TPM Vol. 32, No. S8, 2025
ISSN: 1972-6325
https://www.tpmap.org/

Open Access

THE IMPACT OF GENERATIVE ARTIFICIAL INTELLIGENCE ON COGNITIVE BIAS, CONSUMER TRUST AND ENGAGEMENT IN DIGITAL PLATFORM: EVIDENCE FROM THE UNITED KINGDOM

HARARIA IJAZ[1], AHMED TOUQEER[2*], KIRAN AROOJE[3*], MUHAMMAD MAAZ UL HAQ[4], SYEDA MANAL FATIMA[5], SAIRA MAJID[6]

[1.] LECTURER IN BUSINESS/MODULE LEADER, LONDON SCHOOL OF SCIENCE AND TECHNOLOGY, UNITED KINGDOM
[2.] JARVIS TECHNOLOGIES, ECOMMERCE & DIGITAL MARKETING MANAGER, UNITED KINGDOM
[3.] LECTURER IN BUSINESS/COORDINATOR OF PERSONAL ACADEMIC TUTORS/MODULE LEADER, LONDON SCHOOL OF SCIENCE AND TECHNOLOGY, UNITED KINGDOM
[4.] ASSISTANT MANAGER IT, BS INFORMATION TECHNOLOGY, DEPARTMENT OF IT, UNIVERSITY OF EDUCATION, LAHORE, PAKISTAN
[5.] PHD SCHOLAR (CLINICAL PSYCHOLOGY), DEPARTMENT OF PSYCHOLOGY, UNIVERSITY OF GUJRAT, PAKISTAN
[6.] HEAD OF DEPARTMENT CLINICAL PSYCHOLOGY, THE SUPERIOR UNIVERSITY, LAHORE, PAKISTAN /PHD SCHOLAR CLINICAL PSYCHOLOGY, FACULTY OF MEDICINE AND HEALTH SCIENCES, UNIVERSITY OF PUTRA MALAYSIA

Corresponding Author: ahmed.consultancy8@gmail.com , Kiran.arooje@lsst.ac

Abstract
The fast incorporation of Generative Artificial Intelligence (GenAI) into online applications has created serious implications and uncertainties for consumers regarding how GenAI affects their thoughts and actions. This study utilizes cognitive bias theory and trust-based models of consumer engagement to evaluate the impact of using GenAI on forming cognitive biases, and how cognitive biases affect consumer trust and engagement with GenAI-enabled digital platforms (GenAI-EDPs). Survey data (N=386) from GenAI users conducted through Prolific Academic were analyzed using structural equation modelling (SEM) to test the proposed model. The study's findings indicate a significant increase in cognitive biases when consumers use GenAI ($\beta = 0.57$, $p < 0.001$). Cognitive biases have an adverse effect on consumer confidence in GenAI ($\beta = -0.61$, $p < 0.001$), and consumer confidence, in turn, has a positive influence on consumer engagement with GenAI ($\beta = 0.65$, $p < 0.001$) and serves to mediate the relationship between cognitive bias and consumer engagement (Indirect effect $\beta = -0.40$, 95% CI [-0.48, -0.31]). In addition, consumer perceptions of the level of transparency associated with the use of GenAI systems have an impact on the relationship between cognitive bias and confidence ($\beta = 0.18$, $p = 0.001$). The higher the level of perceived transparency, the less negative impact cognitive bias had on consumer confidence. Based on the conclusion of this study, the results indicate that GenAI improves a user's interaction but also supports the same cognitive biases that may negatively impact a user's trust and engagement with the platform if they are not addressed. Trust and transparency are essential components to the successful and responsible long-term use of Generative Artificial Intelligence by digital platforms.
Keywords: Generative Artificial Intelligence, Cognitive Bias, Consumer Trust, Consumer Engagement,

BACKGROUND OF STUDY

The emergence of Generative Artificial Intelligence (GenAI) has altered the very foundation of how digital platforms operate and changed the methods employed to produce, customize and distribute data on these platforms to users. Unlike many prior versions of AI technology that used only rule-based or predictive approaches to develop their models and data output, GenAI technology produces text-based content, makes recommendations, creates user reviews and comments, etc., in a manner that is very similar to human cognitive capabilities and communication styles (Chughtai et al., 2022). Therefore, the way in which users interact with digital platforms has evolved from being passive-consumption based and simply disseminating information, to an increasing number of cognitive-based; interpretative and relational forms of engagement with these platforms (Dwivedi et al., 2023; Huang & Rust, 2021).

A growing body of research suggests that GenAI goes far beyond just making things more efficient for users in terms of how fast they find information or be able to customize their preferences, and also has a direct impact on users' cognitive decision making and judgement processes through both enhancing or diminishing cognitive biases while utilizing digital platforms (Baird & Maruping, 2021; Logg et al., 2019). This cognitive based influence of GenAI is particularly relevant within digital platforms because users frequently use GenAI produced content for conducting research while experiencing various distractions due to being inundated with large quantities of information; lack of certainty surrounding their decision making and having limited time to devote to conducting thorough research prior to making a purchasing decision.

