

The impact of anonymity in AI service agent interaction on unethical consumer behaviour – A dehumanisation theory perspective

Abstract:

Artificial intelligence (AI) in consumer services add value to businesses in providing better value and satisfaction to consumers. However, this fast development requires critical investigation into the ethical implications of consumer behaviour when interacting with AI, because the anonymity provided by AI can significantly encourage retail consumers to engage in unethical behaviour. We will employ a mixed-methodology approach, including 25 participant interviews and a bottom-up thematic analysis to analyse the qualitative data and explore participants' motivations. The second phase will be conducting eight experimental studies to understand how the motivations of participants when interacting with the anonymity of AI will correlate with unethical consumer behaviour. This paper examines how anonymity in AI service agent interactions contributes to unethical consumer behaviour through the lens of dehumanisation theory, a psychological process that reduces empathy toward non-human entities, thereby justifying unethical actions, with a specific focus on the UK retail market.

Keywords: Artificial intelligence, service agents, unethical consumer behaviour, retail, anonymity

1. Introduction

Integrating AI into personalized or automated customer service, is renovating the marketplace. Anonymous AI systems offer convenience and effectiveness in the retail consumer journey and experience; however, they are sometimes abused by unethical consumers. Consumers may behave unethically in various ways; however, this may cost businesses a lot (Kim et al., 2023; Lei et al., 2024) The deceptive issues, for example, product returns, cost US businesses around 50 billion dollars (Skeldon, 2019); along with this, in 2019, 107,000 deceptive insurance claims were made in the UK (ABI, 2020). Furthermore, unethical consumer behaviour has been detected in attempts to obtain benefits from promotional offers and deceptive product returns, as well as providing false information about selling second-hand cars at inflated prices and making misleading claims about accidents from insurance companies (Zhao and Xu, 2013). Each case involves deceptive actions taken by unethical consumers to gain an economic benefit or advantage.

In these cases, the identification of unethical behaviour by consumers during interactions with human service staff revealed there is a lack of clarity regarding whether consumers behave unethically or deceptively during interactions with anonymous AI services. Therefore, this paper focuses on this under-investigated field, examining the motivations and drives of

unethical consumer behaviour during interactions with anonymous AI service agents. The UK retail context is particularly relevant due to the growth of digital operations in the marketplace and the evolving regulatory landscape surrounding AI (Gross and Stacey, 2024; Gikay, 2024). To show the viability and practical relevance, the study derives from examples of existing AI technologies broadly employed in the retail sector. For example, in the UK and internationally, retailers such as ASOS and Amazon have implemented AI-driven chatbots and automated return systems, enabling consumers to process product returns or submit complaints without interacting with a human agent. These automated return systems often require limited authentication, which decreases perceived accountability but increases the possibility of dishonest behaviour (Skeldon, 2019). Also, retail consumers have been found to exploit promotional tools or recommendation algorithms by creating multiple accounts to claim repetitive discounts or reimbursements. Such technologies embody high levels of anonymity and low human resemblance ideal conditions for testing the psychological effects theorized in our study.

Improving retail customer experience and customer journey is the primary motivation driving retail firms to adopt AI more widely. However, the present research emphasises that adopting anonymous AI to perform marketing and service functions could also encounter unexpected costs due to increasing unethical consumer behaviours. Previous research supports our predictions by reporting recent AI service agent exploitations during service process interactions with AI (Schrage, 2016). The primary objective of this research is to investigate the impact of anonymity in AI service agent interactions on unethical consumer behaviour in retail contexts.

Key contributions of the study are, firstly, this study will extend the dehumanisation theory by applying it to the context of human-AI interactions for the first time, as Dehumanisation theory has traditionally been used in human-to-human interactions. Secondly, it will link AI anonymity to the ethical decision-making processes in the retail context.

2. Literature review

Existing literature on consumer ethics primarily focuses on traditional forms of unethical behaviour, such as shoplifting, fraud, and misrepresentation (Smith & Quelch, 1993). However, one of the primary reasons consumers may act unethically when interacting with AI is the sense of anonymity and the perception that their actions will not have direct consequences. AI systems often hide human identities, making consumers feel less accountable for their behaviour.

