Issues in Information Systems
Volume 27, Issue 2, pp. 533-550, 2026

DOI: https://doi.org/10.48009/2_iis_2026_136

From transparency to trust: Mitigating perceived AI bias to drive adoption

Chandni Bansal, *O.P. Jindal Global University, cbansal@jgu.edu.in*
Krishan Kumar Pandey, *O.P. Jindal Global University, kkpandey@jgu.edu.in*
Rajni Goel, *Howard University, rgoel@howard.edu*
Anuj Sharma, *O.P. Jindal Global University, anuj.sharma@jgu.edu.in*

Abstract

AI has advanced to the forefront of creating efficiency improvements, innovations and competitive advantages in multiple sectors. Although the application of AI into organizations is advancing rapidly, the implementation of AI in most organizations continues to face numerous obstacles including technical challenges, organizational challenges and ethical challenges. As a result of the increasing role of AI in making decisions, many individuals are concerned about the increased likelihood of issues related to bias, fairness and transparency in AI. Previous studies have examined factors affecting the adoption of new technologies which include performance expectations, amount of effort required to implement an innovation, amount of support received from upper management, preparation of the organization for the innovation, perceived risks associated with the innovation, and external pressures created by regulations or laws. To our knowledge however, very little work has been done to examine the ethics of AI, especially AI transparency and algorithmic bias. Therefore, the current research aims to understand the variables motivating managers to accept AI using two underdeveloped but important concepts - AI transparency and perceived algorithmic bias using Technology Acceptance Theory and trust theory. Data was collected from 320 individuals in IT and management fields and utilized PLS-SEM to evaluate the proposed model. Findings indicated that greater algorithmic clarity enhances user trust and reduces perceived bias, whereas perceived bias negatively impacts trust in AI. Additionally, user trust in AI serves as a mediator between both transparency and perceived bias when considering the intent to adopt. Overall, our results highlight the necessity of developing transparent and trustworthy AI systems that increase manager trust and propensity to utilize AI based tools. The research also contributes to theory and practice by providing a solid conceptual model that incorporates integrity-driven AI into the broader context of trust-based adoption.

Keywords: algorithmic bias, user trust, ethical AI, AI transparency, AI bias, AI adoption, PLS-SEM

Introduction

AI is revolutionizing the way decisions are made in organizations by leveraging human judgment with data driven insights (Charles et al., 2025). While AI holds great promise, however, many remain concerned with issues related to bias, transparency and fairness when it comes to using AI in management (Shin, 2021). Such concerns can ultimately affect trust, which is critical to the eventual acceptance of new technology (Bansal et al., 2023; Venkatesh et al., 2012). Thus, it is very important to identify what factors affect trust and the adoption of AI from a management perspective so as to develop approaches to promote the

responsible utilization of AI in firms. Many times, algorithmic systems act as "black boxes," where decisions are made that are difficult for users to comprehend. The lack of clarity in this regard can lead to apprehensions about bias (i.e., believing that certain groups are treated unfairly), ultimately leading to decreased levels of trust and propensity to adopt (Kizilcec, 2016; Lee, 2018). When users perceive AI systems as both transparent and equitable, they generally view them as capable, friendly, and deserving of trust (Shin, 2020). In addition to the positive effect on attitudes toward adopting an innovation (Gefen et al., 2003; Lankton et al., 2015), trust also affects the willingness to utilize an innovation. Although Technology Acceptance Theory (TAM) remains one of the most widely adopted frameworks for explaining technology adoption, its traditional emphasis on perceived usefulness and perceived ease of use may not sufficiently capture the ethical and socio-cognitive challenges associated with contemporary AI systems. Unlike conventional information technologies, AI systems often operate as complex, data-driven decision-support mechanisms whose adoption depends not only on perceptions of functionality but also on users' confidence in the fairness, transparency, and trustworthiness of algorithmic decisions. Consequently, recent AI adoption research increasingly advocates extending TAM with trust-based and ethical constructs to better explain users' behavioral intentions in organizational settings.

There is a significant shortage of literature exploring how ethical and perceptual concerns relate to one another and ultimately influence the adoption of AI in a managerially relevant context. In this study, a conceptual model is proposed that links together the concepts of algorithmic transparency, perceived bias, user trust, and intentions to use AI. Using Partial Least Squares Structural Equation Modeling (PLS-SEM), it is shown how user trust serves as a mediator, demonstrating that perceptions of both transparency and bias directly affect users' behavioral intentions to adopt AI.

