



OPEN ACCESS

EDITED BY

Tour Liu,
Tianjin Normal University, China

REVIEWED BY

Magda Julissa Rojas-Bahamón,
Jorge Eliécer Gaitán, Colombia
Grace Jochan,
St. Joseph's University, India

*CORRESPONDENCE

Sobi Thomas
✉ sobykannalil@gmail.com

RECEIVED 03 July 2025

ACCEPTED 19 September 2025

PUBLISHED 07 October 2025

CITATION

Thomas S and Manalil P (2025) Digital silence:
the psychological impact of being shadow
banned on mental health and self-perception.
Front. Psychol. 16:1659272.
doi: 10.3389/fpsyg.2025.1659272

COPYRIGHT

© 2025 Thomas and Manalil. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Digital silence: the psychological impact of being shadow banned on mental health and self-perception

Sobi Thomas^{1,2*} and Paul Manalil¹

¹School of Social Science, Arts and Humanities, Lincoln University College, Petaling Jaya, Malaysia,

²Department of Communication and Media Studies, Marian College Kuttikkanam, Kuttikkanam, India

KEYWORDS

shadow banning, algorithmic invisibility, self-perception, digital silence, mental health, media psychology, online validation, emotional regulation

1 Introduction: when the feed goes quiet

Imagine that you publish a well-thought-out post, photo, or video—only to watch it silently sink into obscurity. No likes, no comments, no shares. At first, you can brush it off as a fluke or a bad content day. But for many of us, specifically for those who are on active social media platforms like Instagram, Tiktok, or X (previously Twitter), Shadow banning could be the primarily cause resulting in their silence. Shadow banning could be defined as an algorithmic hiding in which the content is quietly de-amplified without no indication (Liu et al., 2023). To differentiate overt censorship from shadow banning, it is an act with a conscious face whereas shadow banning is invisible and creates a sense of social erasure that could potential result in emotional disorientation and psychological distress. Recently, research studies have begun emphasizing the importance to recognize the shadow banning not only as a technical limitation but also on a broader spectrum on digital exclusion and algorithmic marginalization (Delmonaco et al., 2024).

In this paper, we examine shadow banning more as an intensely subjective psycho-existential phenomenon rather than as a technical bug or policy enforcement strategy. Findings of this study show that Shadow banning emotionally affect the self-concept leading to disruptions in digital social feedbacks. The individuals are therefore compelled to rely for validation identification, reinforcement, and social inclusion. This study did a detailed analysis of the literature in media psychology and theories of emotional and digital behavior, and concludes that non-transparency of the social media platforms causes distress of individuals, and it needs to be addressed urgently.

2 Understanding shadow banning and its affective mechanism

Shadow banning also known as Stealth banning, silently prevents or restricts a user's reach in the social media platforms. It is a kind of algorithmic suppression without suspending the account. Unaware of the invisibility of the post in the community the user till continues posting, but the message never appears in search results, hashtags, or regular feeds, leading to decreased engagement. These users are, in fact, speaking to a void. This digital silence can be described as a vocal message within the social media economy.

Feedback is a sustainable rejuvenating factor of the online platforms. The activating responses through “likes, comments, reposts and follows” are emotional assets which indicates self-affirmation. These validations cannot be ignored as they activate neural centres which releases dopamine. When these signals vanish into thin air with no indication of why the users feel lost, rejected and struggle with cognitive dissonance (Politte-Corn et al., 2024). Am I ignored? Is my content awful? Have I done something inappropriate? The withdrawal from engagement is a psychological riddle that upsets the self-worth. From a psychological standpoint, this dynamic activates the mesolimbic dopamine system, reinforcing the role of social affirmation in self-perception (Cross et al., 2025). Cognitive dissonance arises when one's self-image as a socially engaged digital citizen clashes with unexplained algorithmic suppression. A qualitative analysis using Impression Management Theory and Cognitive Dissonance Theory found that teens experience dissonance when their social media presence conflicts with their real-world identity, often leading to discomfort and eventual withdrawal from online activity (Marta and Miletresia, 2022).