Cognitive biases are crucial mechanisms of influence for generative AI technology on consumer behavior. Although current research has established the potential for users to place too much faith in their reliance on AI driven suggested decisions (i.e., automated judgment alternatives); seek out evidence supporting prior beliefs through the AI generated content (i.e., algorithmic confirmation); or base decisions relative to default and ranking suggestions made by the AI (Shin & Park, 2023; Zhang et al., 2022), cognitive biases are not incidental but are inextricably connected to the structure and function of generative AI systems that emphasize fluency, coherence and persuasive realism. Furthermore, recent studies examining cognitive bias generated by GenAI as a key psychological pathway through which subsequent consumer responses are affected are only beginning to gain empirical attention.

One of the most important effects of this avenue is on the formation of trust in consumers; as with most digital platforms, trust has historically been thought of in terms of perceived reliability (performance), perceived transparency (accountability), and perceived integrity (ethical behavior). In an AI-mediated environment, however, trust is increasingly being formed based upon the cognitive experiences of users with algorithmic outputs e.g., perceived accuracy, perceived explainability or transparency, and the degree to which they align with personal values (Shin, 2021; Siau & Wang, 2020). When GenAI has the subtle ability to create cognitive bias within users by either reinforcing pre-existing beliefs or by obscuring alternate perspectives, it produces the dual effect of increasing users' perceived competence and simultaneously decreasing their epistemic vigilance. This creates an ironic tension in which users develop greater levels of trust, behaviorally speaking, but decreased levels of trust cognitively speaking. It raises questions as to the potential for over-trusting and misplaced reliance on AI systems (Langer & König, 2023).

A key factor that influences consumers to interact with brands in an engaging way through technology platforms is the establishment of trust with these brands and their associated digital marketing communications. The engagement of a consumer with a technology-based brand involves not only the economic transaction but also the ongoing sharing of experiences and other forms of interaction such as creating user content and brand loyalty to specific technology platforms (Saif et al., 2024). Recent studies in digital marketing and information systems have shown that trust plays a role as a mediating variable that converts technological activities into an engaging experience for technology users, thus establishing a relationship between technology users and the brands that create these technologies, particularly in the case of AI-based technologies, where there is little or no boundary between the user and the machine (Marinova et al., 2023; Venkatesh et al., 2022). However, most of the current models for engagement treat trust as a byproduct of an AI system's ability or performance level without taking into consideration how trust is built cognitively.

The empirical understanding of these dynamics, particularly as it relates to generative artificial intelligence (AI), is still quite rudimentary. The reason for this is due to how the generative AI is very different qualitatively from prior generations of AI due to its level of reporting, creative, and persuasive capability. Additionally, much of the literature regarding generative AI Post 2023 is based on experimental simulation methodology and non-western data samples, reducing the ability to generalize findings to advanced digital economies such as the UK. The UK presents an excellent example of a highly developed digital economy with a high level of penetration into digital platforms, substantial numbers of users adopting different forms of AI technology and increasing levels of governance surrounding the accountability and responsibility of algorithms (OECD, 2023; UK Department for Science, Innovation and Technology, 2024).

Therefore, within this context, the current study will place cognitive biases as the mediating agent that connects the utilization of generative AI technology on digital platforms to levels of consumer trust and/or engagement. In addition, by using evidence from the UK, the study will respond to recent requests for theory-based, mechanism-based research that examines the interaction between humans and artificial intelligence technology (Dwivedi et al., 2023; Rai et al., 2024), and provide further insight into the effect that generative AI technology has on consumer purchasing behavior and the ways that these behaviors occur through the cognitive and psychological processes defined by cognitive bias.

LITERATURE REVIEW

1. Generative AI and Cognitive Bias

Cognitive biases are tendencies to systematically deviate from what's typical or logical when making decisions. Being like people as a conversational partner and font of information, GenAI is both able to help reduce cognitive biases and create cognitive biases in users. Users may be overly reliant on results from machines and may use GenAI even when there are clear examples of contradictory evidence. Recent studies have demonstrated that the fluent and/or authoritative way in which GenAI communicates contributes to this cognitive bias since users perceive a higher degree of trust when reading GenAI-constructed text when compared to text constructed by other people (Jakesch et al., 2023). In instances where GenAI is used as aide for making decisions, users may accept what GenAI presents without question. GenAI-based algorithms used to personalize content may also manifest "algorithmic echo chambers" since they provide content based on the user's current opinion or idea and create a confirmation bias that minimizes the potential for being exposed to opposite perspectives (Arechar et al, 2023). Although GenAI has the capability of executing many complex tasks detailing both sides of an issue if directed appropriately, making GenAI function as a counterbalance to cognitive bias is still in its infancy. GenAI's conversational characteristics create a significant potential for anthropomorphism. Users may conflate GenAI's natural characteristics with the intent and/or empathy of a human being. This tendency may create an illusion of increased persuasiveness (Araujo, 2020) and can also lead to manipulation or deceptive behaviors and decisions. Because anthropomorphism creates the illusion of personal agency, it impacts users' judgement.

H1a: The use of generative artificial intelligence (GenAI) positively increases automation bias among users.