This study adds to the body of knowledge in the area of management research by:

1. Presenting a validated measurement framework linking AI ethics and user behaviours.
2. Presenting empirical evidence demonstrating transparency's indirect role in diminishing bias and increasing trust.
3. Presenting practical recommendations for managers to incorporate ethical AI tenets into organizational practices.

The study fills a significant void in the body of knowledge regarding how ethical perspectives of AI technologies relate to their acceptance in management.

Literature Review and Hypotheses Development

Previous research on AI adoption has taken place extensively in different areas of AI, specifically the technological, organizational and behavioral aspects, however the majority of previous research has focused on the cognitive and functional aspects of AI, such as usability, ease of use, and expectations regarding the performance of the AI system (Murire, 2024; Tursunbayeva & Gal, 2024). On the organizational side of things, research on AI adoption has mainly dealt with leadership support for AI, the preparedness of the organization to implement AI and the proficiency in using AI (Murire, 2024; Tursunbayeva & Gal, 2024). On the consumer side of things, research on the use of AI has mainly dealt with perceived risk of using AI, resistance to new technologies and aversion to algorithms (Gursoy et al., 2019; Longoni et al., 2019). Alkhwaldi et al. (2025) found that performance expectancy, social influence, hedonic motivation, and anthropomorphism shape HR professionals' intentions of adopting AI-powered recruitment tools in their organization. Other studies highlight the role of cognitive flexibility in AI adoption in higher education (Alshammari et al., 2026). Table 1 provides an overview of existing literature on AI adoption.

Table 1. Summary of Prior Research on Factors Influencing AI Adoption

Author(s) & Year	Context / Industry	Sample Size	Method / Analysis	Key Findings
Miao & Photchanachan (2025)	High-tech enterprises; employee perspective; Shanxi Province, China	Quantitative survey (n=500)	SEM using AMOS; reliability/validity via SPSS	TOE factors (infrastructure, top management support, competitive pressure) significantly drive AI adoption intention. The most influential factor in organizations for the intention to adopt AI is top management support.
Nguyen et al. (2022)	Vietnamese organizations across industries; responses from CIOs, IT managers, and senior executives	Mail survey (n=193 valid)	CFA, SEM using AMOS & SmartPLS	Strongest drivers: Technical Compatibility, Organizational Readiness, Managerial Capability. AI adoption is driven more by internal technological and organizational readiness than by firm size or competition.
Abulail et al. (2025)	Higher education institutions across Türkiye, USA, and Canada (students, faculty, staff)	Quantitative survey (n=367)	CB-SEM using SPSS AMOS; CFA	Various technological, organizational, and environmental factors exert a notable impact on the intention to adopt AI in higher education.; user-centric factors and organizational readiness play critical roles, while competitive pressure and government regulations show weaker or non-significant effects in this context.
Ye et al. (2019)	Public acceptance of ophthalmic AI screening devices in primary/community healthcare settings in China (general citizens, not doctors)	Cross-sectional survey (n=474)	CFA and SEM	Subjective norms were identified as the most robust determinant of AI utilization intentions, followed by perceived usefulness, while resistance bias significantly reduced adoption. Findings highlight that social pressure and resistance to change, rather than technology features, primarily drive AI adoption in this collectivist healthcare context.
Liang & Hongtao (2025)	E-commerce SMEs in Shandong Province, China	Quantitative survey (305 SMEs)	Correlation and Multiple Regression Analysis	Technological factors exert the strongest influence on AI adoption attitudes, followed by organizational and environmental factors. AIaaS, management support, skilled workforce, and competitive pressure significantly drive adoption. Major barriers include high cost and skill shortages despite positive attitudes toward AI.
Yang et al. (2024)	Professional service industry - Australian auditing firms (Big, Medium, Small)	15 semi-structured interviews	Qualitative multiple case study; TOE-guided thematic analysis	Six TOE factors influence AI adoption, but their effect varies by firm size. Large firms adopt AI for operational affordances, small firms for marketing affordance. AI

Author(s) & Year	Context / Industry	Sample Size	Method / Analysis	Key Findings
				constraints (bias, black-box, compatibility) inhibit adoption more in small firms due to low AI readiness. Neutral regulations are a barrier only for large firms with deep AI use. Innovation management structure and linking agents strongly determine adoption depth.
Wang et al. (2021)	Higher education in China - university teachers using AI-based teaching applications (e.g., Mosoteach, Smart Class)	Quantitative survey (n=311 teachers)	CFA and SEM using AMOS	Attitude toward AI is the strongest determinant of teachers' intention to use AI. Teachers adopt AI primarily when they feel confident and develop a positive attitude, rather than merely perceiving AI as useful or easy to use.