LaMothe & Bobek (2020) argue that consumers may be more likely to cheat, lie, or engage in deceptive practices when interacting with AI systems because they perceive a minimal risk of being caught or judged. This is especially true for digital platforms where AI mediates interactions and may not directly enforce ethical norms. In the same vein, Zhong et al., (2010) found that anonymity decreases social accountability and may lead to growing unethical behaviour among consumers, such as deception. However, online anonymity in social media or digital platforms has been the focus of most research but limited attention has been paid to AI service agents driven retail environment. There was no extensive exploration of their influence on consumer ethics (Wirtz et al., 2018). Lastly, Lei et al. (2024) demonstrated that consumers are probably to engage in unethical behaviour when interacting with AI agents perceived as "servants" rather than "partners," due to increased moral disengagement.

2.1. Unethical consumer behaviour

Unethical consumer behaviour is defined as dishonest, deceptive, or illegal activities of consumers that adversely impact businesses. These behaviours can range from minor ethical violations, such as falsely claiming a promotion, to serious offences like shoplifting or fraudulent returns (Skeldon, 2019). Research has explored various factors that influence unethical behaviour, including personal traits such as low-level moral identity and high-level Machiavellianism (Cox et al., 1993). Additionally, the nature of the target—whether perceived as an individual or a group—affects consumer misconduct (Köbis et al., 2019; Yam & Reynolds, 2016). Recent studies suggest that unethical behaviour is more prevalent when consumers interact with AI-driven services (e.g., virtual assistants) instead of human employees (LaMothe & Bobek, 2020).

2.2. Dehumanization Theory and AI Interactions

AI empowers consumers to engage with digital platforms anonymously. This improves privacy and security, but it also opens avenues for dishonest and unethical consumer behaviour. It leads to dehumanization because anonymity can remove social restraints, making it easier for individuals to engage in dehumanizing behaviours, such as reducing empathy, accountability, and moral responsibility. Furthermore, dehumanization occurs when individuals perceive others as lacking human-like qualities and often reduce moral obligations toward them (Haslam, 2006). While dehumanization has been connected to interpersonal exploitation, its role in AI interactions continues underexplored. In human-AI interactions, consumers may dehumanise AI agents and exploit them by attributing less human-like qualities, which can result in unethical behaviour. For instance, consumers might justify fraudulent behaviour

against AI agents because they are perceived as ordinary machines with no ability to suffer or feel sad.

A lack of direct research linking anonymity in AI interactions to unethical consumer behaviour through the lens of dehumanisation theory. However, there is an over-reliance on self-reported survey data and a minimal focus on moderating factors, such as dissimilar cultures, moral identity, and different levels of AI humanisation.

3. Methodology

Our study employs a mixed-methods approach. Eight quantitative experiments and 25 qualitative interviews will be conducted. This approach provides for a clear understanding of how anonymity and dehumanization impact unethical behaviour.

The methodology is proposed to reflect real-world interactions between consumers and AI agents in the retail sector. For example, the automated refund systems of Amazon, virtual assistants of ASOS chatbot, and promotional recommendation engines often let users interact anonymously, submit claims, or receive offers without human authentication. These settings are not only common in the UK's retail sector but provide identical conditions: anonymity, lack of social cues, and low anthropomorphism in which dehumanization and unethical behaviour are hypothesized to occur. Our experimental and qualitative designs replicate these contexts to test psychological responses and behaviours.

An experimental phase will be set out a sample of 1000 consumers (aged 18-60) who have interacted with AI service agents in retail settings will be recruited, ensuring diversity in gender, culture, and prior experience with AI.

3.1. Sampling

To ensure both statistical generalizability and theoretical richness, we employed a dual sampling strategy: stratified random sampling for the experimental phase and purposive sampling for the qualitative phase. Experimental Study (n = 1000): Participants will be recruited through a national panel provider using stratified random sampling (Malhotra & Birks, 2007). Quotas will reflect UK population distributions based on:

Age (18–24, 25–34, 35–44, 45–54, 55–60), to capture age-based variation in technology adoption and ethical framing. **Gender**, balanced to approximate national distribution, with inclusive options for non-binary identities. **Geography**, ensuring coverage from England, Scotland, Wales, and Northern Ireland across both urban and rural settings. **Ethnicity**, reflecting UK Census categories (White, Asian, Black, Mixed, Other) to include cultural differences in ethical judgement. **Socioeconomic standing**, determined via education, income, and employment, which may influence moral reasoning and consumer strategies (Vitell, 2003).

Digital literacy and AI familiarity, screened through a pre-questionnaire to ensure participants regularly engage in online retail and have encountered AI systems such as chatbots, recommendation engines, or automated returns. Random allocation to experimental conditions ensures a balanced distribution across demographics and reduces sampling bias (Cohen, 1992).