H1b: The use of generative artificial intelligence (GenAI) positively increases confirmation bias among users.

2. Generative AI and Consumer Trust

Trusting digital platforms is recognized as a multi-faceted concept, based on an individual's views of their competence, integrity, and benevolence; however, recent advances in Generative AI (GenAI) will create a unique paradox for platforms' users in how they develop and sustain trust through these dimensions. From one perspective, GenAI provides a high level of apparent competence due to the ability to produce fluent written language with a high degree of speed and efficiency, as well as the ability to adaptively personalize responses to the user. Thus, for many users, this would generate a high level of trust in the GenAI. Conversely, this trust can be threatened by GenAI's frequent production of the so-called "hallucination," or outputs that appear fluent but are ultimately incorrect. This clearly threatens the belief in the reliability of GenAI, which is a primary dimension of competence-based trust. Increasingly as users become aware of these failures, particularly in well-developed digital markets, the perception of the trustworthiness of GenAI ultimately becomes increasingly fragile versus cumulative (Jacovi et al., 2024). Additionally, users may be skeptical of the integrity of many GenAI outputs because of the unknown or "black box" nature of the algorithms used to produce the output or the underlying data that serves as the basis for these algorithms. Given the requirements of the General Data Protection Regulation (GDPR) in the UK for the explanation of data media usage, accountability of data usage, and transparency of data usage, the opacity of many GenAI outputs can further erode user confidence in AI-generated content (Felzmann et al., 2020). Moreover, even when users are made aware of the AI influences on the content via explicit labeling, prior studies have found that this increased awareness may also result in perceptions of reduced credibility or value associated with the content generated from AI (Krämer et al., 2022). Beyond competence and integrity, trust is also premised on a belief in the benevolence of the system, or the belief that the system is acting in the user's best interest. In the case of GenAI platform-based applications, the large commercial incentives for maximizing user engagement, influencing user preferences, or influencing user decisions may undermine users' confidence in the degree to which the outputs generated by GenAI represent genuine helpfulness or strategic persuasion; hence, a misalignment between platform goals and user welfare. Thus, these inherent tensions with Generative AI highlight the fact that Generative AI will not only increase or decrease trust but will fundamentally shape the cognitive and moral foundation for that trust in an inherently unstable, context-dependent manner.

H2: Users' cognitive bias formation negatively affects consumer trust in GenAI-enabled digital platforms.

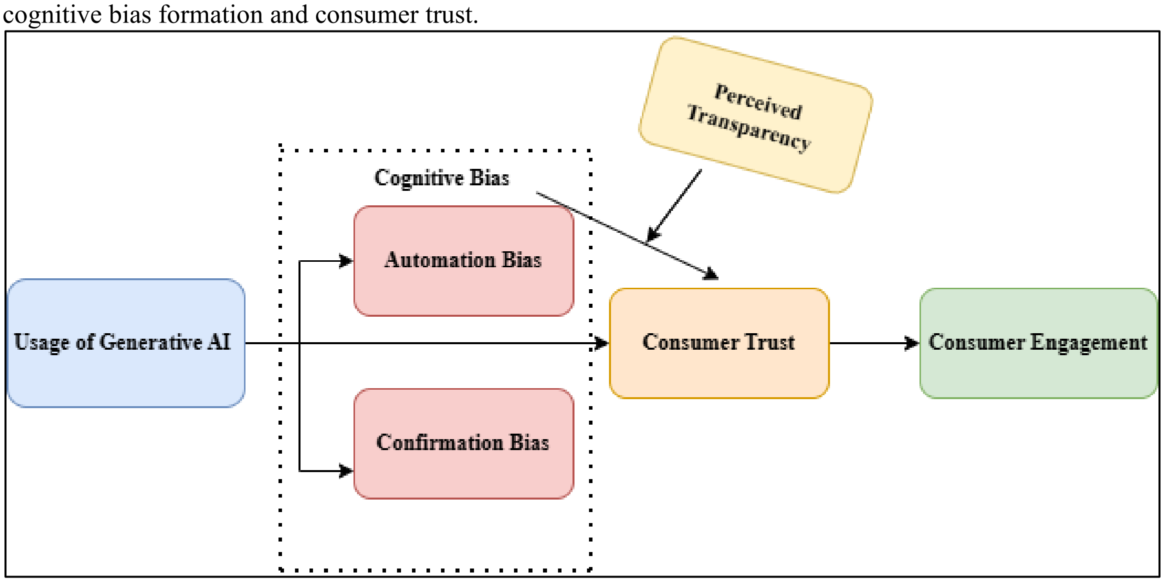
3. Generative AI and Consumer Engagement