Nonetheless, the ethical and perceptual aspects of AI, especially regarding algorithmic transparency and perceived bias are still not fully explored in relation to the adoption of AI in management. While some studies (Shin, 2020; Lee, 2018) have started to incorporate fairness and explainability into models of user trust, empirical validation in organizational decision-making is still limited. Furthermore, earlier models typically regard trust as a straightforward precursor to adoption, overlooking how it influences perceptions of transparency and bias. Therefore, this research aims to fill this significant theoretical and empirical gap by using algorithmic transparency and perceived bias as perceptual ethical precursors of user trust and the intent to use AI. This approach expands current models that integrate ethical concepts of AI perception with behavioral adoption models and therefore provides a broader perspective on the ethical use of AI in management settings. The constructs used in the present study will be briefly described below.

AI Transparency

AI transparency relates to how well users can understand the processes, logic and data usage of an AI system (Kovác & Horváth, 2025; Kizilcec, 2016). When an AI system is transparent, it enables stakeholders to understand why certain decisions were made, thus reducing the amount of bias and uncertainty (Daneshjou et al., 2021; Mehendale, 2023). The research indicates that as the level of transparency of an AI system increases, users tend to view the system as being fairer and more accountable than before, which ultimately contributes to higher levels of trust (Kizilcec, 2016). Furthermore, transparency enhances how users perceive the decisions made by AI systems, enabling them to better understand the underlying logic of these decisions (Lee & Cha, 2025). According to Shin (2020), clear AI interfaces increase trust by decreasing ambiguity and the perceived risks associated with interacting with AI systems. Transparency can enable decision-makers to explain suggestions made by automated systems resulting in higher levels of accountability and acceptance of artificial intelligence systems (Wanner et al., 2022; Lee & Cha, 2025). While the importance of transparency in discussions surrounding ethical AI has been emphasized in recent years, relatively few empirical studies have empirically linked perceived transparency to AI system adoption - particularly in the business, financial services and service industries. Consequently, this research posits the subsequent hypotheses:

H1: AI transparency negatively affects the perceived AI algorithmic bias.
H2: AI transparency positively affects user trust in AI systems
H3: AI transparency positively affects the intention to use AI systems.

Perceived AI Bias

Perceived AI/algorithmic bias refers to the users' belief that an AI system's decision-making process and outcomes are unfairly influenced, favouring some groups over others, or functioning without equal treatment (Bansal et al., 2023; Gudis et al., 2023; Mehrabi et al., 2021). Bias perceptions lead to diminished user trust and may significantly lower the likelihood of AI adoption (Ha & Kim, 2024; Lai et al., 2025). To enable the use of AI in managerial contexts, decision-makers must not only trust the accuracy of the AI system, but also its fairness (De Cremer et al., 2024; Matta & Bolli, 2023). On the same note, Bansal (2025) points out that bias mitigation is an effective way to achieve ethics and fairness in AI and thus ensure AI sustainability. Although bias is increasingly acknowledged as one of the most important ethical issues related to AI, there is still a lack of empirical research that connects perceived bias with behavioral adoption models, especially in practical business and service environments. Therefore, this study suggests the following hypotheses:

H4: Perceived AI bias negatively affects user trust in AI systems.

H5: Perceived AI bias negatively affects the intention to use AI systems.

