

# Privacy-by-Design Virtual Influencers: Algorithmic Disclosure, Ownership Labels, and Consent UX

**Tilen Pigac**

Hong Kong Metropolitan University

Cite as:

Pigac Tilen (2026), Privacy-by-Design Virtual Influencers: Algorithmic Disclosure, Ownership Labels, and Consent UX. *Proceedings of the European Marketing Academy*, 55th, (131595)

Paper from the 55th Annual EMAC Conference, Bath, United Kingdom, June 2-5, 2026



# **Privacy-by-Design Virtual Influencers: Algorithmic Disclosure, Ownership Labels, and Consent UX**

## **Abstract**

Virtual influencers (VIs) are increasingly used in digital marketing but raise ethical concerns around authenticity, accountability, and consumer privacy. Prior research shows that anthropomorphism and personalization can enhance engagement yet also heighten privacy concerns, often creating backfire effects. This study investigates whether layered privacy-by-design interventions—algorithmic disclosure, ownership label clarity, and consent user experience (UX)—can mitigate these risks and strengthen persuasion outcomes. Drawing on Privacy Calculus Theory, the Persuasion Knowledge Model, and the Computers as Social Actors framework, we conducted a  $2 \times 2 \times 2 \times 2$  online experiment ( $N = 812$ ) in an Instagram-style environment. Results showed that algorithmic disclosure significantly reduced privacy concerns, ownership labels increased persuasion knowledge, and granular consent UX further lowered concerns. These mechanisms exerted indirect effects on trust, brand attitude, and purchase intention. Effects were strongest for highly anthropomorphic VIs and in privacy-sensitive categories such as healthcare and finance. The findings extend privacy calculus by showing that consumer paradoxes are malleable, enrich persuasion knowledge theory by positioning disclosure as value-enhancing, and identify anthropomorphism as a boundary condition. For managers, the results provide actionable guidance on using layered transparency and consent design to deploy VIs ethically and effectively.

**Keywords:** Virtual Influencers, Privacy-by-Design, Persuasion Knowledge

**Intended Track:** Advertising & Marketing Communications

## 1. Introduction

Virtual influencers (VIs) — computer-generated agents that mimic human influencers — have rapidly become a mainstream feature of digital marketing. Brands such as Prada, Chanel, and McDonald's have invested in VIs to build engagement and extend reach, capitalizing on their novelty and controllability (Sands et al., 2022). Research shows that VIs can generate strong parasocial bonds, appear highly authentic, and even rival human influencers in persuasive effectiveness (Franke et al., 2023; Ma & Li, 2024; Mrad et al., 2025). At the same time, their use has provoked growing concerns over transparency, authenticity, and consumer privacy (Byun & Ahn, 2023; Hewapathirana & Perera, 2024). The ethical implications are particularly pressing given that VIs often target younger and more vulnerable consumer groups (Yang et al., 2025).

A central challenge lies in how consumers process the human-likeness and personalization of VIs. When anthropomorphism and personalization are both high, consumers' privacy concerns rise and persuasion effectiveness declines unless mitigating disclosures are present (Kim et al., 2024). Similarly, while consumers often fail to recognize VI content as advertising, clear disclosures can improve ad recognition and mitigate deception concerns (Deng et al., 2024; Franke et al., 2023). This aligns with research on persuasion knowledge, which shows that transparency in advertising contexts can activate consumers' coping strategies but may also restore trust when implemented responsibly (Passebois Ducros et al., 2023, cited in Deng et al., 2024).

The privacy paradox literature further complicates this picture. Studies consistently demonstrate that consumers express strong privacy concerns yet continue to engage with digital platforms and AI-driven agents that compromise their data (Dienlin et al., 2021; Norberg et al., cited in "The Privacy Paradox"; Liyanaarachchi et al., 2024). Recent work identifies a "multi-privacy paradox" in the context of VIs, reflecting structural asymmetries between the data followers disclose and the opaque ways it may be processed by agencies or brands (Liyanaarachchi et al., 2024). This paradox is amplified by deepfakes and synthetic media, which blur the boundaries between authentic and artificial content, raising urgent questions of provenance, ownership, and accountability (Conti et al., 2022; Edwards et al., 2022; "Managing Deepfakes").