When people say they are 'engaged,' they usually mean they have an emotional, mental, and physical connection. In the digital world, generative AI is commonly used by digital platforms to increase user engagement. Generative AI allows for much more personalized interactions and offers users dynamically generated content and real-time adaptive recommendations based on their interests and preferences, all of which contribute to increased session length, interaction frequency, and interactions relevant to conversions. More specifically, conversational interfaces like chatbots enable users to continually communicate on demand, reducing the amount of time spent searching for a product or making an educated decision on what to buy, and streamlining the user's journey through the website, thus enhancing the observable engagement behaviors (Chung et al., 2020). In addition to increasing measurable behavioral engagement, generative AI has a cognitive and emotional impact on users as well by giving users the opportunity to be creative and co-productive (e.g., producing images, text, and design concepts), thereby developing cognitive engagement, resource autonomy, and the experience of immersion. However, not all forms of engagement with AI-based platforms are positive. For example, users may be frustrated by inaccurate or inconsistent results, and users may experience anxiety and resistance to AI systems based on the broader issues of automation, data collection, and job loss (Borges et al., 2021). Thus, the outcome of engagement with generative AI is not only a function of whether there is a generative AI component, but also the perceived quality, reliability, and purpose of the experience. In addition, trust and cognitive biases are vital to the relationship between inactive engagement and engagement with generative AI systems. High trust levels likely produce greater degrees of engaged users since they are likely to use AI capabilities, whereas low trust levels may lead to decreased levels of user engagement. Cognitive automation bias may produce substantial, yet low quality, user engagement as users may become accustomed to accepting products produced by AI, and thus will produce repeated engagement, but not critically evaluating the results. At the same time, confirmation bias will increase engagement within content ecosystems; users will participate, if their views and preferences are supported, producing significant, yet polarized, levels of engagement.

H3: Consumer trust positively influences consumer engagement with GenAI-enabled digital forms.

H4: Consumer trust negatively mediates the relationship between cognitive bias formation and consumer engagement.

H5: Perceived transparency of GenAI system positively moderates the negative relationship between cognitive bias formation and consumer trust.



Conceptual Framework of Study

METHODOLOGY

Research Design

The purpose of this study was to quantitatively analyze the impact of using Generative Artificial Intelligence (GenAI) on the formation of cognitive biases among users and how this impacts consumer engagement and trust. A survey using a structured questionnaire was used to gather data from participants. This method can be effectively used to develop and test directional hypotheses based on sound theory, as well as to quantify the latent psychological constructs related to cognitive bias formation, consumer trust, and consumer engagement. To test the proposed conceptual model of how GenAI impacts cognitive bias formation among users, SEM analysis was used, allowing for estimates of direct, mediating, and moderating relationships between the proposed latent constructs.

Sampling and Recruitment

Samples of participants were collected through Prolific Academic; a crowd-sourced platform that provides a diverse and high-quality sample for behavioral & information systems research. Prolific was chosen because it has better prescreen capabilities for participants than other online panels that have less reliability on collecting quality data. Participants who were eligible to participate had to: be at least 18 years old, actively use digital platforms with GenAI features (e.g., AI chatbots, AI recommendation systems, generative content), and speak fluent English. Those who failed the attention check questions or completed the survey in an inappropriate amount of time were excluded from the final sample.

Sample Size

There was a total of 412 responses collected. After invalidating both incomplete questionnaires and low-quality questionnaires the sample size reduced to 386 valid questionnaires therefore this number exceeded the minimum sample size thresholds recommended for SEM analysis which provides adequate statistical power.

Measurement Instrument

All variables were measured using pre-validated, multi-item scales which were adapted to the context of a GenAI-enabled digital platform. Each response was captured on a scale of 1 through 5 on a Likert scale, with 1 representing "strongly disagree" and 5 representing "strongly agree."

1. Generative AI Use

The scales previously used to quantify the use of AI systems and the interaction between humans and algorithms in previous studies by Venkatesh et al. (2012) and Castelo et al. (2019) were adapted to develop a new scale that specifically measures the level and frequency of people's interactions with the features of generative AI. Some sample items from this newly developed scale include these kinds of statements: “I often utilize AI-generated recommendations or content via digital platforms,” and “Generative AI has a significant degree of influence on my use of this platform.” The new scale had considerable psychometric evidence of reliability and validity; specifically, its Cronbach's $\alpha = 0.86$, Composite Reliability (CR) = 0.88, and Average Variance Extracted (AVE) = 0.65; all of which meet the minimum psychometric requirements for scales.

2. Cognitive Bias Formation

Cognitive bias formation was developed as a construct that indicates how vulnerable a person is to using automation and confirmation cognitive biases in the decision-making process with generative AI. The items used to measure this construct were taken from the original research associated with automation and cognitive reliance on algorithmic information by Parasuraman and Riley (1997), and Logg and others (2019). Some of the example items were: “I usually accept the AI-generated input as-is, without question," “AI-generated information reflects what I believe,” and “If I receive a recommendation from AI, I do not look elsewhere." The authors reported good reliability and validity for the cognitive bias formation scale with respect to Cronbach's $\alpha = 0.88$; Composite Reliability (CR) = 0.90; Average Variance Extracted (AVE) = 0.67.