25 Qualitative Interviews: We adopt a purposive, maximum variation sampling strategy (Patton, 2002), selecting participants who meet the following criteria: Aged 18–60, with prior experience using AI in retail customer service (e.g., AI chatbots, automated complaint systems). Variation across gender, ethnicity, digital proficiency, and moral identity (assessed using the scale by Aquino & Reed, 2002). Inclusion of both consumers who report engaging in questionable behaviours and those who do not, enabling contrastive thematic development. Interview recruitment will continue until thematic saturation is achieved, which is expected to occur around the 20th to 25th interview (Guest, Bunce, & Johnson, 2006). This approach provides depth, diversity, and theoretical alignment with the research objective.

3.2. Experimental Procedure

Manipulation of anonymity: participants will be randomly assigned to one of two conditions:

1) Anonymous interaction: AI-driven systems allow consumers to interact anonymously, which can sometimes lead to dishonest or unethical behaviour. Participants interact with AI agents without providing identifying information.

2) Identifiable Interaction: participants interact with AI agents where their information is acknowledged (e.g., through account details or personalization).

Measurement of unethical behaviour: participants will be presented with scenarios (e.g., reporting a fake issue with a product) and will be able to make unethical choices. Their behaviour will be recorded.

Measurement of dehumanization: participants will complete scales assessing their perceptions of the AI agent's human-like qualities, including emotions, intelligence, and empathy.

3.3. Data Analysis

Quantitative Analysis

Mediation Analysis: we will test whether **dehumanization mediates** the relationship between anonymity and unethical consumer behaviour.

Moderation Analysis: we will assess whether individual differences (e.g., moral identity, empathy) moderate the effect of anonymity on unethical behaviour.

Qualitative Analysis: our study employs to collect the interview data and identify key themes related to perceptions of AI and its influence on moral decision-making. This study aims to

provide insights into the psychological mechanisms underlying consumer interactions with AI agents in retail environments. It will explore whether anonymity and dehumanization lead to more unethical consumer behaviour and whether human-like attributes in AI agents can mitigate these effects. Furthermore, the moderating role of personal characteristics (e.g., moral identity) will be explored.

4. Conclusion

This research will make theoretical and practical contributions. The theoretical contribution will expand dehumanization theory to the context of human-AI interactions and unethical consumer behaviour. This study will expand our understanding of how AI anonymity and dehumanization influence unethical consumer behaviour offering a new framework. A practical contribution will be providing findings that will help retailers design AI systems that mitigate unethical behaviour by incorporating human-like qualities, empathy, or transparency. Moreover, retailers can use these findings to design AI systems that incorporate humanising features to mitigate unethical behaviour and foster longer-lasting customer relationships.

Expansion of dehumanization theory

Dehumanization theory traditionally focuses on human-to-human interactions, but this research will explore how the perceived lack of humanity in AI systems leads to reduced empathy and increased unethical consumer behaviour. It will provide a new perspective on how consumers' treatment of non-human agents can be influenced by the perceived "humanness" of AI.

Linking AI anonymity to ethical decision-making: By investigating the relationship between AI's anonymity (i.e., machine-like qualities) and consumer behaviour, this research will contribute to understanding how AI anonymity affects ethical decision-making. The study will offer insights into why consumers might feel less morally accountable when interacting with anonymous or "faceless" AI systems and how this can contribute to unethical behaviour, such as exploitation or dishonesty.

Insights into ethical consumer behaviour: The research will contribute to the field of consumer behaviour by examining the role of AI anonymity in promoting unethical behaviour. It will provide a clear understanding of how consumers' lack of emotional attachment to AI (due to its perceived dehumanization) can influence their ethical stance, leading to unethical actions, such as fraud, unethical consumption, or unfair exploitation of AI features.

Contributions to Dehumanization Theory

Our study expands dehumanisation theory to human-AI interactions by examining how consumers perceive AI service agents in the absence of certain human qualities, thereby reducing empathy and moral accountability. While dehumanization theory has traditionally

focused on interpersonal contexts (Haslam, 2006), we show it also occurs toward non-human agents, empowering unethical behaviour such as deception and exploitation.

In doing so, our study also refines moral disengagement theory (Bandura, 1999) by dividing dehumanization as a key mechanism triggered by AI anonymity and low anthropomorphism. Prior studies emphasize shadowing or risk; our findings suggest that unethical behaviour stems from a perceived moral license to exploit systems viewed as non-emotional.