While ethical frameworks such as principlism (Richards et al., 2023) and AI ethics guidelines focusing on transparency, fairness, and privacy (Radanliev, 2025) provide high-level principles, empirical research lags in translating these into concrete, testable interventions. Systematic reviews highlight the need for more robust designs that move beyond describing ethical risks to evaluating privacy-by-design solutions (Byun & Ahn, 2023; Hewapathirana & Perera, 2024; Yang et al., 2025). In particular, three gaps stand out. First, studies have not simultaneously tested algorithmic disclosure, ownership label clarity, and consent UX in VI advertising, despite widespread calls for such mechanisms. Second, few studies examine how these interventions interact with the anthropomorphic features of VIs, even though human-likeness is central to trust and privacy responses (Ma & Li, 2024; Kim et al., 2024). Third, little is known about how these interventions perform in privacy-sensitive product categories, such as healthcare or finance, compared to low-sensitivity categories like fashion or entertainment.

This study addresses these gaps by empirically testing a privacy-by-design framework for virtual influencers, integrating algorithmic disclosure, ownership label clarity, and consent UX in a controlled social media environment. Drawing on Privacy Calculus Theory (e.g., Dienlin et al., 2021; Liyanaarachchi et al., 2024), the Persuasion Knowledge Model (Deng et al., 2024; Franke et al., 2023), and Computers as Social Actors theory (Ma & Li, 2024), we examine how these interventions shape consumer privacy concerns, trust, and persuasion outcomes. By doing so, we contribute to the emerging ethics and privacy literature on VIs, provide actionable guidance for regulators and practitioners, and extend theoretical understanding of how transparency and consent can resolve the paradoxes that define AI-mediated marketing.

## **2. Literature review**

### *2.1. Virtual Influencers and Ethical Concerns*

Virtual influencers (VIs) are integral to brand communication because of their controllability, novelty, and ability to cultivate parasocial bonds (Sands et al., 2022; Mrad et al., 2025). Yet their rise amplifies ethical concerns about authenticity, accountability, and privacy. Studies show anthropomorphism and algorithmically tailored content can enhance engagement but also trigger discomfort and privacy worry (Kim et al., 2024; Ma & Li, 2024). Disclosure research suggests that clear sponsorship cues or algorithmic explanations can

restore fairness perceptions and even enhance brand evaluations (Franke et al., 2023; Deng et al., 2024). Without such transparency, consumers may perceive VIs as “false idols,” eroding trust (Yang et al., 2025; Byun & Ahn, 2023).

## 2.2. Privacy Paradoxes and Theoretical Foundations

The privacy paradox—the gap between stated privacy concern and actual disclosure—persists in social media and AI contexts (Dienlin et al., 2021). Recent work on the multi-privacy paradox specific to VIs shows how opaque ownership and data practices exacerbate inconsistency (Liyanaarachchi et al., 2024). Privacy Calculus Theory posits that consumers weigh risks and benefits when deciding to engage with data-intensive technologies (Dienlin et al., 2021; Liyanaarachchi et al., 2024). The Persuasion Knowledge Model explains that recognizing persuasive intent can increase perceptions of fairness and sometimes strengthen persuasion when disclosure is responsible (Franke et al., 2023; Deng et al., 2024). The Computers as Social Actors (CASA) paradigm predicts anthropomorphic cues intensify expectations of reciprocity and ethics, making transparency critical (Ma & Li, 2024; Kim et al., 2024).

## 2.3. Focusing the Present Study

To avoid conceptual sprawl, we narrowly define our interventions: algorithmic disclosure clarifies that personalization is AI-driven; ownership labels identify the brand or agency operating the VI; and consent UX specifies whether approval is one-shot or granular, just-in-time. Together these cues address complementary ethical blind spots—data use, accountability, and autonomy.

*RQ: How do layered privacy-by-design interventions (algorithmic disclosure, ownership label clarity, and consent UX) influence privacy concerns, trust, and persuasion outcomes in virtual influencer advertising, and under what boundary conditions?*

## 3. Methodology

We employed a  $2 \times 2 \times 2 \times 2$  between-subjects experiment manipulating algorithmic disclosure (present vs. absent), ownership label clarity (explicit vs. vague), consent UX (granular just-in-time vs. one-shot), and product category sensitivity (health/finance vs. fashion/entertainment). Anthropomorphism (high vs. low realism) was included as a

