Human and AI-Generated Profile Pictures: A Comparative Study of Visual Identity Construction

Ankita Kaushal[1], Dr. Pankaj Garg[2], Dr. Ripudaman Singh[3], Dr. Ranjit Singh Chopra[4]
[1] *PhD Scholar, Chitkara School of Mass Communication Chitkara University, Punjab, India*
[2] *Professor, Chitkara School of Mass Communication Chitkara University, Punjab, India*
[3,4] *Assistant Professor, Chitkara School of Mass Communication Chitkara University, Punjab, India*

Abstract*—Profile pictures function as condensed identity interfaces: they precede interaction, organize first impressions, and often remain publicly visible when other profile information is restricted. Generative artificial intelligence has altered this function by enabling realistic faces that signify personhood without documenting an existing person. This chapter reports an original comparative secondary analysis of 24 empirical studies and datasets—15 examining human or edited profile photographs and nine examining AI-generated faces and synthetic profiles. Evidence units were coded for identity-signalling strength, direction of trust or engagement effects, and authenticity risk. Human and edited photographs produced a high mean identity-signal score (1.87/2) and positive trust or engagement effects in 47% of studies, but their mean authenticity-risk score remained moderate (0.67/2). AI-generated faces produced an equally high identity-signal score (1.89/2) and positive first-impression effects in 44% of studies, while receiving the maximum mean authenticity-risk score (2.00/2). Experimental evidence shows that synthetic faces can be indistinguishable from real faces, judged more trustworthy, and—in the case of White AI faces—perceived as more human than photographs of actual people. In-the-wild evidence simultaneously documents their use in scams, spam, coordinated amplification, and fabricated accounts. The chapter concludes that AI changes the profile picture from selective self-presentation into potentially synthetic identity establishment. It proposes the TRACE model—Transparency, Referential integrity, Audience-context fit, Consent and control, and Engagement accountability—for platform governance and responsible visual identity practice.

***Index Terms*—Artificial Intelligence, Digital Identity, Identity Establishment, Online Engagement, Profile Pictures, Social Networking Sites.**

I. INTRODUCTION

On social networking sites, the profile picture is not merely decorative. It is a persistent visual claim about who is speaking, applying, dating, commenting, or requesting connection. Because it is usually encountered before biography, posts, or relational history, it operates as an identity shortcut. Photographic cues can increase perceived social presence and competence on professional networks, influence judgments of attractiveness and approachability, and interact with popularity and peer-generated information to shape impressions [1]–[3]. Profile images therefore connect self-presentation with platform credibility and online engagement.

Human profile photographs have always involved selection. Users choose expressions, camera angles, companions, backgrounds, filters, and moments that foreground desired attributes. Research on online dating found that users balanced attractiveness enhancement against the risk that a photograph would later be judged deceptive [4]. Facebook and Instagram studies likewise show that visual profiles can communicate actual personality, idealization, or ambiguous combinations of both, depending on context and the observer's interpretive frame [5], [6]. The central issue is not whether a picture is perfectly natural, but whether it retains a referential link to the person whose identity it represents.

AI-generated profile pictures introduce a categorical change. A synthetic face can be photorealistic, socially attractive, demographically specified, and emotionally legible while referring to no embodied individual. Nightingale and Farid [7] found performance near chance when participants

distinguished AI-synthesized from real faces and reported higher trustworthiness for synthetic faces. Miller et al. [8] subsequently demonstrated AI hyperrealism: White AI faces were judged human more frequently than actual White human faces, with the least accurate observers often the most confident. The profile picture can thus maximize the visual conventions of authenticity without possessing biographical authenticity.

This distinction matters because profile pictures are used across interpersonal, professional, political, and commercial settings. A synthetic portrait may be a legitimate privacy-preserving avatar, a creative identity, a branded virtual persona, or an accessibility tool. It may also support impersonation, fabricated expertise, romance fraud, astroturfing, spam, and coordinated influence. Yang, Singh, and Menczer [9] identified 1,420 X accounts using GAN-generated profile faces and estimated a lower-bound prevalence of 0.021%–0.044% among active profiles—approximately 10,000 daily active accounts. The same visual affordance therefore enables identity experimentation and identity deception.