The line graph in the Figure 1 illustrates a noticeable decline in user engagement (likes, comments, shares) following a suspected shadow ban. The data is based on user-reported case studies, showing normal interaction patterns in the days prior (Days 1–15), followed by a significant drop post-event (Days 16–30). This pattern exemplifies the experience of “digital silence,” where content visibility is algorithmically suppressed without user notification, leading to emotional confusion and self-doubt. While this visual is based on informal reports and lacks formal statistical validation, it reflects a recurring pattern documented across multiple user narratives.

3 Emotional dysregulation and self-doubt in a platformed identity

Online, the identity is not just described—it is staged and legitimated in the public sphere. The self is algorithmically discernible, constituted with interaction metrics and validation from followers. When a user is shadow banned, they are systematically excluded from the social world. The shock invisibility disrupts emotional regulatory protocols and can induce depression symptoms, anxiety, and compulsive checking of content behaviors (Wikman et al., 2022).

The concerns of social exclusion were studied by media psychologists in recent years and their findings focus on the “indefiniteness” of shadow banning. The users were not told about the banning and the indefinite nature of such banning. The individuals quite often doubt their perception of reality and the emotional cost of exclusion from the social media platforms in high, particularly for creators of activist postings, often associated with political assertions of minority users (Powers et al., 2013). The freedom of expression of such communities is infringed through shadow banning. As no one is held responsible it makes emotional recuperation more difficult. The lack of feedback from the social media platforms, particularly among the users result in emotional dysregulation or a difficulty in managing emotional responses in accordance with the contextual demands (Rogier et al., 2024).

For the individual users the silent platforms are a failure of their own. Such instances ultimately lead to detrimental thinking patterns like repeated checking of the reach of the posts, resubmission and republishing of posts or immerse in self-critical thinking. It not only frustrates but psychologically damage the user (Da Silva Pinho et al., 2024).

4 Algorithmic inequality and emotional toll of shadow banning

The impact of shadow banning is not equally affected. The posts which are themed on sexuality, racial disturbances, social activism or body-positive are invariably censored. When these posts are not against the rules it reaches the users (Foster et al., 2021). There are many inherent structural inequalities due to algorithmic governance.

The subaltern and fringe groups in the society who are considered marginalised population always feel that their visibility is conditional and carefully crafted. The content provides belonging to queer and fat rights organisers negotiate their own space in the media for interactions and survival protests (Escobar-Viera et al., 2023). Some minority groups like queer had modest following on Instagram, but later when they discussed other general social issues there was sharp drop in views on all subsequent posts. The digital silencing occurs without formal notices and eventually it leads to distress and a temporary social media hiatus; an emotional erasure that sustains systemic silencing. It is a shame on individuals who feel that invisibility is a personal failure than a structure defect of media. As Covin (2021) emphasizes, shadow banning can lead to “unseen shame,” where users privately struggle with feelings of inadequacy, internalizing their online invisibility as a personal failing, despite the lack of explicit criticism from others.

Recent studies on digital exclusion reveal that algorithmic decisions can perpetuate existing social inequalities online, leaving users feeling unfairly penalized for their identity or views. The constant pressure to create content, coupled with the algorithm's silent devaluation of their voice, can be exhausting (Nair et al., 2024).

5 Shadow banning stems from inherent ambiguity?

When uncertainty increases anxiety and causes psychological distress it eventually leads to repetitive negative thoughts and thereby aggravate mental health concerns (Altan-Atalay et al., 2023).

