vedel, I., & Pluye P. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for Information*, 34(4), <https://doi.org/10.3233/EFI-180221>
- *Kobilke, L., & Markiewicz, A. (2021). The Momo Challenge: Measuring the extent to which YouTube portrays harmful and helpful depictions of a suicide game. *SN Social Sciences*, 1, Article 86. <https://doi.org/10.1007/s43545-021-00065-1>
- La Sala, L., Sabo, A., Michail, M., Thorn, P., Lamblin, M., Browne, V., & Robinson, J. (2025). Online safety when considering self-harm and suicide related content: Qualitative focus group study with young people, policy makers, and social media industry professionals. *Journal of Medical Internet Research*, 27, Article e66321. <https://doi.org/10.2196/66321>
- *Latina, D., Goreis, A., Sajko, P., & Kothgassner, O. (2023). Does being ignored on WhatsApp hurt? A pilot study on the effect of a newly developed ostracism task for adolescents. *Journal of Clinical Medicine*, 12(5), Article 2056. <https://doi.org/10.3390/jcm12052056>
- *Lavis, A., & Winter, R. (2020). Online harms or benefits? An ethnographic analysis of the positives and negatives of peer-support around self-harm on social media. *Journal of Child Psychology and Psychiatry*, 61(8), 842-854. <https://doi.org/10.1111/jcpp.13245>
- Lewis, S., Heath, N., Sornberger, M., & Arbuthnott, A. (2014). Googling self-injury: The state of health information obtained through internet search engines. *JAMA Pediatrics*, 168(5), 443-449. <https://doi.org/10.1001/jamapediatrics.2014.187>
- *Lim, S., Ng, Q., Xin, X., Lim, Y., Boon, E., & Liew, T. (2022). Public discourse surrounding suicide during the COVID-19 pandemic: An unsupervised machine learning analysis of Twitter posts over a one-year period. *International Journal of Environmental Research and Public Health*, 19(21), Article 13834. <https://doi.org/10.3390/ijerph192113834>
- *Lookingbill, V., & Le, K. (2024). "There's always a way to get around the guidelines": Nonsuicidal self-injury and content moderation on TikTok. *Social Media + Society*, 10(2), 1-14. <https://doi.org/10.1177/20563051241254371>
- *Martínez-Pastor, E., Atauri-Mezquida, D., Nicolás-Ojeda, M., & Blanco-Ruiz, M. (2023). Visualización e interpretación de las interacciones en los mensajes de autolesiones no suicidas en Twitter. *Redes. Revista Hispana para el Análisis de Redes Sociales*, 34(2), 238-253. <https://doi.org/10.5565/rev/redes.996>
- Muehlenkamp, J. & Gutiérrez, P. (2007). Risk for suicide attempts among adolescents who engage in non-suicidal self-injury. *Archives of Suicide Research*, 11(1), 69-82. <https://doi.org/10.1080/13811110600992902>
- Nock, M. (2010). Self-injury. *Annual Review of Clinical Psychology*, 6, 339-363. <https://doi.org/10.1146/annurev.clinpsy.121208.131258>
- Packer, A. (2024). Multilingualism in literature communicated by journals from the SciELO Brazil Collection. *European Review*, 32(S1), 124-144. <https://doi.org/10.1017/S1062798724000103>
- Page, M., McKenzie, J., Bossuyt, P., Boutron, I., Hoffman, T., Mulrow, C., Shamseer, L., Tetzlaff, J., Akl, E., Brennan, S., Chou, R., Glanville, J., Grimshaw, J., Hróbjartsson, A., Lalu, M., Li, T., Loder, E., Mayo-Wilson, E., McDonald, S., McGuinness, L. ... David. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *PLOS Medicine*, 18(3), Article e1003583. <https://doi.org/10.1371/journal.pmed.1003583>
- *Peart, N., Hetrick, S., Gibson, K., & Stasiak, K. (2023). Exploring counselor practices and risk assessment in a proactive digital intervention through Instagram in young people: Qualitative study. *JMIR Formative Research*, 7, Article e46579. <https://doi.org/10.2196/46579>
- *Record, R., Straub, K., & Stump, N. (2020). #Selfharm on #Instagram: Examining user awareness and use of Instagram's self-harm reporting tool. *Health Communication*, 35(7), 894-901. <https://doi.org/10.1080/10410236.2019.1598738>
- Robinson, J., Cox, G., Bailey, E., Hetrick, S., Rodrigues, M., & Fisher, S. (2016). Social media and suicide prevention: A systematic review. *Early Intervention in Psychiatry*, 10(2), 103-121. <https://doi.org/10.1111/eip.12229>

- Rojas-Jara, C., Henriquez, F., Sanhueza, F., Núñez, P., Inostroza, E., Solís, A., & Contreras, D. (2018). Adicción a internet y uso de redes sociales en adolescentes: Una revisión. *Revisiones*, 43(4), 39-54.