The present chapter moves beyond a narrative review by constructing a comparative evidence matrix. It asks how human and AI-generated profile pictures differ in the kinds of identity they establish, the impressions and engagement they generate, and the risks they create for authenticity and platform trust. The analysis treats immediate trustworthiness as distinct from warranted trust: a face may look trustworthy while the profile remains unverifiable. This separation is necessary for understanding why synthetic images can perform well perceptually while damaging the informational integrity of social networks.

II. RESEARCH OBJECTIVES

The chapter addresses four objectives:

1) To compare the identity-signalling functions of human, edited, avatar-based, and AI-generated profile pictures.
2) To quantify the direction of reported effects on credibility, attractiveness, approachability, participation, and engagement.
3) To identify the authenticity, bias, privacy, and deceptive-use risks associated with synthetic visual identities.
4) To develop an applied framework for responsible profile-picture design, disclosure, and platform governance.

III. METHODOLOGY

A. Comparative Secondary-Data Design

The study used structured comparative secondary analysis. Peer-reviewed experiments, observational studies, computational analyses, and field datasets were included when they examined: (a) profile photographs, selfies, avatars, or AI-generated faces; (b) identity, personality, trust, attractiveness, credibility, employability, engagement, or authenticity; and (c) social-network, dating, professional, or synthetic-profile contexts. Purely technical generative-model papers were excluded unless they evaluated detection in conditions relevant to online profile images.

The coded corpus comprised 24 evidence units: 15 human, edited, or avatar studies and nine AI-face or synthetic-profile studies. Each publication contributed one unit, preventing multi-experiment papers from dominating the descriptive results. The evidence matrix also covered professional screening and human-resource decisions [10], [11]; foundational social-network, self-presentation, self-esteem, deepfake, credibility, and synthetic-face research [12]–[17]; personality, avatar, and face-inference studies [18]–[22]; detection and deceptive self-presentation research [23]–[27]; and selfie-related visual identity practices [28].

B. Coding and Analysis

Three variables were coded. Identity-signalling strength ranged from 0 (minimal information) to 2 (strong, socially interpretable cues). Trust/engagement direction was coded −1 for reduced credibility or harmful engagement, 0 for mixed, null, or diagnostic findings, and +1 for increased credibility, attraction, trust, or interaction intention. Authenticity risk ranged from 0 (low referential concern) to 2 (non-existent identity or high deception potential). These descriptive scores are not standardized effect sizes or a meta-analysis.

TABLE I EVIDENCE CODING RULES

Variable	0	1	2
Identity signal	Minimal identity information	Some identity or presence cues	Strong personality, status, demographic, or relational cues
Trust/engagement	Mixed or null (direction coded separately)	Positive or negative directional evidence	Not applicable to scale
Authenticity risk	Photograph clearly refers to profile owner	Selective editing, stereotyping, or contextual ambiguity	Synthetic/non-existent identity or high deception potential

IV. RESULTS AND FINDINGS

A. Human Profile Pictures Establish Identity Through Selective Evidence

The human-photo studies yielded a mean identity-signal score of 1.87 out of 2. This high value reflects the density of social information carried by facial appearance, pose, setting, companions, photographic style, and surrounding profile cues. Back et al. [5], using 236 profile owners across U.S. and German samples, found observer judgments aligned more closely with actual than idealized personality. Yet later Instagram evidence found inconsistent correspondence between account holders' self-ratings and observers' judgments, showing that visual identity is neither transparently authentic nor uniformly idealized [6].

Profile-picture effects were strongly contextual. In a LinkedIn experiment, the presence of a profile photograph increased perceived social attractiveness and competence compared with a no-picture condition [3]. Scott's [2] experiment with 102 undergraduates showed that popular Facebook targets were perceived as more socially and physically attractive, extroverted, and approachable. These results indicate that a picture is interpreted within a profile ecology: friend counts, tagged photographs, wall activity, and peer appearance alter the meaning attributed to the face [1], [29].

