



To cite this article: Ms. Manilakshmi. S and Dr. C. Vethirajan (2026). AI AD-TRANSPARENCY AND MANIPULATIVE PERSONALIZATION IN E-COMMERCE RECOMMENDER SYSTEMS: A CONCEPTUAL FRAMEWORK, International Journal of Research in Commerce and Management Studies (IJRCMS) 8 (5): 1070-1079 Article No. 921 Sub Id 1496

AI AD-TRANSPARENCY AND MANIPULATIVE PERSONALIZATION IN E-COMMERCE RECOMMENDER SYSTEMS: A CONCEPTUAL FRAMEWORK

Ms. Manilakshmi. S¹ and Dr. C. Vethirajan²

¹Research Scholar (Full – Time)
Alagappa Institute of Management
Alagappa University
Karaikudi – 630 004.
Email id: manilakshmi19mba@gmail.com

²Senior Professor and Head
Department of Corporate Secretaryship
Alagappa University
Karaikudi – 630 004.
Email Id: vethirajanc@alagappauniversity.ac.in

DOI: <https://doi.org/10.38193/IJRCMS.2026.8563>

ABSTRACT

As recommendation systems in online retail increasingly embed paid placements into personalized product discovery, there is a risk that personalization becomes manipulation. This conceptual paper looks at how AI-powered ad-transparency tools can reduce manipulative personalization in e-commerce, without destroying the value of personalization. The paper integrates persuasion knowledge model, reactance theory, trust theory and stimulus-organism-response paradigm with the literature on dark patterns, explainable AI and recent regulation such as the EU Digital Services Act, the EU AI Act and India's Guidelines for Prevention and Regulation of Dark Patterns, 2023. It defines manipulative personalization through three criteria (covertness, vulnerability exploitation and deliberation bypassing) and proposes a four-dimension typology of AI ad-transparency tools (disclosure, explanatory, control and audit transparency). It constructs a framework with three propositions: AI ad-transparency tools activate consumers' persuasion knowledge (P1); when recognized, manipulative tactics increase perceived manipulation, stronger when they are personalized than generic (P2); and control and explanatory transparency increase perceived autonomy and algorithmic trust (P3). The paper repositions transparency from a disclosure label to a multi-layered AI capability that both exposes manipulation and restores consumer agencies, providing a transparency-by-design approach for retailers, substantive criteria for regulators, and testable propositions for future empirical research.

KEYWORDS: Ad transparency; Recommendation systems; Manipulative personalization; Dark patterns; Explainable AI; Persuasion knowledge; Consumer autonomy; Algorithmic trust; E-

commerce

INTRODUCTION

Personalized recommendation is the organizing logic of online retail. Home feeds, search results, “customers also bought” carousels and checkout add-ons are ranked by models trained on each shopper’s behaviour. Recommendation agents reduce search costs and help consumers handle large assortments (Xiao & Benbasat, 2007), but they are now also monetized as advertising inventory: sponsored products are blended into organic results, and the targeting that decides which offer is largely invisible to the shopper.

This convergence creates the risk of *manipulative personalization*: personalization that works by hidden means, exploits predictable weaknesses in decision-making, and steers choices consumers would not endorse on reflection (Susser et al., 2019). Firms that profile consumers on a scale can find and exploit individual vulnerabilities (Calo, 2014), and manipulative interface designs, or dark patterns, are widespread on shopping websites (Mathur et al., 2019) and measurably change consumer choices (Luguri & Strahilevitz, 2021).

Regulators have responded. The EU Digital Services Act (DSA) requires platforms to label advertisements and disclose targeting parameters (Article 26), explain recommender-system parameters (Article 27), avoid manipulative interface design (Article 25) and, for very large platforms, offer a non-profiled recommender option (Article 38) (European Union, 2022). The EU AI Act prohibits manipulative or vulnerability-exploiting AI systems (Article 5) (European Union, 2024), and India’s Central Consumer Protection Authority (2023) has specified thirteen dark patterns, including disguised advertisements, false urgency and drip pricing.