User Trust

User trust pertains to an individual's inclination to depend on an artificial intelligence (AI) system in situations involving ambiguity, susceptibility, or perceived risk. User trust is generally considered to be a critical element in determining whether or not users will accept technology, especially in cases where users do not know how the algorithmic decisions are made (Bedué & Fritzsche, 2022; Daly et al., 2025; Korzyński et al., 2024; Choudhury & Shamszare, 2023). In AI-based contexts, trust is determined not only by traditional variables, such as perceived competence and reliability, but also by ethical and socio-technical variables, including transparency, fairness and perceived bias (Shin, 2021; Araujo et al., 2020). Studies have shown that users are more likely to entrust the decision-making power to an AI system and show greater willingness to accept AI supported services if they perceive the system as fair, understandable and reliable (Bansal et al., 2024; Bansal et al., 2025; Glikson & Woolley, 2020). Conversely, unclear or distorted AI output can diminish trust in AI systems, leading to doubt and lessening the interest in engaging with algorithmic suggestions (Afroogh et al., 2024; Choung et al., 2023; Lee, 2018). Although it is crucial, trust is still not adequately incorporated into empirical AI adoption frameworks in managerial and organizational contexts, emphasizing the necessity to establish it as a fundamental psychological factor in comprehending AI adoption choices. Accordingly, this study does not position trust as a replacement for the fundamental assumptions of TAM. Instead, trust is conceptualized as an extension that complements TAM by capturing the ethical and relational dimensions of AI adoption that traditional technology acceptance variables do not explicitly address. By integrating AI transparency, perceived AI bias, and user trust into the technology acceptance process, the proposed framework extends TAM to provide a more comprehensive explanation of AI adoption intention within organizational contexts. For this reason, current research proposes following hypothesis:

H6: User trust positively affects the intention to use AI systems.

Based on the preceding theoretical arguments, Table 2 summarizes the theoretical constructs, their definitions, supporting literature, and corresponding hypotheses developed in this study.

Table 2. Theoretical Foundations and Hypotheses Development

Construct	Conceptual Definition	Empirical Evidence / Theoretical Rationale	Hypothesis Formulated	Key Supporting Studies
AI Transparency (AIT)	The degree to which an AI system discloses its decision logic, data use, and underlying	AI Transparency reduced AI bias, thereby enhancing user trust in the system, and promotes	H1: AI transparency negatively influences perceived AI bias.	Mehendale (2023); Shin (2020); Kizilcec (2016); Lee & Cha (2025); Wanner et al. (2022)

	processes in a way users can comprehend.	AI adoption intention.	H2: AI transparency positively influences user trust in AI systems. H3: AI transparency positively influences AI adoption intention.	
Perceived AI Bias (AIB)	The extent to which users perceive the AI system as unfair, discriminatory, or systematically favouring certain groups or outcomes.	Perceived bias undermines credibility and damages trust; thereby restricting AI adoption intention.	H4: Perceived AI bias negatively influences user trust in AI systems. H5: Perceived AI bias negatively influences AI adoption intention.	Bansal et al. (2023); Mehrabi et al. (2021); Lai et al. (2025); Ha & Kim (2024); Bansal (2025)
User Trust (UT)	The willingness of users to depend on the AI system with the expectation that it will act competently, reliably, and ethically.	User trust develops on grounds of using transparent AI. This enhanced trust promotes AI adoption intention.	H6: User trust positively influences AI adoption intention.	Bedué & Fritzsche (2022); Shin (2021); Araujo et al. (2020); Bansal et al. (2024); Lee (2018); Choung et al. (2023); Afroogh et al. (2024)
AI Adoption Intention (AAI)	The degree to which users intend to continue or expand their use of AI-based systems in their work or decision-making.	Adoption Intention increases by increased AI transparency and user trust and is hampered by perceived AI biases in the system.	—	Venkatesh et al. (2012); Venkatesh (2022)

Research Methodology

Research Design

This research utilizes a quantitative cross-sectional survey methodology to examine the influence of artificial intelligence transparency and bias on user trust levels and intentions to adopt AI systems. Because of PLS-SEM's capability as both an exploratory and predictive 538odelling technique, it was used for this study with SmartPLS 4. In addition to being a good choice for models that include mediation effects and latent constructs assessed with multiple indicators (Hair et al., 2021), PLS-SEM allows for the simultaneous assessment of the measurement reliability and validity of a model as well as the strength and importance of structural pathways.

To facilitate understanding of the proposed theoretical framework, Figure 1 presents the conceptual structural model depicting all hypothesized relationships (H1-H6) among AI transparency, perceived AI bias, user trust, and AI adoption intention. This conceptual model serves as the basis for the subsequent PLS-SEM analysis.