The shadow banned users repeatedly fall into uncertainties even as they continue the futile exercise of selecting hashtags. The emotional exhaustion produces helplessness and bewilderment. The ambiguity linked to the posting in the social media can impact on trans-diagnostic factors linked to anxiety disorders and obsessive rumination. It renders the users more susceptible to distress (Pinciotti et al., 2021). The intolerant situation caused by uncertainty compels the users to quit the site because silence became unsustainable psychologically. Covin (2021) notes that this hidden shame in digital environments rarely has a reintegrated

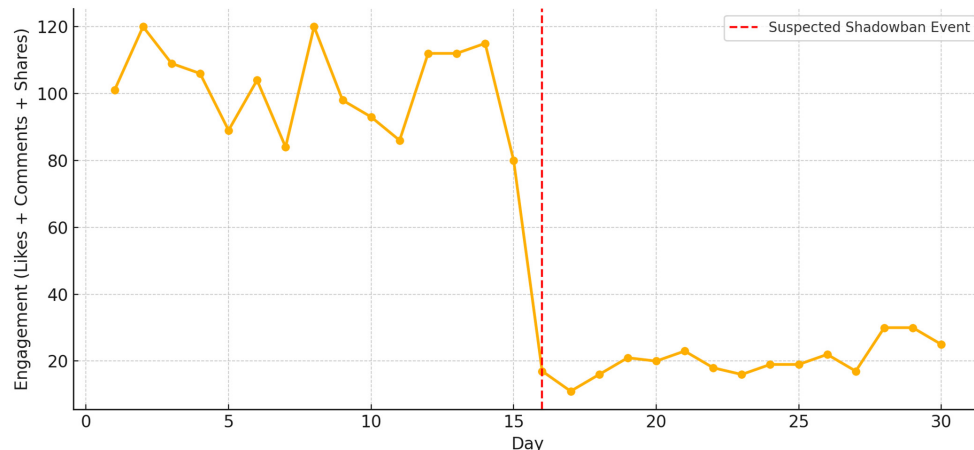


FIGURE 1

Sudden drop in engagement metrics after suspected shadowban event. This figure is based on a composite of self-reported case patterns drawn from user forums and anecdotal experiences. It is presented illustratively to depict a typical engagement trajectory following suspected shadow banning.

function. It isolates the user and increases his or her isolation. This corresponds with Jochan and Banerjee (2021) argument that shame in digital environments rarely has a reintegrated function; instead, it isolates the individual and deepens alienation.

The obscure element in the shadow banning process disrupts digital trust. Though the social media platforms claim freedom of expression they involve in stealth moderation that facilitates self-censorship and self-policing (Wang and Kim, 2023). This phenomenon can subtly persuade unwilling users into altering their tone and the themes, which eventually lead to emotional conformity due to prolonged limitation on the freedom of expression. The present study focuses on the urgent need for specific interventions to address the issues of ambiguity and emotional impact of algorithmic governance related to shadow banning. The negative psychological effects are far-reaching and it includes exclusion, shame and loss of trust. The transparency in the process of algorithmic governance and alleviation of deeply emotional and identity related constraints the users face online must be prioritized in finding solutions (Risius and Blasiak, 2024).

6 A humane platform design and emotional transparency needed

There is an invisible layer of shame in the social media platforms which highlights not only the fundamental issues of algorithmic transparency, but also the hidden psychological costs, ensuring that design responses attend to both external visibility and internal will-being (Covin, 2021). The social media platforms must acknowledge the damage caused by opaque algorithms and adopt transparent practices to reduce the emotional harm done to the users. If the reasons behind the content moderation decisions are explained the platforms can reduce user anxiety and build trust, creating a more open and reliable online environment (Jansen and Krämer, 2023).

Platforms should design with users' mental health in mind, incorporating features such as notifications, appeal options, and transparent explanations for content visibility. Fair governance

demands transparency, due process, and accountability, rather than unexplained penalties (Russ et al., 2014). Openness is not a technical remedy; it is a psychological necessity.

