Social media, self-harm, and suicide

Sebastian Scherr

Abstract

Self-harm- and suicide-related (SHS) content on social media is not clearly defined, and therefore, audience effects remain scattered. This article makes three contributions: First, it offers a definition and taxonomy for SHS content on social media with potentially negative audience effects. SHS content on social media is either explicit, implicit or ambivalent in nature, which makes it hard to regulate, and challenging for media effects research. Second, different forms of social media use are discussed as antecedents to self-harm and suicide. And third, functional social media affordances that shape the exposure to problematic SHS content on social media are reviewed. The regulation of the shape-diverse, problematic SHS content on social media remains a pressing future challenge.

Addresses

Department of Communication, Texas A&M University, 202E Bolton Hall, College Station, TX, 77843, USA

Corresponding author: Scherr, Sebastian (scherr@tamu.edu)

Self-harm- and suicide-related (SHS) content on social media includes more or less explicit depictions of self-injurious thoughts and behaviors—including suicidal ideation, suicide attempts, and non-suicidal self-injury. Importantly, self-injurious thoughts and behaviors (SITBs) rose in the society as social media use gained traction, sparking public and academic discussions about possible connections [1–3]. On social media, users not only passively consume content, but also create and share it with others, which results in SHS content on social media that is manifold, shape-diverse, and generally recognized as particularly impactful on audiences [4–6].

The most recent, comprehensive reviews [1,2,6] have focused on specific forms of social media use and experiences, including cybervictimization, heavy or problematic social media use, sexting, and also exposure to self-harm content on social media, but did not explicitly assess how social media use can be an antecedent of SITBs. Moreover, what users were actually seeing on their social media was neither systematically assessed, nor were the functional affordances of social media mentioned that shape the content. These reviews report mostly positive, small- to medium-sized associations between specific social media use forms or patterns and SITBs, especially among adolescents, and less so among adults. The number of available studies is limited, data from vulnerable groups is particularly scarce, longitudinal studies are virtually absent [7], and the lack of systematic social media use measures is consistently criticized.

However, there is growing interest not only in exploring the effects of SHS content on social media, but also in regulating such content. As there is no clear definition or taxonomy of SHS content on social media yet, it has been difficult to connect findings from different studies and to elaborate on regulation. In this article, I fill this research gap and suggest such a definition and taxonomy. I define SHS social media content as explicit (e.g., suicide memes), implicit (e.g., references to thoughts or symptoms), or ambivalent (e.g., metaphorical use of suicide, as in *suicide mission*) depictions of thoughts or acts of self-harm or suicide in an either fictional or non-fictional context. Importantly, SHS content on social media can be pre-made or livestreamed using audio, video, text, or images. Particularly on social media, this also includes responses to such content (e.g., comments, likes). Table 1 shows the definition and taxonomy and offers more specific examples of SHS content that users can be exposed to on social media.

Recently published reviews consistently report that existing evidence is best described as increasing but still in the preliminary stages. There is only a limited number of studies that yielded barely comparable findings due to a lack of systematic social media measures and an exclusive focus on specific age groups. However, what is especially missing, is an understanding of the underlying conceptual processes that explain *how* social media impacts different users in regard to their tendencies to self-harm and suicide. Reviews generally acknowledge the twofold role of social media for SITBs more [6] or less explicitly [2]: Social media can be a