Computational studies demonstrate that profile pictures expose more than users may intend. Wei and Stillwell [30] analysed 1,122 Facebook users and found that smiling and wearing glasses increased perceived intelligence without corresponding to measured intelligence. Segalin et al. [31] showed that visual features could predict aspects of personality and that computer models outperformed averaged human judgments for extraversion and neuroticism. Profile images therefore support identity establishment partly through valid signals and partly through stereotypes. Public visibility turns visual self-presentation into inferential data.

Enhancement creates a credibility trade-off. Hancock and Toma's [4] analysis of 54 online-dating photographs documented selective photographic self-presentation while highlighting users' need to remain credible for future face-to-face meetings. Appel et al. [32] found that beautified male Tinder photographs increased attractiveness and dating intention but reduced trustworthiness. In the coded evidence, only one human-photo study produced a clearly negative trust direction, but five were mixed or context-dependent. The average authenticity-risk score was 0.67, indicating that human images usually preserve referential identity while permitting substantial optimization.

B. AI-Generated Faces Maximize Social Legibility but Break Referential Identity

AI-generated faces achieved a mean identity-signal score of 1.89, virtually identical to human and edited photographs. Their average authenticity-risk score, however, was 2.00. This combination defines the central empirical difference: synthetic profile pictures preserve social legibility while eliminating the necessary connection between image and person. The viewer sees age, gender presentation, expression, gaze, apparent ethnicity, and personality cues, but these cues may describe a generated statistical composite rather than an accountable individual.

Perceptual experiments show why this difference is difficult to police through ordinary judgment. Nightingale and Farid [7] reported approximately chance-level discrimination between real and StyleGAN2 faces; limited training increased accuracy only modestly, while synthetic faces were rated 7.7% more trustworthy. Miller et al. [8] replicated and

qualified the problem. In Experiment 1 (N=124), White AI faces were judged human more often than actual human faces. In Experiment 2 (N=610), facial proportionality, perceived aliveness, familiarity, memorability, attractiveness, and smoothness explained much of the classification pattern; the visual-attribute model accounted for 62% of variance in judgments of humanness. Hyperrealism was also racially uneven because the training distribution disproportionately represented White faces.

Newer evidence indicates that the trust advantage is not disappearing. McGuire et al. [33] found that diffusion-model faces, despite lower photorealism than GAN faces, were rated as more trustworthy than GAN and real faces. Rehman et al. [34], in an experiment with 184 participants, found that targeted training moderately improved synthetic-face detection but also increased false alarms for authentic faces. Visual literacy therefore helps, but a strategy based only on users spotting artifacts risks both missed fakes and suspicion toward genuine people.

Field evidence moves the problem from laboratory perception to platform integrity. Yang et al. [9] identified 1,420 accounts using GAN-generated faces for scams, spam, and coordinated message amplification. Porcile et al. [35] and Mundra et al. [36] developed detection methods tailored to generated profile photographs, while Ricker et al. [37] analysed AI-generated faces in real-world conditions. Such studies show that the relevant object is not only the image. Account age, posting patterns, network coordination, provenance, and behavioural consistency are necessary for evaluating identity credibility.