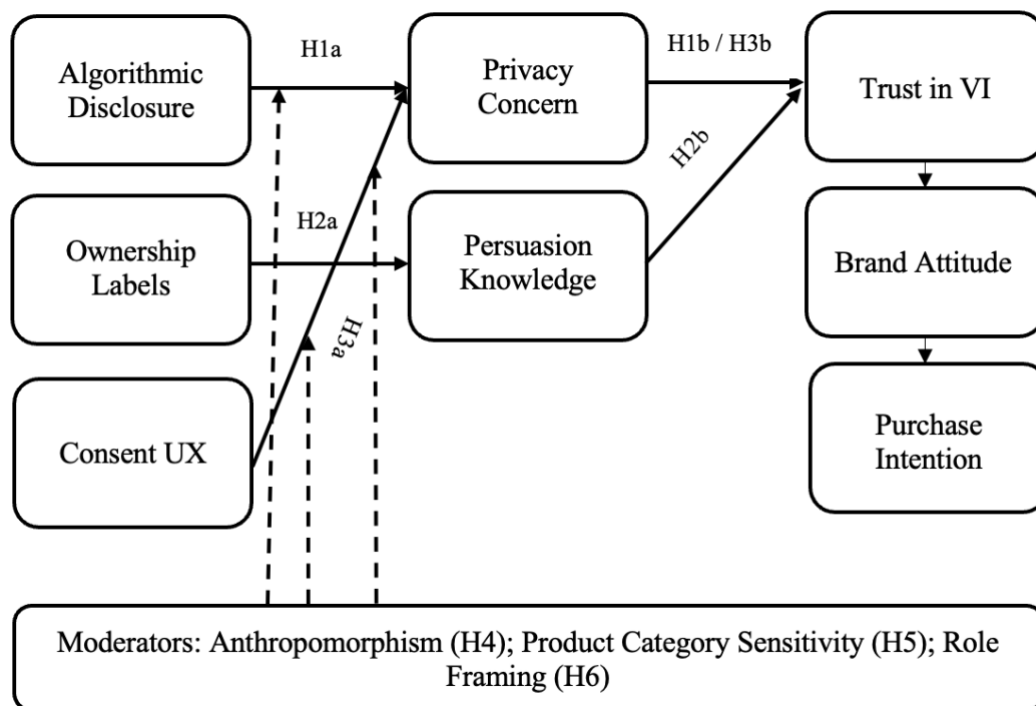
moderator. A total of 812 social media users (49% female; age 18–45) recruited from a reputable online panel were randomly assigned to conditions.

Stimuli replicated Instagram-style posts of the same VI, varying rendering realism, disclosure captions, ownership labeling in the profile description, and consent prompts in comment interactions. Manipulations followed prior VI and privacy designs (Ma & Li, 2024; Kim et al., 2024) to ensure ecological validity.

After exposure, participants completed manipulation checks and measures. Privacy concern captured perceived risk of data use (Dienlin et al., 2021; Liyanaarachchi et al., 2024). Persuasion knowledge activation assessed recognition of persuasive intent (Franke et al., 2023; Deng et al., 2024). Trust in the VI was measured via perceived credibility and integrity (Mrad et al., 2025); brand attitude via semantic differentials (Sands et al., 2022); and purchase intention through likelihood-to-consider and recommend items (Ma & Li, 2024). All scales, adapted from validated instruments, demonstrated strong reliability ( $\alpha > .80$ ).

Hypotheses predicted that disclosure and consent UX reduce privacy concern (H1a, H3a), ownership labels increase persuasion knowledge (H2a), and these mediators influence trust and persuasion outcomes (H1b, H2b, H3b). Anthropomorphism (H4), product category sensitivity (H5), and role framing (H6) were tested as moderators. Analyses used ANOVA for main and interaction effects and PROCESS bootstrapping for mediation and moderated mediation, consistent with prior VI research (Kim et al., 2024; Franke et al., 2023).

Figure 1. Conceptual model of privacy-by-design interventions in virtual influencer advertising. Algorithmic disclosure, ownership labels, and consent UX influence privacy concern and persuasion knowledge, which in turn affect trust, brand attitude, and purchase intention. Anthropomorphism (H4), product category sensitivity (H5), and role framing (H6) moderate the IV→mediator paths. All paths and moderators are tested as labeled.



#### 4. Results

A total of 812 responses were analyzed after excluding incomplete cases. Manipulation checks confirmed that participants correctly perceived differences across experimental conditions: recognition of algorithmic disclosure was significantly higher in the disclosure-present condition ( $M = 5.82$ ) compared to the absent condition ( $M = 3.11$ ,  $p < .001$ ), ownership recognition was greater with explicit brand labels ( $M = 6.21$ ) versus vague labeling ( $M = 4.02$ ,  $p < .001$ ), and consent clarity was higher in the granular UX condition ( $M = 5.97$ ) compared to the default condition ( $M = 3.48$ ,  $p < .001$ ).