Table 1

Definition and taxonomy for suicidal and self-harm (SHS) content on social media

Definition	Type of content	Specific examples of SHS content on social media
SHS content on social media is defined as fictional or non-fictional, pre-made or livestreamed content (audio, video, text, image) depicting thoughts about or acts of self-harm or suicide as well as plans, preparations, arrangements, or detailed instructions for it, including the responses to such content by others.	**Explicit** Social media content that normalizes, glorifies, promotes, encourages, or triggers others to self-harm or attempt suicide, but also discourages and judges it and raises public awareness of the issue, including feelings of hopelessness and not wanting to continue living, with or without references to memories or the user's own lived experience of self-harm or suicide, including mocking/trolling others about such content. **Implicit** Social media content that indirectly refers to, hints at, or figuratively alludes to self-harm or suicide, as well as emotions, cognitions, or behaviors often connected to and commonly understood as co-occurring with self-injurious behaviors. **Ambivalent** Social media content that refers to self-harm and suicide in vague or ambiguous terms using stylistic or literary devices, visual references, or hashtags and codenames that circumvent upload/content filters and/or disguise a reference publicly.	• Goodbye letters, fast-cut "mash-up videos," videos containing self-harm- and suicide-related subliminal messages [29] • Self-harm- or suicide-related artwork or memes [33] • Self-harm- or suicide-related "online challenges" [34] • Triggering imagery, including cuts, wounds, and other forms of self-infliction [32] • Inspirational quotes, pictures, or memes [35] • References to symptoms of mental disorders [36] • References to feeling trapped or desperate [37] • Extremely negative or (self-)destructive content about negative emotions, mental illness, traumatic experiences, low self-esteem, self-hate, or self-loathing • Metaphorical use of suicide (*suicide mission*) or self-harm (*shot in the foot*) • Ironic, cynical, or polemic expressions referencing self-harm or suicide • Jokes, memes, or gifs about self-harm or suicide • Codenames for mental disorders (e.g., using the names *Cat* [if female] or *Sam* [if male] as a code for self-harm or *Sue* or *Dallas* if suicidal)

potential risk factor for SITBs, but can also protect against them.

Based on my definition of SHS content on social media (see Table 1), this article 1) discusses forms of social media use as antecedents of self-harm and suicide, and 2) identifies the specific affordances of social media that contribute or prevent them.

Social media use as an antecedent of self-harm and suicide: Findings and explanations

Two topoi emerge from the literature: First, not all social media screen time is equal for SITBs, and second, the rapid adoption of technology (i.e., Internet, social media, smartphones) seems to be a probable cause of SITBs as part of a cultural shift in how we interact with each other [3]. I discuss how social media use can be conceptualized as an antecedent of SITBs. I will focus on social media use as well as specific social media use phenomena, including cyberbullying, sexting, and suicide-related social media use, and will explain the evidence regarding how social media operates in conjunction with SITBs.

Frequency and intensity of social media use

Studies focusing on social media screen time, averaged across various media and formed into a global measure, found small-to-no associations with well-being measures that included indicators of self-harm and suicide [8,9]. Studies that used the same datasets, but differentiated screen time, and specifically looked at, for example, social media and Internet use, found slightly stronger associations with SITBs, especially among girls [10]. Studies looking at extremely intensive, addictive, and problematic forms of social media use usually found small- to medium-sized associations with SITBs [2].

Cyberbullying

Social media is interpersonal in nature, and so are SITBs [11]. Usually, a social disconnect or a thwarted belongingness to others is a major risk factor for suicide, perceived burdensomeness for others is another. However, all of these factors require that a person has acquired the capability of dying by suicide. Importantly, cybervictimization and cyberbullying perpetration can both affect thwarted belongingness and perceived burdensomeness. While cybervictimization has been robustly associated with SITBs, the studies were very

heterogeneous, and the findings were moderated by factors such as sample age, sex, or the observed timeframe for media effects [2,12]. Studies on the effects of cyberbullying perpetration on SITBs are much more scarce and much less conclusive [2]. The findings are almost exclusively geographically constrained to the Western hemisphere, and longitudinal, primary research, qualitative investigations as well as intersectional aspects involving racial and/or ethnic minorities, different sexual orientations and identities, and/or disabilities are missing [12].

