2026, vol. 12, pp. 69-74
RESEARCH ARTICLE https://doi.org/10.5281/zenodo.22138865

AI-MEDIATED SELF: SOCIAL COMPARISON, SELF-DISCLOSURE, AND ALGORITHMIC VALIDATION IN DIGITAL COMMUNICATION

Andra Seceleanu[1], Alexandrina Cucu[1], Anca Belu[1]

[1] Universitatea Andrei Șaguna, Constanța, România

Abstract

This article examines how social networking sites (SNS) shape self-esteem, self-evaluation, self-disclosure, and social comparison, while extending the analysis to artificial-intelligence (AI)-mediated communication. As an integrative conceptual review, it draws on research in computer-mediated communication, social comparison, self-presentation, and emerging research on AI-mediated communication. The article proposes the AI-mediated self as a working conceptual lens for examining how intelligent systems increasingly participate in self-presentation, comparison, disclosure, and social validation. The argument is not that AI necessarily reduces well-being. Rather, risks may emerge when technological mediation changes the conditions under which individuals evaluate themselves, interpret feedback, experience authorship, or depend on algorithmically structured recognition. A recursive framework links selective self-presentation, social comparison, self-evaluation, disclosure, and validation, while distinguishing forms of augmentation that support human agency from forms of substitution that may weaken it. Future research directions are proposed for psychological resilience, platform design, digital literacy, and human-centered AI governance.

Keywords: social media, self-esteem, self-evaluation, self-disclosure, social comparison, artificial intelligence, AI-mediated communication, authenticity, algorithmic validation

1. Introduction

The explosive growth of social networking sites (SNS) has radically transformed how individuals present themselves and interact. Platforms such as Facebook, Instagram, X, and TikTok have become integral to daily life, offering spaces for self-expression, social comparison, and relationship maintenance. Yet, these digital environments also pose risks to psychological well-being. Most users today are "digital natives," having grown up with digital technology. Carr and Hayes (2015) define social media as "Internet-based channels that allow users to interact opportunistically and present themselves selectively, either in real time or asynchronously, both for broad and narrow audiences, and that derive value from user-generated content and the perception of interaction with others." This highlights SNS as platforms for connection and curated self-presentation.

SNS popularity stems not only from staying connected but also from managing impressions and entertainment (Krämer & Winter, 2008). Social factors like group identification and collective self-esteem predict participation in social networks. Users now engage across multiple platforms, selecting channels based on communicative intent rather than technological or economic constraints, illustrating the complex interplay between social and technological factors. Interpersonal communication competence, the perceived ability to manage relationships effectively, requires reevaluation in online multimedia communication. Interpersonal communication is central to conveying social support, which enhances quality of life. While supportive communication in face-to-face contexts is well studied, less is known about how it operates online.

Contrary to views of online interaction as impersonal, evidence suggests computer-mediated communication (CMC) can facilitate supportive exchanges effectively. Walther and Parks (2002) argue the Internet serves as a successful medium for social support through thousands of online groups. Burleson and Goldsmith (1998) differentiate comforting communication, which addresses emotional distress, from broader social support involving caring relationships. Drawing on Lazarus's cognitive theory of emotion (Lazarus, 1991, 1999; Lazarus & Folkman, 1984), effective comforting communication promotes adaptive cognitive reappraisals that alleviate distress.

Online environments may offer safer spaces to discuss personal topics, as social interaction online often entails fewer personal and social risks than face-to-face conversation (Caplan, 2003). However, online self-image management involves behaviors such as idealized self-presentation, social comparison, seeking validation through likes and comments, exposure to negative feedback, and compulsive message checking. These behaviors can contribute to psychological strain, especially when users feel pressured to maintain perfection or when validation is lacking.

These established mechanisms provide the psychological foundation for the present article. The question developed here is narrower than a general theory of societal security: how are self-presentation, social comparison, self-evaluation, self-disclosure, and validation transformed when AI becomes an active participant in communication?

2. Psychological Processes in Social Networking Sites

The intertwined concepts of self-esteem, self-evaluation, self-disclosure, and social comparison are central to understanding psychological well-being in the context of SNS use. Although analytically distinct, these processes are mutually reinforcing: self-esteem shapes how feedback is interpreted, social comparison affects self-evaluation, self-disclosure determines what information enters the interpersonal environment, and platform responses influence subsequent disclosure and comparison.

2.1 Self-Esteem and Self-Evaluation

The sociometer perspective is also relevant because self-esteem can function as an interpersonal monitor of perceived social inclusion and relational value (Leary et al., 1995).