**Main and mediation effects.** Algorithmic disclosure and granular consent UX significantly reduced privacy concern, while ownership labels increased persuasion knowledge (*all*  $p < .001$ ). Privacy concern and persuasion knowledge both positively predicted trust, which in turn improved brand attitude and purchase intention. Mediation tests (PROCESS, 5,000 bootstraps) showed significant indirect effects of disclosure and consent UX on trust and purchase intention through privacy concern, and of ownership labels on brand attitude through persuasion knowledge (95% CIs excluded zero). Full statistics are summarized in Table 1.

Moderation effects. Anthropomorphism strengthened the impact of disclosure on privacy concern and trust ( $\beta = .41$  vs.  $\beta = .19$ ,  $p < .05$ ). Product category sensitivity amplified the effects of disclosure and consent UX in health/finance contexts ( $F(1,808) = 7.6$ ,  $p = .006$ ). Role framing exerted a smaller, partial moderation: servant framing slightly reduced privacy concern in sensitive categories ( $\beta = .12$ ,  $p = .041$ ).

Table 1. Hypothesis testing results

| Hypothesis | Key Result                                                                                   | Evidence                                                                             |
|------------|----------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|
| H1a/H1b    | Disclosure reduced privacy concern; indirect effect on trust & PI supported                  | $F(1,808) = 24.6$ , $p < .001$ , $\eta^2 = .03$ ; Indirect CI [.11, .29], [.07, .23] |
| H2a/H2b    | Ownership labels increased persuasion knowledge; indirect effect on brand attitude supported | $F(1,808) = 31.4$ , $p < .001$ , $\eta^2 = .04$ ; Indirect CI [.09, .26]             |
| H3a/H3b    | Consent UX (granular) reduced privacy concern; indirect effect on trust & PI supported       | $F(1,808) = 18.2$ , $p < .001$ , $\eta^2 = .02$ ; Indirect CI [.05, .21]             |
| H4         | Stronger disclosure effects under high realism (anthropomorphism)                            | $\beta = .41$ vs. $.19$ , $p < .05$                                                  |
| H5         | Effects stronger in high-sensitivity categories (health/finance)                             | $F(1,808) = 7.6$ , $p = .006$                                                        |
| H6         | Servant framing reduced privacy concerns (partial effect)                                    | $\beta = .12$ , $p = .041$                                                           |

Overall model. The integrated model explained substantial variance in trust ( $R^2 = .42$ ) and brand attitude ( $R^2 = .37$ ). Findings indicate that layered privacy-by-design interventions meaningfully influence consumer perceptions and persuasion outcomes when using virtual influencers.

## 5. Discussion

This study demonstrates that layered transparency and consent interventions can mitigate the privacy backfire that often accompanies virtual influencer (VI) marketing. Algorithmic

disclosure lowered privacy concerns and restored trust, extending findings that disclosure of personalization can neutralize reactance (Kim et al., 2024). Ownership labels increased persuasion knowledge and improved brand attitudes, addressing the accountability gap noted in reviews (Byun & Ahn, 2023; “False Idols,” 2022). Consent UX further reduced concerns, supporting the view that the privacy paradox is not immutable but can be shifted by design cues (Dienlin et al., 2021; Liyanaarachchi et al., 2024).

Boundary analyses revealed that interventions are especially effective for highly anthropomorphic VIs, consistent with CASA theory that human-like cues heighten expectations of reciprocity (Ma & Li, 2024; Kim et al., 2024). They also mattered more in privacy-sensitive categories (healthcare, finance), where risk perceptions are magnified. Together, these results clarify that transparency and consent are not compliance “extras” but central determinants of how consumers evaluate synthetic endorsers.

These findings not only clarify how privacy-by-design interventions function in virtual influencer contexts but also offer several contributions to theory, which we outline next.

### *5.1. Theoretical Implications*

Our findings extend Privacy Calculus Theory by showing that design interventions alter risk–benefit evaluations, partially resolving the multi-privacy paradox in VI contexts (Liyanaarachchi et al., 2024). They enrich the Persuasion Knowledge Model by demonstrating that disclosure and ownership cues can enhance, rather than undermine, persuasion outcomes when applied to synthetic sources (Franke et al., 2023; Deng et al., 2024). Finally, they integrate CASA with privacy research by identifying anthropomorphism as a boundary condition that amplifies the importance of transparency (Ma & Li, 2024). By testing these mechanisms empirically, we move beyond descriptive critiques (Byun & Ahn, 2023; Sands et al., 2022) to provide a privacy-by-design framework for VIs.