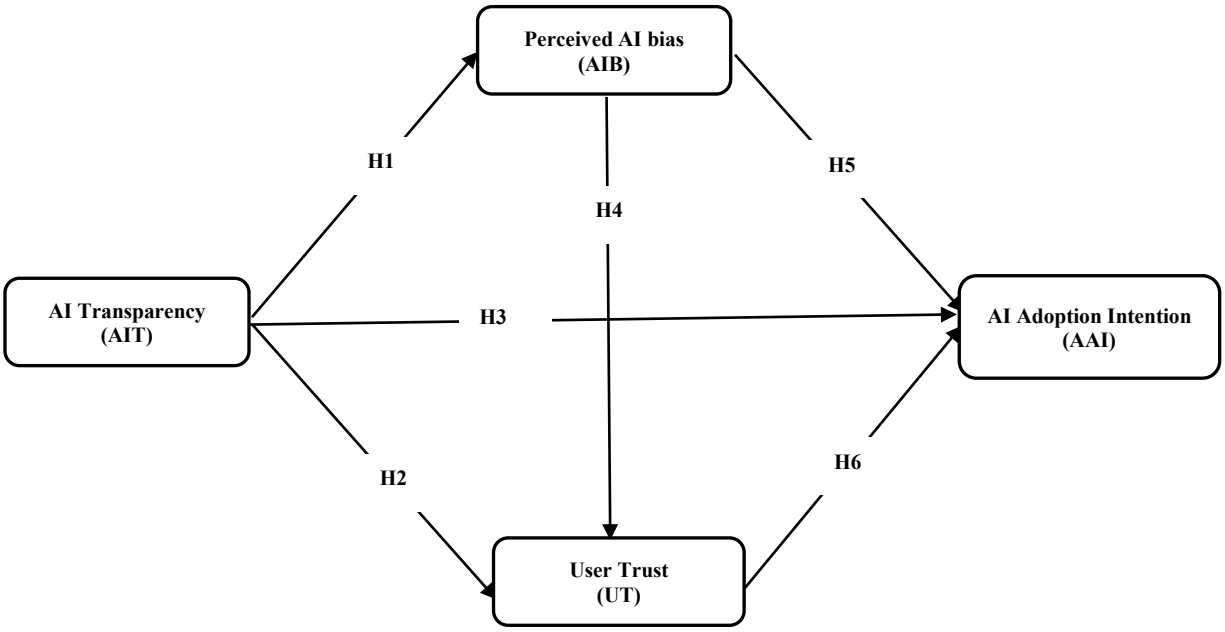

Figure 1. Conceptual Model

Sampling and Data Collection

Data were collected using purposive sampling from professionals employed in IT and management-related functions who possessed prior exposure to AI-enabled applications or decision-support systems in their organizational roles. The term IT and management fields refers to respondents involved in technology-oriented functions (e.g., software development, information systems, digital transformation, and IT services) as well as managerial functions requiring interaction with AI-assisted technologies for operational or strategic decision-making. Consequently, all respondents were considered appropriate informants for evaluating perceptions of AI transparency, perceived AI bias, user trust, and AI adoption intention. The demographics of the participants are listed in Table 3. Of the respondents surveyed, 42% of the participants were female and 58% were male, 47% were between 21 and 30 years old, 55% had a graduate level education, 49% worked in the private sector, and 61% stated they used AI each day.

The target population comprised professionals employed in IT and management-related functions who possessed prior exposure to AI-enabled technologies and have interacted with AI driven systems, which may be classified as recommendation engines, chatbots or predictive analytical tools. Data were collected from participants through an online questionnaire distributed through LinkedIn groups, professional networking sites and through direct e-mail invitation. The purposeful sampling strategy used guaranteed that all participants had at least one prior interaction with an AI system. With the exception of those respondents whose surveys were found to be either partially completed or those who responded to each question in a completely uniform manner, a total of 320 valid responses were received after data cleansing. The number of valid responses collected was sufficient to meet the requirement specified in Hair et al. (2021) of a minimum of 20 observation points per relationship for the 10 times rule, which would result in

a total of 40 observation points per relationship. Having a larger number of participants results in increased statistical power for the mediation analyses and also enhances the model's ability to remain stable.