Transparency is central to these responses, yet the evidence on its effects is mixed. Revealing targeting can reduce ad effectiveness when it exposes unacceptable information flows but increase it when flows are acceptable (Kim et al., 2019); platform explanations are often incomplete (Andreou et al., 2018); and more information does not always produce more trust (Kizilcec, 2016). This article therefore asks: (RQ1) what distinguishes manipulative from legitimate personalization in e-commerce recommendation; (RQ2) what forms AI ad-transparency tools can take; and (RQ3) through which psychological mechanisms (persuasion knowledge, perceived manipulation, and perceived autonomy and algorithmic trust) counter manipulative personalization. The answers are summarized in three propositions.

REVIEW OF LITERATURE

From persuasion to manipulative personalization

While personalization may increase relevance, its consequences depend on how consumers perceive the use of their data (Boerman et al., 2017). Covert data collection lowers click-through (Aguirre et

al., 2015), trust determines whether personalization helps or hurts (Bleier & Eisenbeiss, 2015), and perceived control increases responsiveness to personalized advertising (Tucker, 2014). Building on Susser et al. (2019) and Sunstein (2016), **manipulative personalization** is defined here as the use of personal data and algorithmic targeting to influence choice through means that are (a) *covert*, (b) *vulnerability-exploiting* and (c) *deliberation-bypassing*. In e-commerce it takes three forms: covert sponsorship (paid placements disguised as neutral recommendations), vulnerability targeting (offers timed to inferred states such as fatigue or impulsivity) and algorithmic dark patterns (personalized false scarcity, urgency or pre-selected add-ons) (Gray et al., 2018; Mathur et al., 2019). The same tactics can also be deployed generically, identically for all shoppers; the distinction between personalized and generic tactics is central to P2.

A typology of AI ad-transparency tools

An **AI ad-transparency tool** is any system that uses algorithmic methods to make the commercial nature, logic, data basis or manipulative potential of a recommendation visible, understandable or controllable. Table 1 distinguishes four dimensions. Explanatory tools can draw on explainable-AI methods such as feature attribution (Lundberg & Lee, 2017) and counterfactual explanations (Wachter et al., 2018), and explanation facilities are known to raise trust in recommendation agents (Wang & Benbasat, 2007). Audit tools extend automated dark-pattern detection (Mathur et al., 2019) and internal algorithmic auditing (Raji et al., 2020).

Table 1. Typology of AI ad-transparency tools

Dimension	Makes visible or controllable	Illustrative AI-enabled tools	Main mechanism
Disclosure	That a recommendation is paid, and by whom	Consistent sponsored labels; detection of undisclosed paid content	Persuasion knowledge (P1)
Explanatory	Why an item was recommended or a user targeted	Per-item “Why am I seeing this?”; feature-attribution and counterfactual explanations	Persuasion knowledge (P1); autonomy and algorithmic trust (P3)
Control	How one is profiled and ranked	Editable interests; sensitive-category exclusion; non-profiled feed	Perceived autonomy and algorithmic trust (P3)

Dimension	Makes visible or controllable	Illustrative AI-enabled tools	Main mechanism
Audit	Whether the system manipulates	Dark-pattern classifiers; manipulation-risk scores; ad repositories; independent audits	Recognition of manipulation (P1, P2); reduced prevalence

Source: Author's own work.