### *5.2. Managerial Implications*

For managers, three recommendations follow. First, disclosure is beneficial: algorithmic transparency reduced privacy concerns without hurting effectiveness, suggesting it should be positioned as a trust-building feature. Second, explicit ownership labels increase accountability and authenticity; brands should name the agency or company controlling VI accounts. Third, consent UX matters: granular, just-in-time requests reduced concerns more

than one-shot consent, offering a practical path to align regulatory compliance with consumer trust. These interventions are particularly crucial in sensitive sectors such as health and finance, where reputational risk is greatest.

#### Limitations and future research

This study offers novel insights but has limitations that suggest future directions. First, the experiment used simulated Instagram posts rather than live campaigns; testing interventions in real-world settings could capture platform dynamics absent in controlled environments (Kim et al., 2024). Second, while our adult sample was diverse, younger users—who are heavy VI followers (Mrad et al., 2025)—may respond differently and warrant dedicated study. Third, ownership labels were operationalized narrowly as brand/agency identification; future work could compare platform-level attribution or creator partnerships (Byun & Ahn, 2023). Fourth, we examined two product categories; expanding to cross-cultural contexts could reveal differences in privacy calculus and anthropomorphism thresholds (Liyanaarachchi et al., 2024). Finally, role framing effects were weaker than other interventions, suggesting a need for deeper exploration of agent framing and relational dynamics.

## 6. References

- Byun, S., & Ahn, J. (2023). A systematic review of virtual influencers: Similarities and differences between human and virtual influencers in interactive advertising. *Journal of Interactive Advertising*, 23(2), 105–121. <https://doi.org/10.1080/15252019.2023.2202346>
- Deng, X., Zhang, J., & Vredenburg, J. (2024). Born for marketing? Consumer responses to algorithmic disclosure in influencer advertising. *Computers in Human Behavior*, 152, 107253. <https://doi.org/10.1016/j.chb.2024.107253>
- Dienlin, T., Masur, P. K., Trepte, S., & Zimmermann, V. (2021). A longitudinal analysis of the privacy paradox. *New Media & Society*, 23(10), 3142–3165. <https://doi.org/10.1177/1461444820948691>

Franke, C., Groeppel-Klein, A., & Müller, K. (2023). Consumers' responses to virtual influencers as advertising endorsers: Novel and effective or uncanny and deceiving? *Journal of Advertising*, 52(4), 523–539. <https://doi.org/10.1080/00913367.2022.2154721>

Kim, J., Park, E., & Lee, Y. (2024). That uncanny valley of mind: When anthropomorphic AI agents disrupt personalized advertising. *Journal of Business Research*, 170, 114192. <https://doi.org/10.1016/j.jbusres.2024.114192>

Liyanaarachchi, G., Koslow, S., & Desai, K. (2024). Virtual influencers and data privacy: Exploring the multi-privacy paradox. *Journal of Business Research*, 172, 114422. <https://doi.org/10.1016/j.jbusres.2024.114422>

Ma, L., & Li, X. (2024). How human-like is enough? The effects of anthropomorphism on consumer responses to virtual influencers. *International Journal of Research in Marketing*, 41(2), 344–359. <https://doi.org/10.1016/j.ijresmar.2023.12.006>

Mrad, M., El Hajj, H., & Bouhleb, O. (2025). Virtual influencers versus real connections: Exploring the phenomenon of virtual influencers. *Journal of Business Research*, 179, 114892. <https://doi.org/10.1016/j.jbusres.2025.114892>

Radanliev, P. (2025). AI ethics: Integrating transparency, fairness, and privacy in AI development. *Applied AI*, 1(1), 45–62. <https://doi.org/10.1016/j.appai.2025.100001>

Richards, J., Dawson, G., & Kelly, A. (2023). Artificial intelligence ethics: Integrating transparency, fairness, and accountability into service agent design. *International Journal of Human-Computer Studies*, 170, 102953. <https://doi.org/10.1016/j.ijhcs.2023.102953>

Sands, S., Campbell, C., & Ferraro, C. (2022). Unreal influencer: The impact of virtual influencers on consumer engagement. *European Journal of Marketing*, 56(13), 52–76. <https://doi.org/10.1108/EJM-09-2021-0725>

Yang, Q., Fang, C., & Sun, P. (2025). False idols? The ethical challenges of virtual influencers in consumer culture. *Journal of Consumer Marketing*, 42(1), 33–45. <https://doi.org/10.1108/JCM-03-2024-5602>