- Ryan, R., Deci, E., Vanteenkiste, M., & Soenens, B. (2021). Building a science of motivated persons: Self-determination theory's empirical approach to human experience and the regulation of behavior. *Motivation Science*, 7(2), 97-100. <https://doi.org/10.1037/mot0000194>
- Santos, D. (2011). *Autolesión: Qué es y cómo ayudar*. Edición de Autor.
- *Scherr, S., Arendt, F., Frissen, T., & Oramas, J. (2019). Detecting intentional self-harm on Instagram: development, testing, and validation of an automatic image-recognition algorithm to discover cutting-related posts. *Social Science Computer Review*, 38(6), 673-685. <https://doi.org/10.1177/0894439319836389>
- Schmidt, C., Briones-Buixassa, L., Nicolaou, S., Soler, J., Pascual, J., & Vega, D. (2023). Non-suicidal self-injury in young adults with and without borderline personality disorder: The role of emotion dysregulation and negative urgency. *Anales de Psicología*, 39(3), 345-353. <https://doi.org/10.6018/analesps.492631>
- *Seong, E., Noh, G., Lee, K., Lee, J., Kim, S., Seo, D., Yoo, J., Hwang, H., Choi, C., Han, D., Hong, S., & Kim, J. (2021). Relationship of social and behavioral characteristics to suicidality in community adolescents with self-harm: Considering contagion and connection on social media. *Frontiers in Psychology*, 12, Article 691438. <https://doi.org/10.3389/fpsyg.2021.691438>
- *Szlyk, H., Li, X., Kasson, E., Peoples, J., Montayne, M., Kaiser, N., & Cavazos-Rehg, P. (2023). How do teens with a history of suicidal behavior and self-harm interact with social media? *Journal of Adolescence*, 95(4), 797-810. <https://doi.org/10.1002/jad.12154>
- We Are Social & Meltwater. (2025, 30 de abril). *Digital 2025: Global overview report*. We Are Social y Meltwater.
- West, M., Rice, S., & Vella-Brodrick, D. (2024). Adolescent social media use through a self-determination theory lens: A systematic scoping review. *International Journal of Environmental Research and Public Health*, 21(7), Article 862. <https://doi.org/10.3390/ijerph21070862>
- Woodward, M., McGettrick, C., Dick, O., Ali, M., & Teeters, J. (2025). Time spent on social media and associations with mental health in young adults: Examining TikTok, Twitter, Instagram, Facebook, Youtube, Snapchat, and Reddit. *Journal of Technology in Behavioral Science*, 10, 661-671. <https://doi.org/10.1007/s41347-024-00474-y>
- Yang, Y., Ji, T., Liu, Y., Tian, M., Yang, Y., Zhang, Y., & Lin, L. (2025). Non-suicidal self-injury and emotion regulation strategies among secondary school and university students: A network analysis perspective. *Behavioral Sciences*, 15(11) Article 1517. <https://doi.org/10.3390/bs15111517>