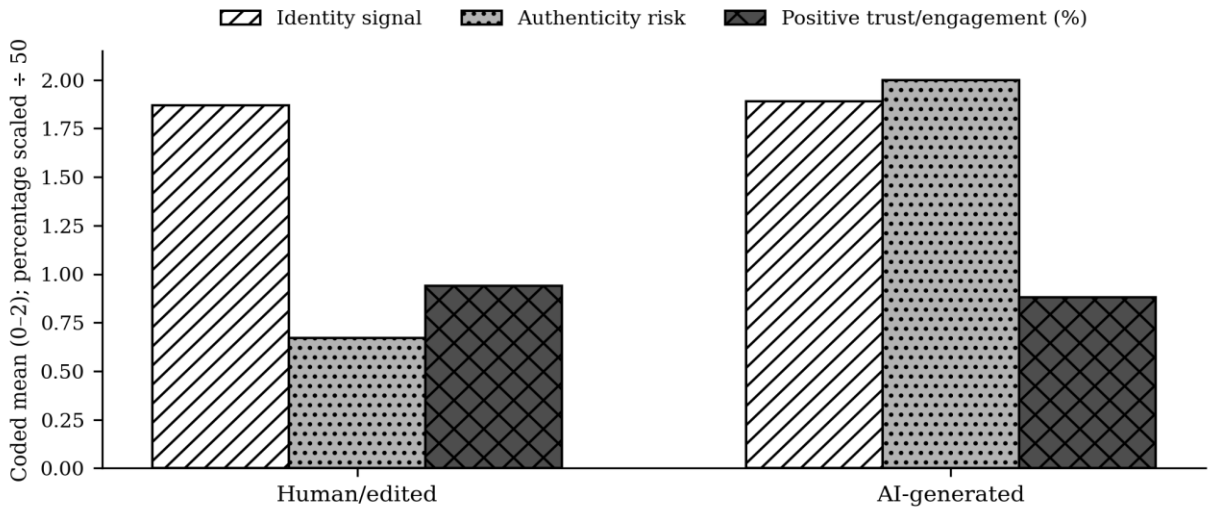

Fig. 1. Comparative coded evidence for human/edited and AI-generated profile pictures.

Note—Positive trust/engagement shares are scaled by dividing percentages by 50 so that 100% equals 2 on the chart. Scores are descriptive author coding, not standardized effect sizes.

TABLE II COMPARATIVE SECONDARY-DATA RESULTS

Indicator	Human/edited (n=15)	AI-generated (n=9)	Interpretation
Mean identity signal (0–2)	1.87	1.89	Both forms create socially rich identity cues.
Positive trust/engagement direction	47%	44%	Synthetic faces can perform comparably in immediate impressions.
Negative direction	7%	22%	AI evidence includes harmful platform use, not only perception.
Mean authenticity risk (0–2)	0.67	2.00	The decisive divergence is referential integrity.

C. Engagement Is Produced by Cue Optimization, Not Guaranteed Authenticity

Across both groups, positive engagement arose from social presence, attractiveness, familiarity, and contextual fit. Human users select photographs that make these cues visible; generative models statistically optimize them. The similarity in positive-direction percentages—47% for human/edited studies and 44% for AI studies—does not mean the forms are equivalent. It means that perceptual response is vulnerable to surface optimization. A synthetic face can secure attention or initial trust without the social costs that constrain human self-presentation, such as consistency across photographs, offline encounters, embodied history, and accountability.

The findings therefore separate three layers of digital identity. Representational identity concerns what the image depicts. Referential identity concerns whether the depicted person exists and is linked to the account. Relational identity concerns whether the profile's behaviour, network, and interaction history are coherent over time. Human photographs may distort representation while retaining reference. AI-generated profile pictures can make representation highly plausible while severing reference. Platform trust depends most heavily on the latter two layers.

Disclosure changes interpretation but is not sufficient by itself. A disclosed AI avatar can support privacy, play, pseudonymity, creative work, or virtual-influencer branding. An undisclosed synthetic portrait that implies a real person can misappropriate the credibility convention of photography. Governance should therefore focus on implied identity claims rather than prohibit synthetic imagery categorically. The ethical question is whether audiences are reasonably able to understand who or what is represented and whether a responsible actor can be identified.

V. DISCUSSION

A. From Self-Presentation to Synthetic Identity Establishment

Traditional profile-picture research assumes that users select among images of themselves or avatars they control. The image is a performance, but the performer exists. Generative AI introduces identity establishment without a prior embodied referent. This changes the analytical unit from impression management to identity infrastructure. The relevant outcomes include not only attractiveness and credibility but also provenance, consent, impersonation, demographic bias, and downstream accountability.

The results also challenge the equation of trustworthiness with trust. Face-perception research shows that observers infer trust from rapid visual cues, often within fractions of a second [38], [39]. These impressions can predict consequential judgments despite weak validity. AI systems exploit the same cue space by generating average, symmetrical, familiar-looking faces. Warranted trust requires evidence outside the portrait: verified ownership, behavioural history, source provenance, and institutional accountability.