THEORETICAL FOUNDATIONS

The *persuasion knowledge model* holds that consumers who recognize an influence attempt re-evaluate it and adjust their response (Friestad & Wright, 1994); manipulative personalization depends on not activating this knowledge, so transparency tools act as persuasion-knowledge activators (P1). *Reactance theory* explains why manipulation, once discovered, provokes a negative appraisal of its source, and why this response is sharper when the influence is directed at the individual (Brehm, 1966; White et al., 2008) (P2). *Trust theory* (Mayer et al., 1995) and self-determination theory's concept of autonomy (Deci & Ryan, 2000) describe the states that manipulation erodes and that control and explanation can rebuild (P3). The *S-O-R paradigm* (Mehrabian & Russell, 1974) integrates these perspectives: transparency tools and manipulative tactics are stimuli; persuasion knowledge, perceived manipulation, and perceived autonomy and algorithmic trust are organismic states; and shopping outcomes are responses (Figure 1). The propositions concern the stimulus–organism links, which is where the theories make their clearest predictions.

CONCEPTUAL FRAMEWORK AND PROPOSITIONS

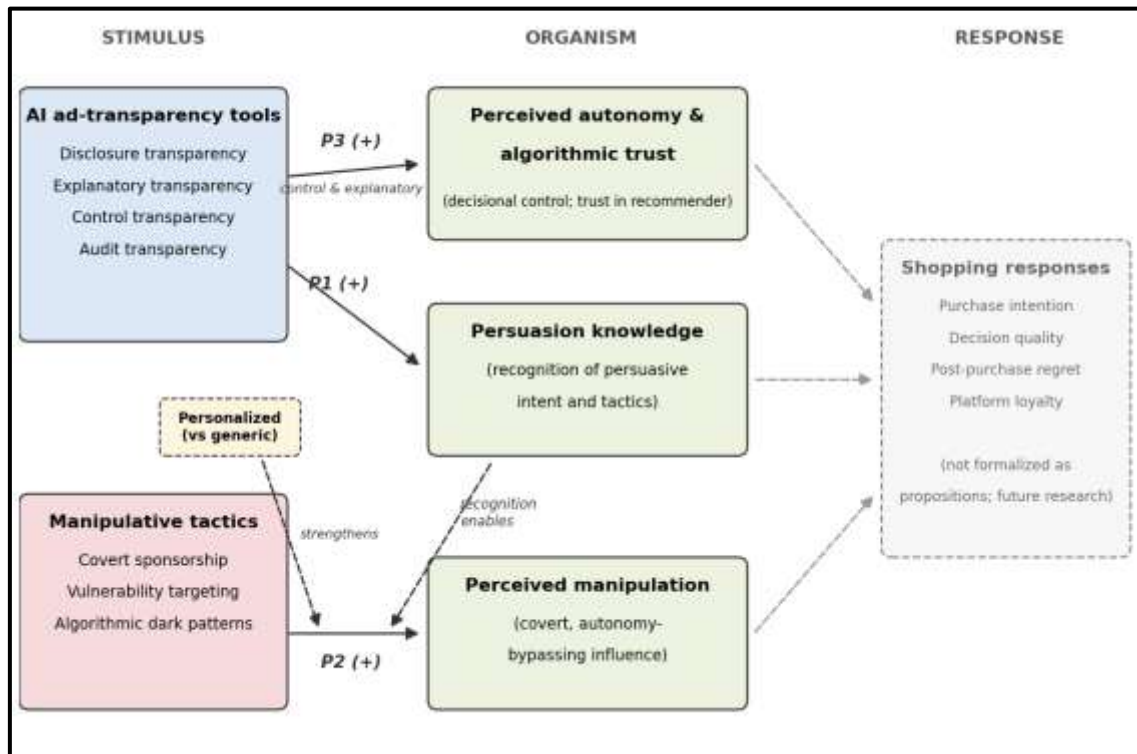


Figure 1. Conceptual framework and propositions (S-O-R structure).

(Source: Author's own work.)