Table 3. Demographic Details of the Respondents

Variable	Categories	Frequency	Percentage
Gender	Male	186	58%
	Female	134	42%
Age	21–30	150	47%
	31–40	122	38%
	41+	48	15%
Education	Undergraduate	90	28%
	Postgraduate	176	55%
	Doctorate	42	13%
	Others	12	4%
Occupation	Business	64	20%
	Private Sector Job	157	49%
	Government Job	74	23%
	Others	25	8%
AI Usage Frequency	Daily	195	61%
	Weekly	86	27%
	Occasionally	39	12%

The demographic profile indicates that the respondents represented a diverse workforce in terms of age, educational qualifications, and occupations. Since the sampling strategy specifically targeted professionals engaged in IT and management-related roles with exposure to AI-supported organizational processes, the respondents were considered well-positioned to evaluate the study constructs despite the absence of detailed industry-level classifications.

Instrument Development

In order to ascertain the validity of the constructs employed in this study, each construct was operationalized through validated scales utilized in prior research. Responses for each item within the instrument were elicited on a five-point Likert-type scale spanning from "strongly disagree" to "strongly agree." In order to maintain context-specific relevance, especially with regards to AI-based adoption in the service industry, slight linguistic modifications and changes in how items were worded were made to improve clarity. The instrument was reviewed by eight subject matter experts prior to data collection to evaluate the clarity, completeness and context-specific relevance of the instrument. A pilot test was also conducted using 30 participants to validate the reliability of the survey and assess its usability. Each scale exhibited a Cronbach's alpha value greater than 0.80, and a composite reliability (CR) value greater than 0.85, thereby validating the reliability of the internal consistency. The skewness and kurtosis statistics were within the acceptable limits, therefore supporting the assumption of normality for survey data based upon Likert-type scales (see Table 4).

Table 4. Descriptive Statistics of Constructs

Construct	Mean	Std. Deviation	Minimum	Maximum	Skewness	Kurtosis
AI Transparency	3.82	0.74	1.00	5.00	−0.41	−0.32
Perceived AI Bias	2.68	0.81	1.00	5.00	0.36	−0.58
User Trust	3.91	0.69	1.00	5.00	−0.48	−0.27
AI Adoption Intention	4.02	0.66	2.00	5.00	−0.52	−0.21

Data Analysis

A partial least squares structural equation modeling (PLS-SEM) approach was used to analyze the data set. PLS-SEM is well-suited for predictive studies, complex structural frameworks and constructs assessed with multiple indicators. PLS-SEM can provide strong parameter estimates for assessing structural pathways and has a robust framework for addressing issues related to non-normality and complex causal relationships. The PLS-SEM analysis was performed using the SmartPLS software and consisted of a two-stage analytical approach. Stage one entailed assessing the quality of the collected data for missing values, normality and common method bias via Harman's Single-Factor Test (Harman, 1976). In stage two, evaluations were performed on the measurement model to assess the reliability and internal consistency of indicators as represented by Cronbach's Alpha and Composite Reliability. Evaluations were also performed on average variance extracted (AVE) for each indicator to verify convergent validity as well as Fornell-Larcker Criterion (Fornell & Larcker, 1981) and HTMT Ratio for discriminant validity. Following these evaluations, the structural model was analysed to evaluate the hypothesized relationships between variables by examining path coefficients, t-values generated from bootstrapping procedures, and R-squared values representing the percentage of explained variation in each endogenous variable. One strength of PLS-SEM lies in providing high-quality parameter estimates while addressing several potential limitations stemming from non-normal distributions and complex structural models.

Analysis and Results

Measurement Model Evaluation

Before performing structural analyses, the reliability and validity of each reflective construct was evaluated. All items in the survey instrument possessed outer loadings ranging from .768 to .910. Every outer loading value exceeded .70, thereby satisfying the recommended minimum requirement (Hair et al., 2021). Internal consistency reliability for each construct was also evidenced at values greater than .80 based upon Cronbach's Alpha values. Similarly, composite reliabilities (CR) values exhibited a range from .889 to .912. Average Variance Extracted (AVE) values for each construct ranged from .562 to .689 and exceeded .50, thereby evidencing convergent validity per Hair et al. (2021). Discriminant validity amongst constructs was further confirmed by Fornell-Larcker Criterion and HTMT Ratios, wherein each construct demonstrated more variance with their respective indicators than with those of other constructs. HTMT Ratios fell below .85 (Henseler et al., 2015), therefore confirming discriminant validity amongst constructs. Collectively, these findings validated the suitability of the measurement model for subsequent structural analysis. Moreover, as illustrated in Table 5, each construct satisfied requirements for reliability and validity. Further validation for discriminant validity existed in outcomes from HTMT ratios and Fornell-Larcker matrices as provided in Table 6.