The mental health practitioners should include algorithmic exclusion within their conceptual framework of digital trauma (Barton et al., 2023). The sudden invisibility resulting from shadow banning can precipitate profound identity crises and emotional distress. Mental health professionals should be trained to address these concerns. Moreover, media literacy initiatives should extend beyond filter bubbles and misinformation to encompass the emotional consequences of algorithmic silence. Further research is warranted to explore the intersections between online trauma and other digital harms, such as cyberbullying, harassment, and community disintegration, to comprehensively understand the phenomenon's scope and implications (Delmonaco et al., 2024).

Despite being dismissed as conspiracy theories, shadow banning can cause real harm. We need more research that combines platform data, user experiences, and signs of psychological distress to understand the true mental health impact of being algorithmically suppressed online.

7 Conclusion: making the invisible visible

Visibility is validation in the social media platforms. Shadow banning turns invisibility into a weapon, and the silent treatment of the feed a tool of emotional coercion. Faith in the platforms erodes, shattering the users' perceptions of the self, and digital neurosis and self-doubt intensify (Van Noordt et al., 2015).

This opinion piece contends that shadow banning transcends content moderation, posing a significant psychological concern. By disrupting emotional regulation, exacerbating social inequalities, and fostering cognitive dissonance, it takes a profound toll on users. To mitigate this, media platforms must prioritize the emotional impact of algorithmic governance, lest users continue to experience silent suffering, overshadowed by both code and emotional distress.

To make the invisible visible is the first step toward justice—technical, social, and psychological. Let that apply not only to content, but to the human costs hidden behind the feed.