3. Consumer Trust

The trust that consumers have in GenAI enabled platforms was assessed using an adaptation of established trust scales found in the research of digital and artificial intelligence systems, namely by using the trust scales developed by McKnight et al. (2002) and Gefen et al. (2003). This scale is intended to assess users' beliefs regarding how competent, trustworthy, and honest the AI appears to be. Some example statements of the trust scale were “I think the AI on this platform is trustworthy”, “I have confidence that this platform is using AI correctly” and “The AI of this platform acts reliably”. The results from this trust scale have excellent reliability and validity with a Cronbach’s α of 0.90, Composite Reliability (CR) of 0.92, and Average Variance Extracted (AVE) of 0.71; therefore, all evidence indicates a cohesive and reliable scale.

4. Consumer Engagement

The definition of consumer engagement has been established as being multidimensional, i.e. by behavioral, cognitive and emotional dimensions. This definition follows conceptual models of Hollebeek et al. (2014) and Vivek et al. (2012). The measurement items used to define consumer engagement have been modified to obtain the three distinct dimensions, in the context of interacting with generative AI. For example: behavioral engagement - “I regularly interact with the AI features provided on this platform”, cognitive engagement - “The use of AI features creates a greater level of mental participation when using the platform”, and emotional engagement - “I experience positive emotional responses when interacting with content that was generated by the AI”. The measurement scale has been validated and shown to have high reliability as indicated by the Cronbach's α of 0.91, Composite Reliability (CR) of 0.93 and Average Variance Extracted (AVE) of 0.69. Thus, the credibility of this three-dimensional measure of consumer engagement has been established.

5. Perceived Transparency

To determine users' perceptions of transparency in generative AI, a set of items was developed based on previous studies that examined transparency and explainability in algorithms. Items included statements like, "I can easily understand how the AI produces results," "I can follow the logic behind what the AI suggests," and "The AI is straightforward about how it comes to its conclusions." The scale was found to have very high internal consistency, with a Cronbach alpha of 0.87; good convergent validity with a Composite Reliability of 0.89; and an Average Variance Extracted (AVE) score of 0.66.

Control Variables

As per accepted methodology, the analysis incorporated several of the following demographic and behavioral variables for inclusion as statistical control variables: age; gender; and education level (as well as use frequency of platform). Control variables were used to help identify and isolate the impact(s) of the primary constructs being studied and to reduce any potential confounding variances that may have existed between the conceptual variables of interest.

Statistical Analysis

After analyzing the data in several phases with SPSS 26 and AMOS. Before Confirmatory Factor Analysis had been carried out to determine the validity of the measurement model; afterwards Structural Equation Modelling was employed to evaluate the research proposition (Hypotheses 1-5). The mediation effects are created by the bootstrapping procedure, while the moderation analyses use interaction terms to create the moderation effects. Alternative model specifications were used to provide further evidence of robustness.

Ethical Consideration

Informed consent was obtained from all the participants prior to their participation in this study as required by the ethical principles to which this research is conducted regarding the use of human subjects. Participation was entirely voluntary, no personal identifying information was collected for this project, and all responses are provided anonymously.

RESULTS

Table 1 Sample Characteristics (N=386)

Characteristic	Category	Frequency	Percentage (%)
Gender	Male	204	52.8
	Female	176	45.6
	Other / Prefer not to say	6	1.6
Age	18–25	98	25.4
	26–35	167	43.3
	36–45	83	21.5
	Above 45	38	9.8
Education	Bachelor’s degree	172	44.6
	Master’s degree	141	36.5
	Doctoral / Professional	73	18.9
Frequency of GenAI Use	Daily	214	55.4
	Weekly	129	33.4
	Occasionally	43	11.2

Table 1 presents demographics of the study sample with a total response from 386 individuals. As seen in Table 1, demographics demonstrate 52.8% identified as male while 45.6% identified as female, with 1.6% identifying as other or wanting to not say. The largest groups for age were 26-35 (43.3%), 18-25 (25.4), and 36-45 (21.5); 9.98% were older than 45 years of age. In terms of education level among participants, most respondents have completed their bachelor’s degrees (44.6%) or a master’s degree (36.5%), and 18.9% completed a Doctoral and/or Professional degree. Finally, in terms of usage patterns, over half of the participants (55.4%) reported using generative AI daily.

Table 2 Reliability

Construct	Cronbach’s α	CR	AVE
Generative AI Use	0.86	0.88	0.65
Cognitive Bias Formation	0.88	0.90	0.67
Consumer Trust	0.90	0.92	0.71
Consumer Engagement	0.91	0.93	0.69
Perceived Transparency	0.87	0.89	0.66

As reported in the table, the reliability and construct validity of all measurement instruments used in the research were found to be satisfactory (see Table 2). Specifically, the data supports the conclusion that each instrument has a Cronbach's alpha score of between .86 and .91 and a composite reliability (CR) score of between .88 and .93; thus, indicating high levels of internal consistency. The average variance extracted (AVE) from the individual measurement instruments was found to range from .65 to .71 and exceed the unacceptable value of .50; therefore, the results support the assumption that these measurement instruments completely capture a significant amount of the variance associated with the corresponding latent variables.