Self-esteem reflects an individual's overall sense of self-worth. Although relatively stable, it can fluctuate in response to social feedback, particularly in interactive environments such as social media (Mitrofan & Gâtej, 2019). Meta-analytic evidence indicates a small negative association between SNS use and self-esteem, but the direction and magnitude of the relationship vary according to the type of use, the content encountered, and individual differences (Saiphoo et al., 2020). Passive observation and repeated exposure to idealized content may be especially relevant when they intensify feelings of inadequacy.

Social rejection is associated with lower self-esteem and may motivate individuals to seek SNS communities that validate their self-image and provide support. Self-evaluation, understood as the assessment of traits and abilities relative to personal or social standards, is bidirectionally linked to self-esteem. Higher self-esteem is generally associated with more favorable self-evaluations, whereas lower self-esteem is more often aligned with negative self-assessment. Exposure to upward comparison targets on social media is associated with less favorable self-evaluations and more negative emotions (McComb et al., 2023).

2.2 Self-Esteem and Social Comparison

Festinger's (1954) social comparison theory proposes that individuals evaluate themselves by comparing themselves with others, especially when objective standards are unavailable. Social media can facilitate upward comparison because users frequently encounter selectively presented or idealized content. Such exposure is associated with negative self-evaluation and lower self-esteem for some users (Saiphoo et al., 2020). A meta-analysis of Facebook social comparison also reported an association between frequent comparison and lower well-being, including feelings of inferiority concerning appearance, lifestyle, and achievement (Yang et al., 2019).

2.3 Self-Disclosure's Role in Self-Esteem and Self-Evaluation

Self-disclosure, the sharing of personal information, plays a nuanced role online. Individuals with higher self-esteem may disclose more freely, whereas those with lower self-esteem may hesitate because of anticipated judgment. Positive feedback received for posted content can be associated with a temporary increase in self-esteem, although this relationship depends on personal factors such as a stable sense of purpose (Burrow & Rainone, 2017). Positive responses may support more favorable self-evaluation, while negative or absent feedback may increase vulnerability when the user's sense of worth is strongly contingent on platform reactions.

2.4 Interaction Between Social Comparison and Self-Disclosure

SNS encourage a recursive cycle of selective self-disclosure and social comparison. Users observe idealized disclosures, compare themselves with the people who produced them, and may then construct their own disclosures in response. A flattering post may temporarily enhance the poster's self-evaluation when it receives positive feedback while simultaneously serving as an upward comparison target for viewers. Those viewers may subsequently curate their own content in an attempt to restore status, recognition, or belonging.

Selective disclosure may support identity exploration, community formation, and social support. Online groups can connect individuals with similar experiences and provide resources unavailable in their immediate environment (Ellison et al., 2007). Concern arises when identity expression becomes strongly contingent on

engagement metrics, when social value is repeatedly interpreted through algorithmic visibility, or when users lack of alternative sources of recognition and belonging.

Risk is therefore more accurately represented as a feedback loop involving user characteristics, platform affordances, encountered content, emotional interpretation, and subsequent behavior. AI-based personalization may accelerate this loop by learning which content captures attention or elicits a response, even when the system has no explicit intention to reduce well-being.

3. From Platform Mediation to AI-Mediated Communication

Beyond the psychological effects associated with social media, artificial intelligence (AI) may reshape individual well-being and digital self-processes by influencing cognitive, motivational, and identity-related conditions. AI-mediated communication differs from traditional computer-mediated communication because an intelligent system may edit, rank, predict, recommend, or generate messages. These functions can improve translation, accessibility, clarity, and creativity, but they may also blur responsibility for expression and make authorship more difficult to evaluate (Hancock et al., 2020).

Conversational systems may also be experienced as responsive social partners rather than only as tools. Emerging evidence suggests that psychosocial outcomes vary according to communication mode, conversation type, attachment tendencies, trust, and intensity of use. In one recent randomized study, higher voluntary use was associated with greater loneliness, emotional dependence, problematic use, and reduced real-world socialization; because the study was released as a preprint and use intensity was not experimentally manipulated, causal conclusions remain premature (Fang et al., 2025). The security paradigm should therefore expand beyond physical and economic concerns to include cognitive-societal dimensions that protect autonomy, identity continuity, and meaningful participation in AI-mediated environments (Buzan et al., 1998; Ienca, 2023).

3.1 AI-Mediated Self-Presentation and Authorship

AI-mediated communication differs from traditional computer-mediated communication because an intelligent system may edit, rank, predict, recommend, or generate messages (Hancock et al., 2020). These functions can improve translation, accessibility, personalization, and creative assistance, but they also alter the relationship between sender, message, and receiver.