Sexting

It remains unclear as to how sexting is related to SITBs. Some argue that sexting generates a normative pressure to respond, which can promote the perception that it is necessary to engage in sexting in order to sustain romantic relationships [1]. Feelings of being entrapped can increase distress for vulnerable individuals and thus contribute to SITBs [13]: Individual vulnerabilities (e.g., mental health) and environmental stressors (e.g., derailed sexting episodes) can contribute to feelings of entrapment—defined perceptions of defeat and humiliation—from which the only escape seems to be through suicide. Further progression from suicidal ideation to suicidal intent depends on motivational factors, such as belongingness or social connectedness [14]. Non-reciprocated sexting can thwart belongingness, and sexts forwarded to others they were not intended for may not only destroy a reputation and contribute to feelings of shame, entrapment, and perceptions of being a burden on one's family, but may also indirectly contribute to cyberbullying effects on SITBs.

However, sexting might also boost relational trust and confidence and boost self-esteem, which has been associated with lowered SITBs [15]. Evidence suggests a range of small-to-large associations between sexting and SITBs based only on a few studies, which warrants caution regarding interpretation and bears the risk that the true effects are actually less stable than they currently seem to be [2].

Browsing SHS content on social media

SHS social media use has been differentiated into 1) actively posting or talking about SITBs on social media (i.e., social media engagement) and 2) being exposed to SHS content on social media [1,2]. Regarding the exposure to SHS content, it is crucial to keep in mind that a definition such as the one presented in Table 1 had been missing, resulting in a lack of comparability across findings. Most observed effects of SHS social media engagement or SHS content exposure were of a small-to-medium size [1,2]. Conditional effects and individual vulnerabilities for social media effects were largely disregarded and were only more recently streamlined into existing conceptualizations.

Scherr [16] suggested a *media-related diathesis-stress model* that integrates influences of depression as a vulnerability factor for suicide, attitudes, social norms, self-efficacy, and emotions, which may all ultimately contribute to self-harm and suicide. Importantly, an ability to perform self-harm and suicide behaviors must be present and external restrictions to performing the behaviors must be absent. Furthermore, one could argue that SHS content on social media resonates more strongly with vulnerable audiences [17]. It might well be that SHS content resonates particularly well with vulnerable audiences' cognitions, attitudes, and experiences, which arguably increases both potentially harmful (e.g., making suicide plans) and protective effects (e.g., help-seeking).

Posting SHS content on social media

Studies investigating the relationship between popular social media interactions with others and SITBs [2,18] repeatedly found associations between problematic Internet use (oftentimes defined as excessive use with similarities to addiction disorders) and SITBs; interestingly, SITBs were most prevalent among those, who used the Internet for chatting and blogging and these are more popular activities among girls. But things are more complicated: Nesi et al. [19] found that girls with a suicide attempt history not only saw and posted more SHS content on social media than boys or than girls without suicide attempt history, they also reported to have received the most social support, encouragement, and distraction from their problems on social media. One aspect that might come into play for social media interactions is *self-effects* [20] that is, the effects of message composition on an individual's own cognitions, emotions, attitudes, and behaviors, in contrast to *reception effects* (i.e., the effects of *others'* messages on oneself). Similar to the mental health benefits of expressive writing [21], social media self-effects can increase perceived social support, well-being, and self-esteem, and reduce depressive symptoms [20,22]. Expressive writing (e.g., journaling, blogging) can push individuals to re-evaluate aspects of their life and reflect upon their emotions and emotional responses to their environment, thereby yielding changes.

Posting about SHS content has been associated with SITBs, and vice versa. Adolescents with higher SITB rates commented, reposted, or talked more frequently about suicide and suicide-related news on social media [1]. However, one longitudinal study [23] also showed that SITBs can predict the use of online health forums and support groups, but their use does not automatically translate into fewer SITBs a month later. This finding and scarce additional evidence [15] leave room for speculation about what other factors might have been responsible for protective social media effects, with social media self-effects being one plausible explanation.

Thus, as a more general concern, future studies should look more deeply into what people are specifically engaging with on social media. Research on extreme forms of social media interactions with others, including stress disclosures (e.g., "I want to die"), for example, found them to be associated with increased SITBs [1]. In line with social media self-effects, such extreme self-disclosures might bring even more troublesome facets of existing problems to the surface, resulting in harmful backfire effects, especially when shared on social media.

