INTERNATIONAL REVIEW OF ARTIFICIAL INTELLIGENCE IN MARKETING

NAWALA

Volume 4, Number 1, 2025

Trusting the Machine: Consumer Perceptions of AI Versus Human Recommendations

Annisa Rahma Dianti [1]

[1] Universitas Diponegoro, Semarang, Indonesia

Abstract

Article history:
Received: January 9, 2025
Revised: February 14, 2025
Accepted: April 8, 2025
Published: June 30, 2025

Keywords:
Algorithm Aversion, Artificial Intelligence, Consumer Trust, Human Recommendation, Systematic Literature Review.

Identifier:
Nawala
Page: 27-37
https://nawala.io/index.php/iraim

This study investigates how consumers perceive and trust artificial intelligence based recommendation systems compared with human recommendations across decision contexts. Using a systematic literature review method, the article synthesises empirical evidence on perceived competence, impartiality, empathy, transparency, learning capability, and privacy concerns linked to artificial intelligence driven advice. The review identifies trust as a multidimensional judgement that combines beliefs about technical performance with inferences about benevolence and integrity. The findings show that consumers may value artificial intelligence recommenders for their efficiency and perceived objectivity, yet often experience them as opaque and threatening to personal control and privacy. Algorithm aversion emerges when visible errors lead consumers to penalise artificial intelligence more harshly than human advisors, although demonstrations of learning and clear explanations can partially restore trust. Overall, the review concludes that willingness to follow artificial intelligence recommendations depends on how systems are governed and communicated, and on whether human warmth and fairness are seen as necessary in the specific decision domain.

*Corresponding author:
(Annisa Rahma Dianti)

1. Introduction

Artificial intelligence driven recommendation systems have become deeply embedded in consumer decision journeys, from e-commerce product suggestions and content curation to financial advice and healthcare support. As firms increasingly replace or augment human frontline advisors with algorithmic systems, many choice situations that were once mediated by salespeople, experts, or peers are now shaped by machine-generated recommendations. Recent work on trust in artificial intelligence shows that consumers do not evaluate these systems purely on technical performance; instead, they form trust judgements based on perceptions of competence, reliability, and integrity, much as they do with human advisors (Bitkina et al., 2020; Henrique & Santos, 2024). At the same time, the rapid diffusion of AI into everyday services has intensified public debate about whether algorithms can be trusted to act in consumers' best interests, particularly when decisions are opaque or data-intensive.

A growing stream of research documents that consumers often react differently to decisions made by algorithms versus humans, even when objective accuracy is comparable. Experimental evidence shows that people sometimes penalize algorithmic decision makers more harshly for errors and may be less willing to accept unfavorable outcomes from an AI than from a human, a phenomenon broadly discussed as algorithm aversion (Reich et al., 2022; Yalcin et al., 2022). Yet, trust in AI is not uniformly low: studies of AI service adoption suggest that when consumers perceive AI as efficient, objective, and properly governed, trust in the company and in the system can support high willingness to use AI-based services

(Frank et al., 2023). Recent literature reviews similarly highlight that trust in AI is shaped by perceptions of transparency, fairness, and the system's ability to learn and improve, and that miscalibrated trust either over or under trust can undermine the benefits of algorithmic decision support (Henrique & Santos, 2024, Bitkina et al., 2020).

Emerging work directly comparing AI and human agents suggests that the identity of the recommender can systematically alter consumer responses. For instance, research on information disclosure finds that consumers may trust brands less when they are asked to share personal data with an AI rather than a human, because they infer a larger, less controllable audience and feel more exploited (Lefkeli et al., 2024). Other studies show that disclosing the use of AI in message creation or prosocial advertising can change attitudes and behavioral intentions, indicating that consumers hold distinct lay beliefs about the motives, capabilities, and limitations of machine versus human sources (Baek et al., 2024). However, existing research remains fragmented across domains and often focuses either on generic trust in AI technologies or on single contexts, offering limited insight into how consumers consciously weigh AI versus human recommendations when making everyday choices. Against this backdrop, the present study examines how consumers perceive and trust AI-based recommendations relative to human recommendations, and how these perceptions shape their intention to follow the advice. By unpacking the roles of perceived competence, impartiality, empathy, and privacy concerns in these comparative judgements, the study seeks to clarify when consumers are willing to "trust the machine" and when they continue to prefer human advice.

2. Literature Review

Research on trust in artificial intelligence has established that users do not only evaluate AI systems on their technical accuracy, but also on relational cues such as benevolence, integrity, and transparency. Glikson and Woolley (2020) synthesize empirical findings across multiple domains and show that trust in AI depends on how the system is represented (robot, virtual agent, or embedded system), its perceived level of machine intelligence, and its reliability over time. They emphasize that transparency, reliability, and immediacy behaviours shape cognitive trust, while anthropomorphic cues are especially important for emotional trust, suggesting that AI recommenders may be trusted differently depending on how "human-like" or opaque they appear.