Table 3 Validity

Construct	GAI	CB	Trust	Engagement	Transparency
Generative AI Use (GAI)	**0.81**				
Cognitive Bias (CB)	0.54	**0.82**			
Consumer Trust	-0.41	-0.59	**0.84**		
Consumer Engagement	0.46	-0.44	0.63	**0.83**	

Perceived Transparency	0.39	-0.36	0.52	0.48	**0.81**

In Table 3, the results of the validity assessment are displayed, which confirms that the square root of the AVE for each construct (the number in parentheses on the diagonal) are larger than the correlation coefficient of all other constructs when measuring discriminant Validity. Overall, the correlation coefficient between cognitive bias and consumer trust was very negative (-0.59) indicating a very strong negative effect of the cognitive bias on consumer trust, while the correlation coefficient of consumer trust and engagement was very strong (0.63). The remaining coefficients were all moderate and support the distinction between the constructs measured.

Table 4 Structural Path Estimates and Hypothesis Testing

	Path	**β**	**t**	**p**	**Result**
H1	GenAI Use → Cognitive Bias	0.57	9.84	<0.001	Supported
H2	Cognitive Bias → Consumer Trust	-0.61	-10.7	<0.001	Supported
H3	Consumer Trust → Engagement	0.65	11.3	<0.001	Supported

The findings from the structural model, as seen in Table 4, provide overwhelming evidence in favor of the proposed relationships. H1, which stated that Use of Generative AI leads to Cognitive Bias Formation was very well supported ($\beta = 0.57$, $t = 9.84$, $p < 0.001$). H2, which argued that there is a significant inverse correlation between Cognitive Bias and Consumer Trust, was confirmed at an equal level ($\beta = - 0.61$, $t = -10.7$, $p < 0.001$). Lastly, H3, which predicted that there would be a positive correlation between Consumer Trust and Consumer Engagement, was equally corroborated as well ($\beta = 0.65$, $t = 11.3$, $p < 0.001$). Thus, these results suggest that an increase in the use of Generative AI is associated with increases in cognitive bias, which subsequently decreases trust, and that trust is a significant direct predictor of consumer engagement.

Table 5 Mediation Results

Effect	**β**	**95% CI LL**	**UL**	**Result**
Cognitive Bias → Trust → Engagement	−0.40	-0.48	-0.31	Significant

The findings in Table 5 showed that there is a significant negative indirect effect of cognitive bias on consumer engagement, which is mediated by consumer trust ($\beta = -0.40$, 95% CI [-0.48, -0.31]). The fact that the confidence interval does not include zero provides support for the assertion that an erosion of trust is one of the main pathways for how increased cognitive bias leads to decreased engagement with the GenAI platform.

Table 6 Moderation Results

Path	**β**	**t-value**	**p-value**
Cognitive Bias × Transparency → Trust	0.18	3.21	0.001

Table 6 shows the moderation analysis results that indicate perceived transparency is an important buffer against the negative effects of cognitive bias on consumer trust. The interaction term between cognitive bias and perceived transparency was found to be statistically significant ($\beta = 0.18$, $t = 3.21$, $p = 0.001$). This means when consumers perceive a GenAI system as more transparent, then the effect of cognitive bias on consumer trust is reduced in strength. Therefore, greater transparency of the system will act as a shield or mitigate the decline in trust related to consumer vulnerability to automation or confirmation bias.

DISCUSSION

The study focused on the relationship between Generative Artificial Intelligence (GenAI), cognitive biases and how these biases influence trust and engagement in consumers. Findings provide insight into not only how GenAI can affect user behavior but there are psychological processes which facilitate this effect. All hypotheses have been supported by previous research and are presented in detail.

Evidence shows that hypothesis H1 is supported in that the cognitive bias formation of users toward automation systems increases with GenAI users. The results relate to previous research in automation (Parasuraman & Riley, 1997; Logg et al., 2019), indicating a tendency for users to rely too heavily on algorithm-based automation systems, especially in instances where the perceived accuracy of algorithms is very high, and sophistication level is also very high. Unlike conventional forms of decision-making subsystems, GenAI provides users with seamless, human morphologies in the form of written or spoken materials (completely generated by AIs), this further reinforces the user's reliance on the information generated by an AI without appropriate critical evaluation, supporting the extension of the literature regarding the extent to which GenAI assists users in their decision-making. While previous research examined the development of bias associated with the use of AI systems based on rules or predictive models, this empirical study demonstrates that GenAI systems produce heightened levels of cognitive bias by enabling user involvement in content co-creation and interactive persuasion, further supporting recent theoretical works in human-AI interactions (Ullah et al., 2024).

Cognitive bias formation as explained in H2 has been shown to decrease a person's belief in the competence of GenAi systems and has been shown as negative for peoples trust towards gen-ai based products. The evidence is in line with prior literature on Trust, where individuals develop a level of distrust or "stepping back" from trust when they realize their dependence on automating technologies, which feels like they have lost their control of their lives and independence (Lee and See 2004; Dietvorst, et al. 2015). Moreover, the findings in this study provide insight into the way the trust versus automation studies had been collected prior. Some studies have indicated that dependence on AI systems has increased an individual's level of trust in AI in the short term (Castelo et al. 2019). However, this study shows that bias can erode trust over the long term, specifically if the outputs are being interpreted as only reinforcing narrow viewpoints and removing independent thought. Ultimately, there is an inverse relationship for the way GenAi systems can create cognitive dependence on the AI in a way that erodes trust in the AI and its outputs (Saif et al., 2023).