AI ad-transparency tools and persuasion knowledge

Manipulative personalization is effective largely because it is not perceived as an influence attempt: sponsored items are mixed in with organic rankings, stealing credibility from neutral recommendations, and the targeting logic behind them is opaque (Susser et al., 2019). The persuasion knowledge model suggests that when cues signal persuasive intent, this activates consumers' knowledge of the motives and tactics of persuasion (Friestad & Wright, 1994). These cues are all aspects of AI ad-transparency. Disclosure reveals a paid placement and by whom; explanation reveals why a shopper was targeted (Eslami et al., 2018; Kim et al., 2019). Audit tools can identify personalized dark patterns that would otherwise remain hidden (Mathur et al., 2019). Transparency thus converts hidden commercial influence into a recognizable attempt at persuasion.

P1: *The availability of AI ad-transparency tools increases consumers' activation of persuasion knowledge.*

Recognized manipulation and perceived manipulation

Recognition is needed. But not enough. When persuasion knowledge is activated, consumers evaluate a tactic. When the tactic is covert, exploits a vulnerability, or bypasses deliberation, that evaluation results in a judgment of manipulation, rather than an acceptance of legitimate persuasion (Sunstein, 2016). This judgment should be increased by personalization. A sitewide countdown banner is a generic tactic that can be seen as run-of-the-mill marketing, while a personalized one, such as a scarcity message or add-on timed to a shopper's inferred impulsivity, suggests the firm has used personal data to exploit that individual's very own weaknesses (Calo, 2014). Such tactics are more intrusive and more directly threaten personal freedom, and reactance theory predicts that this will elicit a stronger negative response (Brehm, 1966; White et al., 2008). Similarly, covert use of personal data depresses responses toward personalized advertising (Aguirre et al., 2015).

P2: Once recognized, manipulative tactics increase consumers' perceived manipulation, and this effect is stronger when the tactics are personalized than when they are generic.

Control and explanatory transparency, autonomy and algorithmic trust

Disclosure and audit mainly show manipulation. Control and explanatory transparency, however, also give something back to the consumer. Control tools, such as editable interest profiles, exclusion of sensitive categories and a non-profiled feed, restore decisional control, satisfying the basic need for autonomy (Deci & Ryan, 2000) and increasing responsiveness to personalized advertising (Tucker, 2014). Explanatory tools help to understand the logic of a recommendation, indicating the integrity and benevolence of the platform (Mayer et al., 1995) and increasing trust in recommendation agents (Wang & Benbasat, 2007; Shin, 2021). Together they equip consumers to see and shape how they are profiled, so that personalization feels like a service, not a stealthy influence.

P3: Control and explanatory transparency increase consumers' perceived autonomy and algorithmic trust.

DISCUSSION AND IMPLICATIONS**Theoretical contributions**

The article contributes in three ways. It first distinguishes between manipulation and persuasion along three explicit criteria, and between personalized and generic manipulative tactics. Second, it helps explain mixed findings on transparency by showing its four dimensions work through different mechanisms: disclosure, explanation and audit make influence recognizable by activating persuasion knowledge (P1), while control and explanation restore perceived autonomy and algorithmic trust (P3). Transparency therefore has an exposing function and an empowering function, and studies that treat it as a single label conflate the two. Third, we extend the persuasion knowledge model to algorithmic

retail, proposing that personalization intensifies the manipulation judgments that recognition makes possible (P2).

Managerial implications

Retailers should view transparency as a trust asset embedded in the recommendation pipeline, rather than as a label slapped on after the fact. P1 and P2 jointly suggest that the more pervasive transparency is, the more of a liability personalized dark patterns are: they are more likely to be identified, and once identified, are found to be more egregious than generic tactics. P3 shows the upside, because control and explanation are the things that create autonomy and trust. A practical transparency-by-design architecture consists of five layers; data governance (excluding sensitive inferences from targeting), manipulation audit (scoring and suppressing high-risk personalized recommendations before display), disclosure (consistent, salient sponsored labels), explanation (short, recommendation-specific and expandable, with audited fidelity, since post-hoc explanations can be unfaithful; Rudin, 2019) and consumer control (editable profiles and a non-profiled feed with visible effect). Furthermore, privacy practices have implications for customer and firm performance over time (Martin et al., 2017), so performance indicators should include returns, repeat purchase and satisfaction, and not just click-through on sponsored placements.