B. The TRACE Framework

The empirical findings support a five-part framework for responsible visual identity:

TABLE III TRACE FRAMEWORK FOR TRUSTWORTHY PROFILE-PICTURE ECOSYSTEMS

Principle	Requirement	Operational indicators
T — Transparency	Disclose synthetic or materially altered identity imagery when it could reasonably be understood as a real person.	Visible label; machine-readable provenance; explanation available without leaving the profile.
R — Referential integrity	Establish whether the portrait corresponds to an accountable person, organization, fictional persona, or automated agent.	Identity verification separated from public-name disclosure; organization/agent ownership field; consistency checks.
A — Audience-context fit	Apply stronger standards where pictures affect employment, dating, finance, health, politics, or public authority.	Risk-tiered policy; stricter rules for high-stakes categories; contextual warnings.

Principle	Requirement	Operational indicators
C — Consent and control	Protect the likeness, demographic identity, and biometric privacy of represented or training-data subjects.	Consent records; impersonation reporting; deletion and appeal channels; data-minimization.
E — Engagement accountability	Assess networks and behaviour rather than judging credibility from faces alone.	Coordinated-behaviour detection; account-history signals; complaint response; audit of recommender amplification.

For users, TRACE implies choosing pictures that fit the intended identity context and disclosing AI generation where a reasonable viewer could mistake a synthetic person for the account holder. For platforms, it implies provenance and behavioural verification rather than face-based suspicion. For researchers, it requires separating visual trust, account credibility, and identity authenticity as distinct dependent variables. For organizations, especially recruitment and professional networking services, profile images should not be used as unexamined proxies for competence, intelligence, or personality.

VI. LIMITATIONS

The comparative coding is deliberately transparent but limited. The included studies vary in platform, sample, stimulus quality, cultural context, and outcome measurement. The ordinal scores are interpretive and were not independently double-coded. Positive-direction results do not measure equal effect sizes, and laboratory trustworthiness ratings are not equivalent to real interaction. Research remains disproportionately focused on Western participants and White faces, while contemporary diffusion models change faster than publication cycles. Future work should preregister direct comparisons of human, edited, disclosed-AI, and undisclosed-AI profile pictures across Asian, African, Latin American, and multilingual platform contexts; measure actual clicking, following, hiring, and disclosure behaviour; and evaluate provenance labels under realistic scrolling conditions.

VII. CONCLUSION

Human and AI-generated profile pictures are similar in one important respect: both are powerful, compressed identity signals. They organize first impressions and can increase social presence, attractiveness, credibility, and willingness to interact. They differ fundamentally in referential status. A human profile photograph generally represents a selectively presented person; an AI-generated portrait may establish the appearance of a person who has never existed.

The secondary-data analysis demonstrates a demand–integrity asymmetry. Synthetic faces can equal or exceed human portraits on perceived realism and trustworthiness, yet carry the highest authenticity risk. This asymmetry explains why reliance on visual intuition is no longer a defensible platform-safety strategy. Training can modestly improve detection but can also increase false suspicion toward genuine images. Trust must migrate from appearance to provenance, behavioural consistency, and accountable ownership.

The practical objective is not to eliminate synthetic profile pictures. AI avatars can protect privacy, enable creativity, and support legitimate mediated identities. The objective is to prevent synthetic visual fluency from being misread as human verification. The TRACE framework offers a route toward that distinction by aligning transparency, referential integrity, context, consent, and engagement accountability. In contemporary social networks, the most trustworthy profile ecosystem will not be the one with the most realistic faces; it will be the one in which viewers can understand what those faces represent and who remains responsible for the interactions conducted through them.