The study makes three contributions: it reconceptualizes dehumanization theory beyond human targets, identifies a specific psychological pathway in moral disengagement, and offers a new explanation for unethical consumer behaviour in AI-mediated retail environments. These insights advance theory at the intersection of consumer ethics, AI design, and moral psychology.

References

- ABI (2020). Detected insurance fraud – new data shows that every five minutes a fraudulent claims discovered. Retrieved June 10, 2021 from <https://www.abi.org.uk/news/news-articles/2020/09/detectedinsurance-fraud/>.
- Aquino, K., & Reed, A. II. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, 83(6), 1423–1440. <https://doi.org/10.1037/0022-3514.83.6.1423>
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>
- Cox, A. D., Cox, D., Anderson, R. D., & Moschis, G. P. (1993). Research note: Social influences on adolescent shoplifting—Theory, evidence, and implications for the retail industry. *Journal of Retailing*, 69(2), 234–246.
- Gikay, A. A. (2024). Risks, innovation, and adaptability in the UK's incrementalism versus the European Union's comprehensive artificial intelligence regulation. *International Journal of Law and Information Technology*, 32(1). <https://doi.org/10.1093/ijlit/eaee013>
- Gross, A., & Stacey, S. (2024). UK will legislate against AI risks in next year, pledges Kyle, *Financial Times*, Accessed through: [UK will legislate against AI risks in next year, pledges Kyle](#).
- Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? *Field Methods*, 18(1), 59–82. <https://doi.org/10.1177/1525822X05279903>

- Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, 10(3), 252–264. https://doi.org/10.1207/s15327957pspr1003_4
- Malhotra, N. K., & Birks, D. F. (2007). *Marketing Research: An Applied Approach* (3rd ed.). Pearson Education.
- Patton, M. Q. (2002). *Qualitative Research and Evaluation Methods* (3rd ed.). Sage Publications.
- Kim, T., Lee, H., Kim, M.Y. *et al.* AI increases unethical consumer behavior due to reduced anticipatory guilt. *J. of the Acad. Mark. Sci.* **51**, 785–801 (2023). <https://doi.org/10.1007/s11747-021-00832-9>
- Köbis, N. C., Verschuere, B., Bereby-Meyer, Y., Rand, D., & Shalvi, S. (2019). Intuitive honesty versus dishonesty: Meta-analytic evidence. *Perspectives on Psychological Science*, 14(5), 778–796.
- LaMothe, E., & Bobek, D. (2020). Are individuals more willing to lie to a computer or a human? Evidence from a tax compliance wetting. *Journal of Business Ethics*, 167, 157–180.
- Lei, S., Xie, L., & Peng, J. (2024). Unethical Consumer Behavior Following Artificial Intelligence Agent Encounters: The Differential Effect of AI Agent Roles and its Boundary Conditions. *Journal of Service Research*, 0(0). <https://doi.org/10.1177/10946705241278837>
- Schrage, M. (2016). Why you shouldn't swear at Siri. Retrieved June 10, 2021 from <https://hbr.org/2016/10/why-you-shouldnt-swear-at-siri>.
- Skeldon, P. (2019). ASOS leads the way in ending 'Wardrobing' trend that costs retail £60bn a year. Retrieved November 4, 2021 from <https://internetretailing.net/industry/industry/asos-leads-the-way-inending-wardrobing-trend-that-costs-retail-60bn-a-year-19443>.
- Smith, N. C., & Quelch, J. A. (1993). *Ethics in marketing*. Homewood, IL: Irwin.
- Wirtz, J., Patterson, P. G., Kunz, W. H., Gruber, T., Lu, V. N., Paluch, S., & Martins, A. (2018). Brave new world: Service robots in the frontline. *Journal of Service Management*, 29(5), 907–931.
- Vitell, S. J. (2003). Consumer ethics research: Review, synthesis and suggestions for the future. *Journal of Business Ethics*, 43(1–2), 33–47. <https://doi.org/10.1023/A:1022918515436>
- Yam, K. C., & Reynolds, S. J. (2016). The effects of victim anonymity on unethical behaviour. *Journal of Business Ethics*, 136(1), 13–22.

- Zhao, B., & Xu, S. (2013). Does consumer unethical behaviour relate to birthplace? Evidence from China. *Journal of Business Ethics*, 113(3), 475–488.
- Zhong, C.-B., Bohns, V. K., & Gino, F. (2010). Good Lamps Are the Best Police: Darkness Increases Dishonesty and Self-Interested Behavior. *Psychological Science*, 21(3), 311-314. <https://doi.org/10.1177/0956797609360754>