Functional social media affordances and their link with self-harm and suicide

Social media enables the broadcasting of personal, self-related information to large, potentially global audiences, which exacerbates the previously discussed SHS antecedents. Specifically, social media comes with affordances that can be categorized as functional, social, identity, cognitive, and emotional affordances [24]. I will focus on the functional affordances here. Basic social media functions that define how content is exchanged include its virtually boundless replicability, searchability, scalability, permanence, editability, and composition time [24–27].

Boundless replicability, searchability, and scalability

SHS content can be re-used by other users, to the extent that content is shared to such an extent that it goes "viral." Users can find SHS content because of its searchability and interconnectedness using hashtags to categorize content. Hashtags can be explicit (e.g., #selfharm, # suicide), implicit (e.g., #sad, #heartbroken) or ambivalent (e.g., #cat, #mysecretfamily) and are typically self-selected by users to label and connect uploaded content. Thereby, for example cyberbullying is no longer limited to one school; self-disclosures can go viral, and so can derailed sexting.

Permanence and cultural boundlessness

SHS posts potentially reach global audiences for an indefinite timespan and irrespective of the cultural background and social norms related to SITBs. Users can be willingly or accidently exposed to more or less elaborately crafted and tagged SHS social media posts with potentially much stronger self- and audience effects [20]. SHS content is not only able to reach much larger audiences, but is also available and accessible 24/7, and once online, it is hard to extinguish, ever (cf. Cyberbullying, sexting). Postings about SITBs on social media might also create unprecedented forms of public commitment that might further increase the pressure on individuals to go through with an announced plan [28].

Editability and boundless composition time

Finally, SHS content can be carefully prepared, pre-recorded and/or posted with a delay [27]. The editability of SHS content may contribute to a selective and exaggerated public image of SITBs. This image can deviate from the felt inner states that suicidal individuals experience for themselves, which might further increase confusion and uncertainty. Crafting SHS content to post on social media, on the other hand, might also help individuals to reflect on their SITBs and to become clearer. The asynchronicity [27] of SHS content can prevent others from live interventions if it remains unclear when SHS posts were actually made. Users might choose to delay their posts or have edited content in a way that challenges automatic upload filters to detect hidden messages [29]. In the worst case, this could delay or prevent posts that were meant as a cry for help from being online.

Taken together, functional social media affordances can make it harder to interpret findings from content analyses [25], as we do not necessarily know how the social media posts were created and when. One way to better understand the purpose of extreme SHS social media posts [30] would be to follow up on those users who posted extreme content, for example, by directly contacting them on the platform [31].

Outlook

Self-harm- and suicide-related content on social media is both prevalent and problematic. Scherr et al. developed and applied an image recognition algorithm to a 48 h sample of $N = 13{,}132$ pictures of self-harm posted on Instagram in June 2018 and found between 30.0% and 41.7% to be classified as containing explicit depictions of self-harm [32]. At around the same, regulators have started to acknowledge this as problematic, and in 2019, Instagram implemented upload and content filters to its platform. It remains unknown to what extent the implementation improved the situation. From a regulatory perspective, clear-cut definitions and taxonomies (see Table 1) about what content should be controlled and what should not will become most relevant moving forward.

From a media effects perspective, some social media effects are harmful, some are helpful. Psychological mechanisms involving various social media affordances deserve further attention to support individuals in seeking help when needed. Therefore, regulatory blanket solutions seem inappropriate. Due to the sheer amount of potentially problematic SHS content on social media, upload and content filters are needed that are capable of adaptation to the continuous flood of new forms of SHS content [28]. Without an independent monitoring of the available content by entities that are not the social networking companies themselves, it is impossible to know what users experience on these platforms. An individual, algorithm-driven content curation further complicates such observations. True

media effects are impossible to assess without finding this important piece of the puzzle.

Conflict of interest statement

Nothing declared.