Author contributions

ST: Conceptualization, Validation, Visualization, Writing – original draft, Writing – review & editing. PM: Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Altan-Atalay, A., Tuncer, I., King, N., Önel, B., Sözeri, Y., and Tezel, S. (2023). Repetitive negative thinking during ambiguous situations: interactive roles of looming cognitive style and intolerance of uncertainty. *J. Behav. Ther. Exp. Psychiatry* 79:101840. doi: 10.1016/j.jbtep.2023.101840
- Barton, B. B., Reinhard, M. A., Goerigk, S., Wüstenberg, T., Musil, R., Ehring, T., et al. (2023). Association between the behavioral response during social exclusion and recalled childhood maltreatment. *Behav. Res. Ther.* 160:104232. doi: 10.1016/j.brat.2022.104232
- Covin, R. (2021). *The Private Shame of Social Media Users*. Medium. Available online at: <https://rogercovin.medium.com/internet-shame-a-common-and-unseen-private-experience-5bcceb757a> (Accessed September 15, 2025).
- Cross, E. A., Borland, J. M., Shaughnessy, E. K., Lee, S. D., Vu, V., Sambor, E. A., et al. (2025). Distinct subcircuits within the mesolimbic dopamine system encode the salience and valence of social stimuli. *Psychopharmacology* 242, 2219–2232. doi: 10.1007/s00213-025-06793-z
- Da Silva Pinho, A., Céspedes Izquierdo, V., Lindström, B., and Van Den Bos, W. (2024). Youths' sensitivity to social media feedback: a computational account. *Sci. Adv.* 10:eadp8775. doi: 10.1126/sciadv.adp8775
- Delmonaco, D., Mayworm, S., Thach, H., Guberman, J., Augusta, A., and Haimson, O. L. (2024). "What are you doing, TikTok?": how marginalized social media users perceive, theorize, and "prove" shadowbanning. *Proc. ACM Hum. Comput. Interact.* 8, 1–39. doi: 10.1145/3637431
- Escobar-Viera, C. G., Porta, G., Coulter, R. W. S., Martina, J., Goldbach, J., and Rollman, B. L. (2023). A chatbot-delivered intervention for optimizing social media use and reducing perceived isolation among rural-living LGBTQ+ youth: development, acceptability, usability, satisfaction, and utility. *Internet Interv.* 34:100668. doi: 10.1016/j.invent.2023.100668
- Foster, M. D., Tassone, A., and Matheson, K. (2021). Tweeting about sexism motivates further activism: a social identity perspective. *Br. J. Soc. Psychol.* 60, 741–764. doi: 10.1111/bjso.12431
- Jansen, M.-P., and Krämer, N. C. (2023). Balancing perceptions of targeting: an investigation of political microtargeting transparency through a calculus approach. *PLoS ONE* 18:e0295329. doi: 10.1371/journal.pone.0295329
- Jochan, G. M., and Banerjee, T. (2021). "Shaming in the internet era: evaluating the reintegrative function of shame in digital spaces," in *Shame 4.0*, eds. C.-H. Mayer, E. Vanderheiden, and P. T. P. Wong (Cham: Springer International Publishing), 431–454.
- Liu, J., Lee, D. N., and Stevens, E. M. (2023). Characteristics associated with young adults' intentions to engage with anti-vaping Instagram posts. *Int. J. Environ. Res. Public Health* 20:6054. doi: 10.3390/ijerph20116054
- Marta, R. F., Miletresia, and Fernandes, M. (2022). Reflections of today's teens behavior: from impressions on social media to cognitive dissonance. *Salus Cultura* 2, 81–91. doi: 10.55480/saluscultura.v2i1.49
- Nair, K., Mosleh, M., and Kouchaki, M. (2024). Racial minorities face discrimination from across the political spectrum when seeking to form ties on social media: evidence from a field experiment. *Psychol. Sci.* 35, 1278–1286. doi: 10.1177/09567976241274738
- Pinciotti, C. M., Riemann, B. C., and Abramowitz, J. S. (2021). Intolerance of uncertainty and obsessive-compulsive disorder dimensions. *J. Anxiety Disord.* 81:102417. doi: 10.1016/j.janxdis.2021.102417
- Polite-Corn, M., Pegg, S., Dickey, L., and Kujawa, A. (2024). Neural reactivity to social reward moderates the association between social media use and momentary positive affect in adolescents. *Affect. Sci.* 5, 281–294. doi: 10.1007/s42761-024-00237-1
- Powers, K. E., Somerville, L. H., Kelley, W. M., and Heatherton, T. F. (2013). Rejection sensitivity polarizes striatal-medial prefrontal activity when anticipating social feedback. *J. Cogn. Neurosci.* 25, 1887–1895. doi: 10.1162/jocn_a_00446
- Risius, M., and Blasiak, K. M. (2024). Shadowbanning: an opaque form of content moderation. *Bus. Inf. Syst. Eng.* 66, 817–829. doi: 10.1007/s12599-024-00905-3
- Rogier, G., Muzi, S., and Pace, C. S. (2024). Social media misuse explained by emotion dysregulation and self-concept: an ecological momentary assessment approach. *Cogn. Emot.* 38, 1261–1270. doi: 10.1080/02699931.2024.2363413
- Russ, A. L., Zillich, A. J., Melton, B. L., Russell, S. A., Chen, S., Spina, J. R., et al. (2014). Applying human factors principles to alert design increases efficiency and reduces prescribing errors in a scenario-based simulation. *J. Am. Med. Inform. Assoc.* 21, e287–e296. doi: 10.1136/amiajnl-2013-002045
- Van Noordt, S. J. R., White, L. O., Wu, J., Mayes, L. C., and Crowley, M. J. (2015). Social exclusion modulates event-related frontal theta and tracks ostracism distress in children. *NeuroImage* 118, 248–255. doi: 10.1016/j.neuroimage.2015.05.085
- Wang, S., and Kim, K. J. (2023). Content moderation on social media: does it matter who and why moderates hate speech? *Cyberpsychol. Behav. Soc. Netw.* 26, 527–534. doi: 10.1089/cyber.2022.0158
- Wikman, P., Moisala, M., Ylinen, A., Lindblom, J., Leikas, S., Salmela-Aro, K., et al. (2022). Brain responses to peer feedback in social media are modulated by valence in late adolescence. *Front. Behav. Neurosci.* 16:790478. doi: 10.3389/fnbeh.2022.790478

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.