From a consumer perspective, Puntoni et al. (2021) argue that AI should be understood through four experiential modes data capture, classification, delegation, and social interaction which jointly shape whether consumers see AI as empowering or exploitative. Their review highlights that, even when AI recommendations improve efficiency and personalization, consumers often worry about loss of control, surveillance, and misaligned motives, which can undermine trust in AI-driven advice and recommendation systems.

A key barrier to trusting AI recommenders is "algorithm aversion", whereby people avoid or penalize algorithmic advice when it makes errors, even if its average performance is superior. Reich et al. (2022) demonstrate that consumers are particularly sensitive to visible algorithmic mistakes, but that showing evidence the algorithm can learn from its errors can significantly reduce algorithm aversion and

increase reliance on AI advice over time. Complementary work by Wesche et al. (2024) compares human versus algorithmic decision makers in workplace selection contexts and finds that people generally react more favourably to human decision makers, especially in terms of fairness and acceptance, although clear explanations can partly mitigate negative reactions to algorithmic decisions. These findings suggest that the identity of the decision agent (human versus AI), perceived learning capability, and the availability of explanations jointly shape trust and willingness to accept AI-based recommendations.

Another stream of research examines how explicit disclosure of AI involvement influences trust. Baek et al. (2024) show that disclosing that prosocial advertising messages are AI-generated initially leads to less favourable ad evaluations and lower donation intentions, primarily because disclosure reduces perceived ad credibility. However, the negative effect of disclosure is weaker when consumers perceive AI as more human-like and when message credibility is high, indicating that trust in AI sources can be repaired under specific conditions. Taken together, these studies indicate that consumer trust in AI-based recommendations is shaped by perceived competence, fairness, transparency, learning capability, and source identity, yet existing work remains fragmented across contexts and rarely examines how consumers explicitly weigh AI versus human recommenders when deciding whether to follow advice. This fragmentation motivates further comparative research on when consumers are willing to "trust the machine" and when they still prefer human advisors.

3. Methods

The present study adopts a systematic literature review (SLR) method to synthesise existing evidence on how consumers perceive and trust AI-based recommendations relative to human recommendations. The review followed the standard SLR stages of planning, searching, screening, quality appraisal, and synthesis. First, a review protocol was developed that defined the research questions, conceptual focus (trust, algorithm aversion, perceptions of competence, impartiality, empathy, transparency, learning capability, and privacy concerns), and inclusion criteria. Second, a comprehensive search was carried out in major academic databases such as Scopus, Web of Science, ScienceDirect, and Google Scholar using combinations of keywords related to artificial intelligence, recommendation systems, consumer trust, algorithm aversion, and human versus algorithmic decision makers. Only peer-reviewed journal articles written in English and directly examining consumer responses to AI or algorithmic systems in decision or recommendation contexts were included, while non-scholarly sources, dissertations, books, and purely technical model-development studies without a behavioural component were excluded. After removing duplicates, titles and abstracts were screened for relevance, followed by full-text assessment against the inclusion criteria.

The methodological quality of the retained articles was appraised using a structured checklist that considered clarity of research design, adequacy of sample and context description, transparency of measurement and analysis, and robustness of conclusions. For each study, key data were extracted into a coding template covering context, type of AI application, presence or absence of a human

comparison agent, operationalisation of trust and related constructs, and main findings regarding consumers willingness to accept or follow AI-based recommendations. Finally, a narrative and thematic synthesis was conducted to identify recurring patterns and divergences across studies, organise the evidence around core antecedents and outcomes of trust in AI versus human recommenders, and highlight conceptual gaps and directions for future research.

4. Results and Discussion

The systematic review reveals that consumer trust in AI-based recommendation systems is jointly shaped by perceptions of technical competence, relational qualities, and the social meaning of delegating decisions to machines. Across the reviewed studies, trust emerges not as a single attitude but as a composite judgement that combines beliefs about reliability and performance with inferences about benevolence, integrity, and transparency (Bitkina et al., 2020; Henrique & Santos, 2024). In line with Glikson and Woolley's (2020) emphasis on representation and perceived machine intelligence, the findings show that consumers differentiate between AI systems embedded in everyday services and more "visible" agents such as chatbots or virtual assistants. When AI recommenders are experienced as opaque "black boxes", concerns about control, surveillance, and misaligned motives become salient, echoing Puntoni et al. (2021) view that data capture and classification can make AI feel more exploitative than empowering. Conversely, when firms communicate clear governance, safeguards, and learning capabilities, consumers are

more likely to treat AI as a competent and objective advisor, which supports higher willingness to adopt AI based services (Frank et al., 2023; Henrique & Santos, 2024).