The results provide a strong basis for accepting Hypothesis 3 (H3). This confirms that Trust has a significant, positive impact on Engagement with consumers. The literature supports this finding: both Hollebeek et al. (2014) and Vivek et al. (2012) note that Trust is a key element to fostering Engagement on digital platforms. Consumers are more likely to engage with an AI system (that appears to be Reliable/Ethical) when they feel that they can trust the AI system. Trust promotes additional Cognitive Effort, Emotional Attachment, and Behavioral Participation by consumers. The current study indicates that Engagement in the context of GenAI is not solely a function of Novelty and Personalization, but rather that Trust is a critical component of creating sustainable and meaningful Engagement. As previously discussed, Engagement without Trust is often short-term/superficial.

The analysis of mediation supports H4, showing that consumer trust mediates the connection between the formation of cognitive biases and consumer engagement behavior. The results support theories that place trust at the heart of how cognitive evaluations translate into behaviors (McKnight et al, 2002). More specifically, the findings point to cognitive bias reducing engagement through decreased levels of trust. In other words, cognitive biases may result in initial increases in interaction frequency, but distort users' confidence in the GenAI platform, thus decreasing engagement overall. As such, this finding expands upon the existing engagement theory by demonstrating that engagement based upon biased cognition is not psychologically stable and states that developing GenAI systems to protect and sustain trust is of critical importance.

H5 was also confirmed by the outcomes because of perceived transparency decreasing the effect of cognitive bias development on consumer trust rather than increasing it like many people believe. The current results support previous research on the effects of algorithmic transparency, where the ability to explain an algorithm will result in less discomfort from users and an increased perception of fairness and accountability (Kizilcec, 2016; Schnackenberg & Tomlinson, 2016). In this study, the authors provided the first empirical support for transparency acting as a moderator, expanding the body of research on explainable AI, by showing that transparency can both directly enhance trust and mitigate the negative impact of cognitive bias upon trust. Users are less likely to blindly depend on the outputs generated by GenAI systems if they understand how those outputs are produced, reducing the possibility of having their trust eroded by cognitive bias. Therefore, transparency is an important mediating factor for human interactions with GenAI systems.

Limitations

Several limitations exist in this research which should be highlighted. A first limitation to be addressed is the fact that the research used a cross-sectional study design which makes it difficult to draw accurate causal conclusions about the interactions between the use of GenAI, cognitive biases, trust and engagement; this suggests that future research should use longitudinal or experimental methodologies to observe the influences of different aspects over time. The second limitation is related to the fact that the data were collected using Prolific, an online panel that, although provides high quality responses, may restrict the external validity of the results to other sample groups/cultures. Another limitation to the study's findings was that the research only examined certain cognitive biases (the primary biases being related to confirmation and automation); biases such as anchoring or over-confidence were excluded from study analysis. Lastly, self-reported measures were used in this study; even though procedures were implemented to help alleviate the chances of common method bias, these procedures are not fool proof. Future studies should therefore also consider using system-generated or behaviorally collected data to provide additional evidence of study results.

Implications

GenAIs may have positive effects on user interaction; however, through the cognitive bias of the GenAI, user trust and long-term engagement may be compromised. To mitigate blind dependence and uphold user trust, platform managers should prioritize the development of GenAi's transparent and explainable AI design. From a policy standpoint, the findings inform policymakers that AI governance frameworks that prioritize transparency and user knowledge provide a critical guarantee to ensure responsible and sustainable application of GenAI technologies.

CONCLUSION

The current research focused on use of Generative Artificial Intelligence impacts and users develop cognitive biases and ultimately how these cognitive biases affect users' levels of trust and engagement with digital platforms as consumers. The findings indicate that increased use of GenAI has an exponential increase in cognitive bias for users that results in a decrease in users' overall trust. Furthermore, the study confirms that trust is a major factor in user engagement, supporting the notion that trust is critical for encouraging long-lasting and meaningful user interactions in GenAI-enabled environments. The results suggest that trust serves as a mediating variable between the negative effects of cognitive bias toward GenAI, as well as a decrease in user engagement through a degradation of trust resulting from a reliance on bias toward GenAI. Although there is evidence of a potential relationship between trust and cognitive bias, perceptions of transparency of GenAI systems counter the negative effects of cognitive bias on trust, indicating that the ability of a user to perceive the system's transparency will act as a protective factor. In conclusion, the research suggests that the positive impact of the use of GenAI on user engagement depends on the way in which cognitive bias and trust are handled and that the necessity for designing GenAI systems to be perceived as transparent and trustworthy is essential for the sustainability of user engagement in a responsible manner and the responsible deployment of generative technologies.