When AI assists with wording, images, profiles, or other forms of self-presentation, the relevant psychological issue is not simply whether content is artificial. It is also whether users experience meaningful authorship, authenticity, and ownership of the resulting self-presentation. Experimental work has shown that perceived AI involvement in profile authorship can reduce trustworthiness under mixed human/AI conditions (Jakesch et al., 2019), while recent theory argues that AI-mediated communication reshapes selective self-presentation through co-creation and adaptability in message production (Zhu et al., 2026).

Related conceptual work further suggests that AI-mediated communication can function as an extension of the self by modifying, augmenting, or generating communication, with implications for how users present and understand themselves (Campbell et al., 2025).

3.2 Algorithmically Amplified Social Comparison

Festinger's (1954) social comparison theory proposes that individuals evaluate themselves by comparing themselves with others, especially when objective standards are unavailable. Social media can facilitate upward comparison because users frequently encounter selectively presented or idealized content. Such exposure is associated with negative self-evaluation and lower self-esteem for some users (Saiphoo et al., 2020). A meta-analysis of Facebook social comparison also reported an association between frequent comparison and lower well-being, including feelings of inferiority concerning appearance, lifestyle, and achievement (Yang et al., 2019).

Risk is therefore more accurately represented as a feedback loop involving user characteristics, platform affordances, encountered content, emotional interpretation, and subsequent behavior. AI-based personalization may accelerate this loop by learning which content captures attention or elicits a response, even when the system has no explicit intention to reduce well-being.

In AI-mediated environments, comparison targets may additionally be AI-enhanced, AI-generated, or algorithmically amplified. This extends the established social-comparison problem: the user may be comparing the self not only with selectively presented human lives, but also with technologically constructed or disproportionately promoted standards.

3.3 AI-Mediated Self-Disclosure and Algorithmic Validation

Self-disclosure, the sharing of personal information, plays a nuanced role online. Individuals with higher self-esteem may disclose more freely, whereas those with lower self-esteem may hesitate because of anticipated judgment. Positive feedback received for posted content can be associated with a temporary increase in self-esteem, although this relationship depends on personal factors such as a stable sense of purpose (Burrow & Rainone, 2017). Positive responses may support more favorable self-evaluation, while negative or absent feedback may increase vulnerability when the user's sense of worth is strongly contingent on platform reactions.

Selective disclosure may support identity exploration, community formation, and social support. Online groups can connect individuals with similar experiences and provide resources unavailable in their immediate environment (Ellison et al., 2007). Concern arises when identity expression becomes strongly contingent on engagement metrics, when social value is repeatedly interpreted through algorithmic visibility, or when users lack of alternative sources of recognition and belonging.

The concept of algorithmic validation is used here to distinguish perceived recognition that occurs in environments where visibility, reach, and feedback are partly structured by recommendation, ranking, personalization, or generative systems. Such validation may still be experienced as interpersonal even when the opportunity to receive it has been strongly shaped by platform or AI design. This distinction is also relevant to trust and authenticity: AI-mediated communication can complicate epistemic trust in social-media environments (Sahebi & Formosa, 2025), and recent experimental evidence reports lower perceived trustworthiness and authenticity for AI-assisted or AI-authored communication than for human-authored communication, together with a disclosure paradox surrounding AI use (Sahebi et al., 2026).

4. The AI-Mediated Self: A Recursive Conceptual Framework

The proposed framework retains the relationships already identified in the social-media literature but considers AI as a potential intervening layer that may influence multiple stages of this process. Selective self-presentation may shape exposure to comparison targets, which in turn can influence self-evaluation and the interpretation of social feedback and validation. These experiences may subsequently affect how individuals disclose information and present themselves online, creating a recursive process in which AI-mediated personalization, recommendation, and communication can intervene at different points. The AI-mediated self is therefore proposed as a working conceptual lens rather than as a new personality construct. It refers to situations in which AI systems increasingly participate in how individuals present themselves, encounter comparison standards, disclose personal information, receive feedback, and interpret social recognition. This formulation is consistent with emerging work on AI-mediated self-extension and AI-assisted selective self-presentation, while retaining a narrower focus on the recursive psychological processes examined here (Campbell et al., 2025; Zhu et al., 2026).

The framework should be understood as a set of testable relationships rather than as evidence that AI-mediated communication inevitably produces psychological harm.

The relevant distinction is not between human action and machine assistance in the abstract, but between forms of augmentation that support human agency and forms of substitution that systematically weaken it.

5. Governance and Design Implications

Platform governance and the legal security of digital ecosystems are important to the preservation of democratic safeguards (Ghervase et al., 2026; Papari et al., 2026).