References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest

1. Macrynikola N, Auad E, Menjivar J, Miranda R: **Does social media use confer suicide risk? A systematic review of the evidence**. *Comput Hum Behav Rep* 2021, **3**:100094.

2. Nesi J, Burke TA, Bettis AH, Kudinova AY, Thompson EC, MacPherson HA, Fox KA, Lawrence HR, Thomas SA, Wolff JC, *et al.*: **Social media use and self-injurious thoughts and behaviors: a systematic review and meta-analysis**. *Clin Psychol Rev* 2021, **87**:102038.

3. Twenge JM: **Why increases in adolescent depression may be linked to the technological environment**. *Curr Opin in Psychol* 2020, **32**:89–94.

4. ** Picardo J, McKenzie SK, Collings S, Jenkin G: **Suicide and self-harm content on Instagram: a systematic scoping review**. *PLoS One* 2020, **15**. e0238603.

This is an excellent scoping review about suicide and self-harm *content on social media* that acknowledges the conceptual and terminological shortcomings of the field, presents the main findings broadly, and also reflects upon the rare observations in this research area, which makes it an excellent starting point for those interested in learning more about the topic.

5. Sedgwick R, Epstein S, Dutta R, Ougrin D: **Social media, internet use and suicide attempts in adolescents**. *Curr Opin Psychiatr* 2019, **32**:534–541.

6. Biernesser C, Sewall CJR, Brent D, Bear T, Mair C, Trauth J: **Social media use and deliberate self-harm among youth: a systematized narrative review**. *Child Youth Serv Rev* 2020, **116**, e105054.

7. * Arendt F, Scherr S, Romer D: **Effects of exposure to self-harm on social media: evidence from a two-wave panel study among young adults**. *New Media Soc* 2019, **21**:2422–2442.

This highly downloaded study presents rare, longitudinal (1 month) evidence for the effects of self-reported exposure to self-harm on social media among young adults in the United States. It sheds light on the temporal order and unique detrimental impact of exposure to self-harm on social media on suicidal ideation, self-harm, and emotional disturbance among young users, even if other sources with similar content are controlled for.

8. Orben A, Przybylski AK: **The association between adolescent well-being and digital technology use**. *Nat Hum Behav* 2019, **3**:173–182.

9. Orben A, Przybylski AK: **Screens, teens, and psychological well-being: evidence from three time-use-diary studies**. *Psychol Sci* 2019, **30**:682–696.

10. Twenge JM, Farley E: **Not all screen time is created equal: associations with mental health vary by activity and gender**. *Soc Psychiatr Psychiatr Epidemiol* 2021, **56**:207–217.

11. Joiner TE: *Why people die by suicide*. Harvard University Press; 2005.

12. Kwan I, Dickson K, Richardson M, MacDowall W, Burchett H, Stansfield C, Brunton G, Sutcliffe K, Thomas J: **Cyberbullying and children and young people's mental health: a systematic map of systematic reviews**. *Cyberpsychol, Behav Soc Netw* 2020, **23**:72–82.

13. O'Connor RC: **The integrated motivational-volitional model of suicidal behavior**. *Crisis* 2011, **32**:295–298.

14. O'Connor RC, Kirtley OJ: **The integrated motivational–volitional model of suicidal behaviour**. *Phil Trans Biol Sci* 2018, **373**:20170268.

15. Choi D-H, Noh G-Y: **Associations between social media use and suicidal ideation in South Korea: mediating roles of social capital and self-esteem**. *Health Commun* 2020, **35**:1754–1761.

16. Scherr S: *Depression—medien—Suizid: Zur empirischen Relevanz von Depressionen und Medien für die Suizidalität [Depression—media—Suicide: on the Empirical Relevance of Depression and Media for Suicidality]*. Springer VS; 2016.

17. Valkenburg PM, Peter J: **The differential susceptibility to media effects model**. *J Commun* 2013, **63**:221–243.

18. Kim KM, Kim H, Choi JW, Kim SY, Kim JW: **What types of Internet services make adolescents addicted? Correlates of problematic Internet use**. *Neuropsychiatric Dis Treat* 2020, **16**:1031–1041.