Table 5. Evaluation of Reliability and Validity of Measurement Model

Constructs	Items	Factor Loading	Cronbach's alpha	Composite Reliability	AVE
AI Transparency (AIT)	AIT1	0.796	0.835	0.890	0.670
	AIT2	0.818			
	AIT3	0.801			
	AIT4	0.857			
Perceived AI Bias (AIB)	AIB1	0.863	0.833	0.889	0.667
	AIB2	0.768			
	AIB3	0.835			
	AIB4	0.796			
User Trust (UT)	UT1	0.853	0.868	0.910	0.716
	UT2	0.854			
	UT3	0.822			
	UT4	0.855			
AI Adoption Intention (AAI)	AAI1	0.910	0.856	0.912	0.777
	AAI2	0.851			
	AAI3	0.882			

Table 6. Discriminant Validity

Fornell-Larcker Criterion				
Variables	**AAI**	**AIT**	**AIB**	**UT**
AAI	**0.881**			
AIT	0.281	**0.818**		
AIB	0.304	0.444	**0.816**	
UT	0.536	0.408	0.490	**0.846**
HTMT Criterion				
AAI				
AIT	0.332			
AIB	0.354	0.531		
UT	0.620	0.478	0.571	

Structural Model Evaluation

The structural model was assessed using bootstrapping with 5,000 resamples to ascertain the significance of hypothesized relationships. Multicollinearity was not an issue since all VIF values were under 3.0 (Table 7). The coefficient of determination (Table 8) shows that the model explains 20% of the variance in perceived AI bias, 29% in user trust, and 63% in AI adoption intention. These values reflect moderate to substantial explanatory power according to the criteria of Hair et al. (2021).

Table 7. VIF Values of the Model

Outer Model	VIF
AIT1	1.665

AIT2	1.785
AIT3	1.786
AIT4	2.125
AIB1	2.09
AIB2	1.616
AIB3	1.969
AIB4	1.674
UT1	2.264
UT2	2.129
UT3	1.981
UT4	2.187
AAI1	2.521
AAI2	1.909
AAI3	2.22

Table 8. Coefficient of Determination (R^2) for Endogenous Constructs

Endogenous Construct	R^2 Value	Interpretation
Perceived AI Bias	0.20	Moderate explanatory power
User Trust	0.29	Moderate explanatory power
AI Adoption Intention	0.63	Substantial explanatory power

Hypothesis Testing

After confirming the validity and consistency of the variables, the structural model underwent assessment to test he proposed relationships in the hypotheses. A bootstrapped analysis, utilizing 5000 bootstrap samples, was performed to explore both the direct and indirect impacts within the model. The standardized path coefficients, t-values and confidence intervals for each relationship examined in the model are shown in Table 9. The outcomes derived from the structural equation modeling demonstrate a notable adverse impact of transparency in artificial intelligence decision-making on the perception of AI bias ($\beta = –0.444$, $t = 9.662$, $p < 0.001$). Therefore, transparency in AI decision making decreases users' perception of AI bias. Additionally, transparency in AI decision making also has a positive effect on user trust ($\beta = 0.237$, $t = 4.641$, $p < 0.001$), thus providing evidence that when users understand how AI systems function, they are more likely to trust those systems. Furthermore, transparency in AI decision making has a statistically significant effect on AI adoption intention ($\beta = 0.065$, $t = 1.164$, $p < .01$), therefore supporting hypothesis three. Finally, perceived AI bias has a significant negative effect on user trust ($\beta = –0.385$, $t = 7.915$, $p < 0.001$), and therefore users perceive bias as a reduction of their trust in AI systems. Similarly, the negative effect of perceived AI bias on AI adoption intention ($\beta = –0.033$, $t = 1.588$, $p < .01$), indicates that AI bias produces negative feelings toward the system and negatively impacts AI adoption intentions. Finally, user trust has a significant and positive effect on AI adoption intentions ($\beta = 0.493$, $t = 9.914$, $p < 0.001$), therefore indicating that users who have faith in AI systems are more likely to adopt those systems in organizational environments.