REFERENCES

1. Araujo, T. (2020). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior, 85*, 183-189.
2. Arechar, A. A., Allen, J., Berinsky, A. J., Cole, R., Epstein, Z., Garimella, K., ... & Zhang, A. X. (2023). Understanding and evaluating the impact of large language models on the study of human behavior. *Science Advances, 9*(42), eadj4453.
3. Baird, A., & Maruping, L. M. (2021). The next generation of research on IS use: A theoretical framework of delegation to and from agentic IS artifacts. MIS Quarterly, 45(1), 315–341.
4. Borges, A. F., Laurindo, F. J., Spínola, M. M., Gonçalves, R. F., & Mattos, C. A. (2021). The strategic use of artificial intelligence in the digital era: Systematic literature review and future research directions. *International Journal of Information Management, 57*, 102225.
5. Chughtai, M. S., Mushtaque , I., Waqas, H., Raza, H., & Luis Angulo-Cabanillas. (2022). Knowledge hiding behaviors as moderator between machiavellianism, professional envy and research productivity: Empirical evidence from emerging economy. *Knowledge Management & E-Learning: An International Journal*, 510–535.
6. Chung, M., Ko, E., Joung, H., & Kim, S. J. (2020). Chatbot e-service and customer satisfaction regarding luxury brands. *Journal of Business Research, 117*, 587-595.
7. Dwivedi, Y. K., et al. (2023). So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI. International Journal of Information Management, 71, 102642.
8. Felzmann, H., Fosch-Villaronga, E., Lutz, C., & Tamò-Larrieux, A. (2020). Towards transparency by design for artificial intelligence. *Science and Engineering Ethics, 26*(6), 3333-3361.
9. Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM. *MIS Quarterly*.
10. Hollebeek, L. D., Glynn, M. S., & Brodie, R. J. (2014). Consumer brand engagement. *Journal of Interactive Marketing*.
11. Huang, M.-H., & Rust, R. T. (2021). Artificial intelligence in service. Journal of Service Research, 24(1), 3–19.
12. Jacovi, A., Caciularu, A., Goldman, O., & Goldberg, Y. (2024). Reading between the lines: Towards quantifying uncertainty in large language model responses. *Transactions of the Association for Computational Linguistics, 12*, 818-836.
13. Jakesch, M., Hancock, J. T., & Naaman, M. (2023). Human heuristics for AI-generated language are flawed. *Proceedings of the National Academy of Sciences, 120*(11), e2208839120.
14. Kizilcec, R. F. (2016). Transparency effects in AI systems. *CHI Proceedings*.
15. Krämer, N., Zerres, C., & Stürmer, S. (2022). When trust in AI fails: The influence of anthropomorphism and situational privacy concerns on the experience of threat. *Computers in Human Behavior, 137*, 107390.
16. Langer, M., & König, C. J. (2023). Introducing artificial intelligence in organizations: A review and research agenda. Human Resource Management Review, 33(1), 100809.

17. Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation. *Management Science*.
18. Marinova, D., de Ruyter, K., Huang, M.-H., Meuter, M. L., & Challagalla, G. (2023). Getting smart: Learning from technology-empowered frontline interactions. Journal of Marketing, 87(1), 1–22.
19. McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Trust in e-commerce. *Information Systems Research*.
20. OECD. (2023). Trust in artificial intelligence. OECD Publishing.
21. Parasuraman, R., & Riley, V. (1997). Automation bias. *Human Factors*.
22. Rai, A., Constantinides, P., & Sarker, S. (2024). Next-generation digital platforms: Toward human–AI symbiosis. MIS Quarterly, 48(1), 1–26.
23. Saif, S., Tahir, U., Anjum, S., & Jabeen, A. (2023). *Retail brand personality and brand relationship quality as initiators of electronic word of mouth*. NICE Research Journal of Social Sciences, 16(3).
24. Saif, S., Zameer, H., Wang, Y., & Ali, Q. (2024). *The effect of retailer CSR and consumer environmental responsibility on green consumption behaviors: Mediation of environmental concern and customer trust*. Marketing Intelligence & Planning, 42(1), 149–167.
25. Schnackenberg, A. K., & Tomlinson, E. C. (2016). Organizational transparency. *Academy of Management Review*.
26. Shin, D. (2021). The effects of explainability and causability on trust in AI. Telematics and Informatics, 55, 101494.
27. Shin, D., & Park, Y. J. (2023). The role of algorithmic cognition in human–AI interaction. Information, Communication & Society, 26(9), 1853–1872.
28. Siau, K., & Wang, W. (2020). Building trust in artificial intelligence, machine learning, and robotics. CUTTER Business Technology Journal, 33(2), 47–53.
29. Ullah, S., Jianjun, Z., Saif, S., Hayat, K., & Ali, S. (2024). *The influence of corporate social responsibility on impulse buying*. Management Decision, 62(6), 2002–2028.
30. Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). UTAUT2. *MIS Quarterly*.
31. Venkatesh, V., Thong, J. Y. L., & Xu, X. (2022). Consumer acceptance and use of information technology: Extending the unified theory. MIS Quarterly, 46(1), 547–568.
32. Zhang, B., Liao, Q. V., & Bellamy, R. K. E. (2022). Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making. Proceedings of the ACM on Human-Computer Interaction, 6(CSCW2), 1–28.