A second key pattern concerns the tension between perceived impartiality and emotional warmth in AI versus human recommendations. Consistent with work on algorithm aversion, the reviewed studies indicate that consumers tend to penalise AI more harshly than humans for visible errors and are less willing to accept unfavourable outcomes from algorithmic decisions, even when average accuracy is comparable (Yalcin et al., 2022). This aligns with Reich et al.'s (2022) evidence that demonstrating an algorithm's ability to learn from its mistakes can partially restore trust and increase reliance on AI advice over time, suggesting that dynamic learning signals are crucial for calibrating trust. At the same time, research comparing human and algorithmic decision makers shows that people generally perceive human agents as fairer and more acceptable, particularly in consequential contexts, although transparent explanations can mitigate resistance to algorithmic outcomes (Wesche et al., 2024). Studies on disclosure and data sharing further refine this picture: when personal data are requested or when messages are explicitly labelled as AI-generated, consumers often report lower brand trust and weaker behavioural intentions, driven by fears of a wider, less controllable audience and reduced message credibility (Baek et al., 2024; Lefkeli et al., 2024). Taken together, these findings suggest that consumers see AI recommenders as potentially more competent and impartial but less empathetic and more threatening to privacy, and that their intention to "trust the machine" versus preferring human advice depends on how transparency,

governance, learning capability, and source identity are communicated and experienced across specific decision contexts.

5. Conclusion

This study shows that consumer trust in AI-based recommendation systems is shaped by a complex interplay of technical, relational, and contextual factors. Consumers do not simply evaluate whether AI produces accurate outputs, but also whether it appears transparent, fair, controllable, and aligned with their interests. The synthesis highlights that AI recommenders can be seen as efficient and impartial advisors when they are framed as well-governed, explainable, and capable of learning from mistakes. At the same time, many consumers still experience AI as a "black box" that amplifies concerns about surveillance, loss of control, and exploitation, especially in data-intensive contexts. These ambivalent perceptions help explain why trust in AI is often fragile and why willingness to follow AI-based recommendations varies across situations, applications, and disclosure conditions.

The review also underscores that human and AI recommenders are not evaluated on the same psychological dimensions. While AI may be perceived as more objective, humans are generally granted more empathy, moral understanding, and fairness, particularly in consequential or sensitive decisions. Consumers tend to penalise AI more severely for visible errors and react negatively when AI involvement is highlighted in ways that trigger privacy concerns or reduce message credibility. For practitioners, these findings imply that successful deployment of AI recommendation systems requires more than technical optimisation: it demands

careful design of transparency, communication, and governance that explicitly addresses consumer concerns about control, privacy, and accountability. For researchers, the fragmented evidence base points to the need for more comparative, context-sensitive studies that examine how consumers actively choose between "trusting the machine" and relying on human advice across different domains, stakes, and emotional climates.

References

Baek, T. H., Kim, J., & Kim, J. H. (2024). Effect of disclosing AI-generated content on prosocial advertising evaluation. *International Journal of Advertising*, 1–22.

Bitkina, O. V., Jeong, H., Lee, B. C., Park, J., Park, J., & Kim, H. K. (2020). Perceived trust in artificial intelligence technologies: A preliminary study. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 30(4), 282–290.

Frank, D. A., Jacobsen, L. F., Søndergaard, H. A., & Otterbring, T. (2023). In companies we trust: Consumer adoption of artificial intelligence services and the role of trust in companies and AI autonomy. *Information Technology & People*, 36(8), 155–173.

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, 14(2), 627–660.

Henrique, B. M., & Santos Jr, E. (2024). Trust in artificial intelligence: Literature review and main path analysis. *Computers in Human Behavior: Artificial Humans*, 2(1), 100043.

Lefkeli, D., Karataş, M., & Gürhan-Canli, Z. (2024). Sharing information with AI (versus a human) impairs brand trust: The role of audience size inferences and sense of exploitation. *International Journal of Research in Marketing*, 41(1), 138–155.

Puntoni, S., Reczek, R. W., Giesler, M., & Botti, S. (2021). Consumers and artificial intelligence: An experiential perspective. *Journal of Marketing*, 85(1), 131–151.

Reich, T., Kaju, A., & Maglio, S. J. (2023). How to overcome algorithm aversion: Learning from mistakes. *Journal of Consumer Psychology*, 33(2), 285–302.

Wesche, J. S., Hennig, F., Kollhed, C. S., Quade, J., Kluge, S., & Sonderegger, A. (2024). People's reactions to decisions by human vs. algorithmic decision-makers: The role of explanations and type of selection tests. *European Journal of Work and Organizational Psychology*, 33(2), 146–157.

Yalcin, G., Lim, S., Puntoni, S., & van Osselaer, S. M. (2022). Thumbs up or down: Consumer reactions to decisions by algorithms versus humans. *Journal of Marketing Research*, 59(4), 696–717.