Within the European Union, the Digital Services Act requires very large platforms to assess and mitigate systemic risks and establishes duties related to recommender-system transparency, auditing, and fundamental rights, particularly through Articles 34 and 35 (Regulation (EU) 2022/2065, 2022). The Artificial Intelligence Act establishes a risk-based framework, prohibits certain manipulative or exploitative practices under Article 5, and introduces transparency obligations for particular AI systems under Article 50 (Regulation (EU) 2024/1689, 2024).

Platform governance should promote transparency, accountability, and democratic safeguards while preserving beneficial uses of digital technology. Governments and organizations could require algorithmic impact assessments that consider potential psychological and social effects, support independent access to data for systemic-risk research, and establish oversight mechanisms capable of auditing compliance. Digital-literacy programs should include cognitive self-defense: understanding personalization, recognizing emotional targeting, calibrating trust in AI, verifying generated content, and maintaining skills that should not be fully delegated.

Participatory policy-making mechanisms, including citizen panels, stakeholder consultations, and the involvement of vulnerable groups, may improve legitimacy and identify harms that technical audits overlook. Regulation should remain proportionate and design-sensitive. Its purpose should be to distinguish augmentation from dependency, support from substitution, and informed persuasion from manipulation, rather than to restrict the positive potential of social media and AI.

For the present framework, these governance questions are relevant insofar as platform and AI-system design shape the conditions of self-presentation, comparison, disclosure, visibility, and validation. They are not treated here as evidence for a separate national-security model.

6. Future Research

Future research should directly test whether AI-mediated communication changes the established relationships among self-presentation, social comparison, self-evaluation, self-disclosure, and feedback. Longitudinal surveys, experience-sampling, qualitative interviews, experimental comparisons, mixed-methods research, and cross-cultural studies could clarify these relationships.

Particularly useful comparisons include human-generated versus AI-generated comparison targets; AI-assisted versus unassisted self-presentation; interpersonal versus algorithmically structured feedback; and disclosure to human versus conversational-AI partners. Research should also examine whether attachment tendencies, trust, intensity of use, self-esteem, and social-comparison orientation moderate these effects, given emerging evidence that psychosocial outcomes of chatbot use vary according to communication mode and patterns of use (Fang et al., 2025).

7. Conclusion

The convergence of social media dynamics and AI-mediated communication reveals a new stage in the study of digital self-processes.

The relationships among self-esteem, self-evaluation, self-disclosure, and social comparison remain central, but AI increasingly changes the environment in which these processes occur by participating in message production, personalization, visibility, comparison, and feedback.

The relevant distinction is not between human action and machine assistance in the abstract, but between forms of augmentation that support human agency and forms of substitution that systematically weaken it.

This article has several limitations. It is a conceptual synthesis rather than a systematic review; the selected literature supports theory development. Research on social media reports small average effects and substantial heterogeneity, while evidence on generative and relational AI remains emergent.

The proposed AI-mediated self framework therefore requires direct empirical testing and should not be interpreted as evidence that AI necessarily undermines authenticity, self-esteem, or psychological well-being.

8. References

Burleson, B. R., & Goldsmith, D. J. (1998). How the comforting process works: Alleviating emotional distress through conversationally induced reappraisals. In P. A. Andersen & L. K. Guerrero (Eds.), Handbook of communication and emotion: Research, theory, applications, and contexts (pp. 245–280). Academic Press. https://doi.org/10.1016/B978-012057770-5/50011-4

Burrow, A. L., & Rainone, N. (2017). How many likes did I get? Purpose moderates links between positive social media feedback and self-esteem. Journal of Experimental Social Psychology, 69, 232–236. https://doi.org/10.1016/j.jesp.2016.09.005

Campbell, S. W., Ellison, N. B., & Ross, M. Q. (2025). Extending the self through AI-mediated communication: Functional, ontological, and anthropomorphic extensions. Communication and Change, 1, Article 3. https://doi.org/10.1007/s44382-025-00003-2

Caplan, S. E. (2003). Preference for online social interaction: A theory of problematic Internet use and psychosocial well-being. Communication Research, 30(6), 625–648. https://doi.org/10.1177/0093650203257842

Carr, C. T., & Hayes, R. A. (2015). Social media: Defining, developing, and divining. Atlantic Journal of Communication, 23(1), 46–65. https://doi.org/10.1080/15456870.2015.972282

Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The benefits of Facebook "friends": Social capital and college students' use of online social network sites. Journal of Computer-Mediated Communication, 12(4), 1143–1168. https://doi.org/10.1111/j.1083-6101.2007.00367.x