REFERENCES

[1] Walther, J. B., Van Der Heide, B., Kim, S.-Y., Westerman, D., & Tong, S. T. (2008). The role of friends' appearance and behavior on evaluations of individuals on Facebook. Human Communication Research, 34(1), 28–49. doi:10.1111/j.1468-2958.2007.00312.x

[2] Scott, G. G. (2014). More than friends: Popularity on Facebook and its role in impression formation. Journal of Computer-Mediated Communication, 19(3), 358–372. doi:10.1111/jcc4.12067

[3] Edwards, C., Stoll, B., Faculak, N., & Karman, S. (2015). Social presence on LinkedIn: Perceived credibility and interpersonal attractiveness based on user profile picture. Online Journal of Communication and Media Technologies, 5(4), 102–115. doi:10.29333/ojcmt/2528

[4] Hancock, J. T., & Toma, C. L. (2009). Putting your best face forward: The accuracy of online dating photographs. Journal of Communication, 59(2), 367–386. doi:10.1111/j.1460-2466.2009.01420.x

[5] Back, M. D., Stopfer, J. M., Vazire, S., et al. (2010). Facebook profiles reflect actual personality, not self-idealization. Psychological Science, 21(3), 372–374. doi:10.1177/0956797609360756

[6] Harris, E., & Bardey, A. C. (2019). Do Instagram profiles accurately portray personality? An investigation into idealized online self-presentation. Frontiers in Psychology, 10, 871. doi:10.3389/fpsyg.2019.00871

[7] Nightingale, S. J., & Farid, H. (2022). AI-synthesized faces are indistinguishable from real faces and more trustworthy. Proceedings of the National Academy of Sciences, 119(8), e2120481119. doi:10.1073/pnas.2120481119

[8] Miller, E. J., Steward, B. A., Witkower, Z., et al. (2023). AI hyperrealism: Why AI faces are perceived as more real than human ones. Psychological Science, 34(12), 1390–1403. doi:10.1177/09567976231207095

[9] Yang, K.-C., Singh, D., & Menczer, F. (2024). Characteristics and prevalence of fake social media profiles with AI-generated faces. Journal of Online Trust and Safety, 2(4). doi:10.54501/jots.v2i4.197

[10] Bohnert, D., & Ross, W. H. (2010). The influence of social networking web sites on the evaluation of job candidates. Cyberpsychology, Behavior, and Social Networking, 13(3), 341–347. doi:10.1089/cyber.2009.0193

[11] Davison, H. K., Maraist, C., & Bing, M. N. (2011). Friend or foe? The promise and pitfalls of using social networking sites for HR decisions. Journal of Business and Psychology, 26, 153–159. doi:10.1007/s10869-011-9215-8

[12] boyd, d. m., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. Journal of Computer-Mediated Communication, 13(1), 210–230. doi:10.1111/j.1083-6101.2007.00393.x

[13] Gonzales, A. L., & Hancock, J. T. (2011). Mirror, mirror on my Facebook wall: Effects of exposure to Facebook on self-esteem. Cyberpsychology, Behavior, and Social Networking, 14(1–2), 79–83. doi:10.1089/cyber.2009.0411

[14] 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. doi:10.1027/1864-1105.20.3.106

[15] Farid, H. (2022). Creating, using, misusing, and detecting deep fakes. Journal of Online Trust and Safety, 1(4). doi:10.54501/jots.v1i4.56

[16] Lago, F., Pasquini, C., Böhme, R., et al. (2022). More real than real: A study on human visual perception of synthetic faces. IEEE Signal Processing Magazine, 39(1), 109–116. doi:10.1109/MSP.2021.3120982

[17] Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. Journal of Communication, 60(3), 413–439. doi:10.1111/j.1460-2466.2010.01488.x

[18] Naumann, L. P., Vazire, S., Rentfrow, P. J., & Gosling, S. D. (2009). Personality judgments based on physical appearance. Personality and Social Psychology Bulletin, 35(12), 1661–1671. doi:10.1177/0146167209346309

[19] Nowak, K. L. (2013). Choosing buddy icons that look like me or represent my personality: Using buddy icons for social presence. Computers in Human Behavior, 29(4), 1456–1464. doi:10.1016/j.chb.2013.01.027

[20] Nowak, K. L., Hamilton, M. A., & Hammond, C. C. (2009). The effect of image features on judgments of homophily, credibility, and intention to use as avatars in future interactions. Media Psychology, 12(1), 50–76. doi:10.1080/15213260802669433

[21] Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., et al. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. Cognition, 127(1), 105–118. doi:10.1016/j.cognition.2012.09.001