Policy implications

Regulators should judge transparency by substance rather than form. A label that is not noticed, understood, or linked to real control satisfies the rule without activating persuasion knowledge (P1) or restoring autonomy (P3), and visibility alone does not produce accountability (Ananny & Crawford, 2018). P2 highlights that personalized dark patterns should be a priority for enforcement, as they exploit individual vulnerabilities and erode trust to a greater extent than generic ones. Comprehension testing and independent audits of transparency features, along with protective defaults, are especially important for less digitally literate consumers, who are most vulnerable to dark patterns (Luguri & Strahilevitz, 2021) and make up a large proportion of new online shoppers in emerging markets like India.

LIMITATIONS AND FUTURE RESEARCH

The propositions are untested as a conceptual study. The framework is limited to the psychological states that transparency changes and does not formalize how these psychological states affect shopping outcomes or under what conditions transparency may backfire, such as unacceptable information flows, explanation overload or low digital proficiency (Kim et al., 2019; Kizilcec, 2016). It also focuses on individual consumers rather than platform competition, assumes good-faith implementation and draws mainly on Western and Indian regulation. Future research can (1) test P1 and P3 with experiments manipulating each transparency dimension in realistic shopping interfaces,

(2) test P2 with factorial designs crossing personalized and generic tactics with the presence or absence of recognition cues, (3) link the three organismic states to purchase intention, decision quality, post-purchase regret and loyalty through survey-based structural equation modeling, and (4) develop and validate AI classifiers to detect personalized dark patterns.

CONCLUSION

As advertising moves into the storefront of online retail, personalization can easily become manipulation. AI ad-transparency tools can counter this, but only when they are understood as more than labels. This paper proposes that transparency activates consumers' persuasion knowledge (P1), that manipulation, once recognized, is judged more harshly when it is personalized than when it is generic (P2), and that control and explanatory transparency restore perceived autonomy and algorithmic trust (P3). Built into the recommendation pipeline, transparency can turn a regulatory obligation into a lasting source of consumer trust.

REFERENCES

1. Aguirre, E., Mahr, D., Grewal, D., de Ruyter, K., & Wetzels, M. (2015). Unraveling the personalization paradox: The effect of information collection and trust-building strategies on online advertisement effectiveness. *Journal of Retailing*, 91(1), 34–49.
<https://doi.org/10.1016/j.jretai.2014.09.005>
2. Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 20(3), 973–989.
<https://doi.org/10.1177/1461444816676645>
3. Andreou, A., Venkatadri, G., Goga, O., Gummadi, K. P., Loiseau, P., & Mislove, A. (2018). Investigating ad transparency mechanisms in social media: A case study of Facebook's explanations. In *Proceedings of the Network and Distributed System Security Symposium (NDSS)*. <https://doi.org/10.14722/ndss.2018.23191>
4. Bleier, A., & Eisenbeiss, M. (2015). The importance of trust for personalized online advertising. *Journal of Retailing*, 91(3), 390–409. <https://doi.org/10.1016/j.jretai.2015.04.001>
5. Boerman, S. C., Kruikemeier, S., & Zuiderveen Borgesius, F. J. (2017). Online behavioral advertising: A literature review and research agenda. *Journal of Advertising*, 46(3), 363–376.
<https://doi.org/10.1080/00913367.2017.1339368>
6. Brehm, J. W. (1966). *A theory of psychological reactance*. Academic Press.
7. Calo, R. (2014). Digital market manipulation. *George Washington Law Review*, 82(4), 995–1051.

8. Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227–268.
https://doi.org/10.1207/S15327965PLI1104_01
9. Eslami, M., Krishna Kumaran, S. R., Sandvig, C., & Karahalios, K. (2018). Communicating algorithmic process in online behavioral advertising. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Paper 432). ACM.
<https://doi.org/10.1145/3173574.3174006>
10. Friestad, M., & Wright, P. (1994). The persuasion knowledge model: How people cope with persuasion attempts. *Journal of Consumer Research*, 21(1), 1–31.
<https://doi.org/10.1086/209380>
11. Gray, C. M., Kou, Y., Battles, B., Hoggatt, J., & Toombs, A. L. (2018). The dark (patterns) side of UX design. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Paper 534). ACM. <https://doi.org/10.1145/3173574.3174108>
12. Kim, T., Barasz, K., & John, L. K. (2019). Why am I seeing this ad? The effect of ad transparency on ad effectiveness. *Journal of Consumer Research*, 45(5), 906–932.
<https://doi.org/10.1093/jcr/ucy039>
13. Kizilcec, R. F. (2016). How much information? Effects of transparency on trust in an algorithmic interface. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 2390–2395). ACM. <https://doi.org/10.1145/2858036.2858402>
14. Luguri, J., & Strahilevitz, L. J. (2021). Shining a light on dark patterns. *Journal of Legal Analysis*, 13(1), 43–109. <https://doi.org/10.1093/jla/laaa006>
15. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems 30* (pp. 4765–4774).
16. Martin, K. D., Borah, A., & Palmatier, R. W. (2017). Data privacy: Effects on customer and firm performance. *Journal of Marketing*, 81(1), 36–58. <https://doi.org/10.1509/jm.15.0497>
17. Mathur, A., Acar, G., Friedman, M. J., Lucherini, E., Mayer, J., Chetty, M., & Narayanan, A. (2019). Dark patterns at scale: Findings from a crawl of 11K shopping websites. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), Article 81.
<https://doi.org/10.1145/3359183>
18. Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20(3), 709–734.
<https://doi.org/10.5465/amr.1995.9508080335>
19. Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. MIT Press.
20. Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end

- framework for internal algorithmic auditing. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 33–44). ACM. <https://doi.org/10.1145/3351095.3372873>
21. Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206–215. <https://doi.org/10.1038/s42256-019-0048-x>
22. Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, 146, 102551. <https://doi.org/10.1016/j.ijhcs.2020.102551>
23. Sunstein, C. R. (2016). Fifty shades of manipulation. *Journal of Marketing Behavior*, 1(3–4), 213–244. <https://doi.org/10.1561/107.00000014>
24. Susser, D., Roessler, B., & Nissenbaum, H. (2019). Online manipulation: Hidden influences in a digital world. *Georgetown Law Technology Review*, 4(1), 1–45.
25. Tucker, C. E. (2014). Social networks, personalized advertising, and privacy controls. *Journal of Marketing Research*, 51(5), 546–562. <https://doi.org/10.1509/jmr.10.0355>
26. Wachter, S., Mittelstadt, B., & Russell, C. (2018). Counterfactual explanations without opening the black box: Automated decisions and the GDPR. *Harvard Journal of Law & Technology*, 31(2), 841–887.
27. Wang, W., & Benbasat, I. (2007). Recommendation agents for electronic commerce: Effects of explanation facilities on trusting beliefs. *Journal of Management Information Systems*, 23(4), 217–246. <https://doi.org/10.2753/MIS0742-1222230410>
28. White, T. B., Zahay, D. L., Thorbjørnsen, H., & Shavitt, S. (2008). Getting too personal: Reactance to highly personalized email solicitations. *Marketing Letters*, 19(1), 39–50. <https://doi.org/10.1007/s11002-007-9027-9>
29. Xiao, B., & Benbasat, I. (2007). E-commerce product recommendation agents: Use, characteristics, and impact. *MIS Quarterly*, 31(1), 137–209. <https://doi.org/10.2307/25148784>