Fang, C. M., Liu, A. R., Danry, V., Lee, E., Chan, S. W. T., Pataranutaporn, P., Maes, P., Phang, J., Lampe, M., Ahmad, L., & Agarwal, S. (2025). How AI and human behaviors shape psychosocial effects of chatbot use: A longitudinal randomized controlled study [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2503.17473

Festinger, L. (1954). A theory of social comparison processes. Human Relations, 7(2), 117–140. https://doi.org/10.1177/001872675400700202

Ghervase, D. G., Ilincuța, A. M., & Geamalinga, N. R. (2026). The legal security of digital ecosystems. Veredas do Direito, 23(5), e235572. https://doi.org/10.18623/rvd.v23.5572

Hancock, J. T., Naaman, M., & Levy, K. (2020). AI-mediated communication: Definition, research agenda, and ethical considerations. Journal of Computer-Mediated Communication, 25(1), 89–100. https://doi.org/10.1093/jcmc/zmz022

Jakesch, M., French, M., Ma, X., Hancock, J. T., & Naaman, M. (2019). AI-mediated communication: How the perception that profile text was written by AI affects trustworthiness. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (pp. 1–13). Association for Computing Machinery. https://doi.org/10.1145/3290605.3300469

Krämer, N. C., & Winter, S. (2008). Impression management 2.0: The relationship of self-esteem, extraversion, self-efficacy, and self-presentation within social networking sites. Journal of Media Psychology, 20(3), 106–116. https://doi.org/10.1027/1864-1105.20.3.106

Lazarus, R. S. (1991). Emotion and adaptation. Oxford University Press.

Lazarus, R. S. (1999). Stress and emotion: A new synthesis. Springer Publishing Company.

Lazarus, R. S., & Folkman, S. (1984). Stress, appraisal, and coping. Springer Publishing Company.

Leary, M. R., Tambor, E. S., Terdal, S. K., & Downs, D. L. (1995). Self-esteem as an interpersonal monitor: The sociometer hypothesis. Journal of Personality and Social Psychology, 68(3), 518–530. https://doi.org/10.1037/0022-3514.68.3.518

McComb, C. A., Vanman, E. J., & Tobin, S. J. (2023). A meta-analysis of the effects of social media exposure to upward comparison targets on self-evaluations and emotions. Media Psychology, 26(5), 612–635. https://doi.org/10.1080/15213269.2023.2180647

Mitrofan, L., & Gâtej, E.-R. (2019). Psihologie și psihopatologie socială. SPER.

Papari, A. C., Ghervase, D. G., Ilincuța, A. M., Geamalinga, N. R., & Șunda, I. (2026). Neurolaw and social media governance: Neurorights, platform design, and the EU regulatory response: A comprehensive scientific literature review. Veredas do Direito, 23(9), e6859. https://doi.org/10.18623/rvd.v23.6859

Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act). (2022). Official Journal of the European Union, L 277, 1–102. https://eur-lex.europa.eu/eli/reg/2022/2065/oj

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). (2024). Official Journal of the European Union, L 2024/1689. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Sahebi, S., & Formosa, P. (2025). The AI-mediated communication dilemma: Epistemic trust, social media, and the challenge of generative artificial intelligence. Synthese, 205, Article 128. https://doi.org/10.1007/s11229-025-04963-2

Sahebi, S., Formosa, P., & Bankins, S. (2026). The AI penalty and disclosure paradox: Trust, authenticity and knowledge uptake in AI-mediated communication. Computers in Human Behavior: Artificial Humans, 8, Article 100304. https://doi.org/10.1016/j.chbah.2026.100304

Saiphoo, A. N., Dahoah Halevi, L., & Vahedi, Z. (2020). Social networking site use and self-esteem: A meta-analytic review. Personality and Individual Differences, 153, Article 109639. https://doi.org/10.1016/j.paid.2019.109639

Walther, J. B., & Parks, M. R. (2002). Cues filtered out, cues filtered in: Computer-mediated communication and relationships. In M. L. Knapp & J. A. Daly (Eds.), Handbook of interpersonal communication (3rd ed., pp. 529–563). Sage.

Yang, F.-R., Wei, C.-F., & Tang, J.-H. (2019). Effect of Facebook social comparison on well-being: A meta-analysis. Journal of Internet Technology, 20(6), 1829–1836. https://doi.org/10.3966/160792642019102006013

Zhu, R. (C.), Markowitz, D. M., & Van Der Heide, B. (2026). Presenting the augmented self: Extending sender effects of the hyperpersonal model in artificial intelligence-mediated communication. Annals of the International Communication Association, wlag029. https://doi.org/10.1093/anncom/wlag029