[22] Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. Science, 308(5728), 1623–1626. doi:10.1126/science.1110589

[23] Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A., & Ortega-Garcia, J. (2020). Deepfakes and beyond: A survey of face manipulation and fake detection. Information Fusion, 64, 131–148. doi:10.1016/j.inffus.2020.06.014

[24] Toma, C. L., Hancock, J. T., & Ellison, N. B. (2008). Separating fact from fiction: An examination of deceptive self-presentation in online dating profiles. Personality and Social Psychology Bulletin, 34(8), 1023–1036. doi:10.1177/0146167208318067

[25] Van Der Heide, B., D'Angelo, J. D., & Schumaker, E. M. (2012). The effects of verbal versus photographic self-presentation on impression formation in Facebook. Journal of Communication, 62(1), 98–116. doi:10.1111/j.1460-2466.2011.01617.x

[26] Vazire, S., & Gosling, S. D. (2004). e-Perceptions: Personality impressions based on personal websites. Journal of Personality and Social Psychology, 87(1), 123–132. doi:10.1037/0022-3514.87.1.123

[27] Wang, R., Yang, F., & Haigh, M. M. (2017). Let me take a selfie: Exploring the psychological effects of posting and viewing selfies and groupies on social media. Telematics and Informatics, 34(4), 274–283. doi:10.1016/j.tele.2016.07.004

[28] Weiser, E. B. (2015). #Me: Narcissism and its facets as predictors of selfie-posting frequency. Personality and Individual Differences, 86, 477–481. doi:10.1016/j.paid.2015.07.007

[29] Tong, S. T., Van Der Heide, B., Langwell, L., & Walther, J. B. (2008). Too much of a good thing? The relationship between number of friends and interpersonal impressions on Facebook. Journal of Computer-Mediated Communication, 13(3), 531–549. doi:10.1111/j.1083-6101.2008.00409.x

[30] Wei, X., & Stillwell, D. (2017). How smart does your profile image look? Estimating intelligence from social network profile images. Proceedings of WSDM 2017, 33–40. doi:10.1145/3018661.3018663

[31] Segalin, C., Celli, F., Polonio, L., et al. (2017). What your Facebook profile picture reveals about your personality. Proceedings of the 25th ACM International Conference on Multimedia, 460–468. doi:10.1145/3123266.3123331

[32] Appel, M., et al. (2023). Swipe right? Using beauty filters in male Tinder profiles increases physical attractiveness but decreases perceived trustworthiness. Computers in Human Behavior, 149, 107871. doi:10.1016/j.chb.2023.107871

[33] McGuire, A. A., Nightingale, S., Taylor, P., Farid, H., & Bohacek, M. (2026). AI-generated faces are becoming more trustworthy. Journal of Vision, 26(7), 3. doi:10.1167/jov.26.7.3

[34] Rehman, R. S., Meier, E., Ibsen, M., Rathgeb, C., Nichols, R., & Busch, C. (2025). Training humans for synthetic face image detection. Frontiers in Artificial Intelligence, 8, 1568267. doi:10.3389/frai.2025.1568267

[35] Porcile, G. J. A., Gindi, J., Mundra, S., Verbus, J. R., & Farid, H. (2024). Finding AI-generated faces in the wild. Proceedings of the IEEE/CVF CVPR Workshops, 4297–4305. doi:10.1109/CVPRW63382.2024.00433

[36] Mundra, S., Porcile, G. J. A., Marvaniya, S., Verbus, J. R., & Farid, H. (2023). Exposing GAN-generated profile photos from compact embeddings. Proceedings of the IEEE/CVF CVPR Workshops, 884–892. doi:10.1109/CVPRW59228.2023.00095

[37] Ricker, J., et al. (2024). AI-generated faces in the real world: A large-scale case study of Twitter profile images. Proceedings of the ACM Workshop on Multimedia AI against Disinformation. doi:10.1145/3678890.3678922

[38] Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. Psychological Science, 17(7), 592–598. doi:10.1111/j.1467-9280.2006.01750.x

[39] Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. Proceedings

of the National Academy of Sciences, 105(32), 11087–11092. doi:10.1073/pnas.0805664105