19. Nesi J, Wolff JC, Hunt J: **Patterns of social media use among adolescents who are psychiatrically hospitalized**. *J Am Acad Child Adolesc Psychiatry* 2019, **58**:635–639.

20. Valkenburg PM: **Understanding self-effects in social media**. *Hum Commun Res* 2017, **43**:477–490.

21. Pennebaker JW: **Writing about emotional experiences as a therapeutic process**. *Psychol Sci* 1997, **8**:162–166.

22. Lee SW, Kim I, Yoo J, Park S, Jeong B, Cha M: **Insights from an expressive writing intervention on Facebook to help alleviate depressive symptoms**. *Comput Hum Behav* 2016, **62**:613–619.

23. Scherr S, Reinemann C: **First do no harm: cross-sectional and longitudinal evidence for the impact of individual suicidality on the use of online health forums and support groups**. *Comput Hum Behav* 2016, **61**:80–88.

24. Moreno MA, Uhls YT: **Applying an affordances approach and a developmental lens to approach adolescent social media use**. *Digit. Health* 2019, **5**. 2055207619826678.

25. Treem JW, Leonardi PM: **Social media use in organizations: exploring the affordances of visibility, editability, persistence, and association**. *Ann Int Commun Assoc* 2013, **36**:143–189.

26. boyd d: **Social network sites as networked publics: affordances, dynamics, and implications**. In *Networked Self: Identity, Community, and Culture on Social Network Sites*. Edited by Papacharissi Z, Routledge; 2011:39–58.

27. Valkenburg PM, Peter J: **Online communication among adolescents: an integrated model of its attraction, opportunities, and risks**. *J Adolesc Health* 2011, **48**:121–127.

28. Luxton DD, June JD, Fairall JM: **Social media and suicide: a public health perspective**. *Am J Publ Health* 2012, **102**:S195–S200.

29. Arendt F, Markiewitz A, Scherr S: **Investigating suicide-related subliminal messages on Instagram**. *Crisis* 2020, **42**:263–269.

30. ** Brown RC, Fischer T, Goldwich AD, Keller F, Young R, Plener PL: **#cutting: non-suicidal self-injury (NSSI) on Instagram**. *Psychol Med* 2018, **48**:337–346.

Paul Plener is a pioneer in studying self-harm on social media. This study exemplifies the methodological breadth and creativity of his research group and the research field.

31. Brown RC, Fischer T, Goldwich DA, Plener PL: **"I just finally wanted to belong somewhere": qualitative analysis of experiences with posting pictures of self-injury on Instagram**. *Front Psychiatr* 2020, **11**.

32. Scherr S, Arendt F, Frissen T, Oramas MJ: **Detecting intentional self-harm on Instagram: development, testing, and validation of an automatic image-recognition algorithm to discover cutting-related posts**. *Soc Sci Comput Rev* 2020, **38**:673–685.

33. Shanahan N, Brennan C, House A: **Self-harm and social media: thematic analysis of images posted on three social media sites**. *BMJ Open* 2019, **9**, e027006.

34. Kobilke L, Markiewitz A: **The Momo Challenge: measuring the extent to which YouTube portrays harmful and helpful depictions of a suicide game**. *SN SS* 2021, **1**:86.

35. Carlyle KE, Guidry JPD, Williams K, Tabaac A, Perrin PB: **Suicide conversations on Instagram™: contagion or caring?** *J Commun Healthc* 2018, **11**:12–18.

36. Moreno MA, Ton A, Selkie E, Evans Y: **Secret society 123: understanding the language of self-harm on Instagram**. *J Adolesc Health* 2016, **58**:78–84.

37. Seko Y, Lewis SP: **The self—harmed, visualized, and reblogged: remaking of self-injury narratives on Tumblr**. *New Media Soc* 2016, **20**:180–198.