Table 9. Structural Model Results

Hypothesis	Relationship	Path Coefficient (β)	t-values	P values (<0.05)	Supported
H1	AIT → AIB	-0.444	9.662	0.000	✓
H2	AIT → UT	0.237	4.641	0.000	✓
H3	AIT → AAI	0.065	1.164	0.001	✓
H4	AIB → UT	-0.385	7.915	0.000	✓
H5	AIB → AAI	-0.033	1.588	0.001	✓
H6	UT → AAI	0.493	9.914	0.000	✓

Discussion

The results of this research provide an increased understanding of the interdependent factors associated with AI transparency, perceived bias, and user trust that influence the use of AI in managerial settings. The findings of this study contribute significantly to the body of literature concerning the adoption of AI and the acceptance of trust-oriented technology. First, the negative and statistically significant relationship between AI transparency and perceived AI bias, provides evidence that when organizations communicate clearly and accurately, why and how an AI system arrived at a decision, users will perceive those decisions as being made fairly and responsibly. Therefore, it is essential to incorporate explainability and traceability into the design of AI systems to foster trust among end-users. Second, the findings support the concept that AI transparency will have a positive effect on user trust; therefore, when an organization provides clear explanations, rationale, and insight into the processes and algorithms that make up its AI system, users will feel less uncertain and psychologically comfortable. This supports previous research which indicates that transparency is a primary mechanism for reducing information asymmetry and increasing user trust in automated decision-making systems. While the prior research demonstrated the value of transparency in increasing user trust in automation, this research demonstrates that transparency is not only a useful design feature but also a necessary prerequisite for developing trust in real-world organizational environments. Third, this research demonstrates that perceived bias in AI will negatively affect user trust; therefore, fairness perceptions play an important role in how users react to the output from algorithms. While the literature has described the concept of algorithmic fairness at length and theoretically, there is a dearth of empirical research on the relationship between user's perceptions of bias and trust in AI systems. Moreover, the results indicate that users' perceptions of even small amounts of bias or inequity in the output from AI systems can cause a substantial reduction in user trust, which can result in a greater reluctance for widespread acceptance of AI systems in both managerial and operational settings. Furthermore, the positive influence that user trust has on the intention to utilize AI systems underscores the importance of developing trustworthy AI environments. Therefore, managers should continually be aware of bias within AI systems, support equitable processes within algorithms and implement and enforce open communications policies regarding the operation of AI tools. The results also provide additional theoretical insight into the significance of trust as a mediating variable linking transparency and bias perceptions to the intention to use AI systems. Specifically, the mediation results demonstrate that user trust mediates the relationship between transparency and adoption as well as bias and adoption; however, the mediation results also show that user trust only partially mediates the relationship between transparency and adoption and between bias and adoption. Therefore, the results of this study suggest that while transparency and bias directly affect the attitudes toward adoption of AI systems among users, the effects of transparency and bias are magnified when mediated by trust. Therefore, these results add to the theoretical discussions on trust as the primary

explanatory variable that connects the characteristics of AI systems to the behaviours related to adoption of users.

In summary, this study contributes to the current conversation regarding user-centric and responsible AI by illustrating that transparency and fairness can serve as practical drivers of adoption as well as basis for consideration regarding ethics.

Implications and Conclusion

Theoretical Implications

This research provides many theoretical contributions to the body of knowledge on AI adoption and technology trust. First, the study demonstrates how the two important components of responsible AI (AI transparency and perceived bias) are interrelated in an empirically tested model demonstrating how system-related traits affect user reactions. Most previous studies have investigated each of these concepts separately; therefore, this study contributes to the literature by examining their combined and comparative impacts on user trust and adoption. Second, this study contributes by extending the traditional technology acceptance theory through the integration of trust theory within an AI-specific context. Unlike conventional technologies, AI adoption involves ethical evaluations concerning transparency and algorithmic bias that influence users' trust before behavioral intentions are formed. By demonstrating the mediating role of user trust between perceived AI bias and AI adoption intention, the proposed framework expands TAM beyond its original technology-centric assumptions and provides a richer explanation of AI adoption behavior in organizational settings. Third, the study contributes to the growing discussion on responsible and ethical AI, by providing empirical evidence that fairness and explainability are not only a matter of ethics and morality, but are also critical variables impacting users' adoption of AI systems. Fourth, by treating bias as a perceptual construct rather than solely as a technical problem, this study extends the conceptualization of AI bias studies beyond fairness in algorithms, to include user perceptions and organizational trust dynamics.

Managerial Implications

Findings of the study are highly relevant to practitioners, managers, developers of AI systems, and policymakers. First, organizations that plan to implement AI systems should focus on transparency enhancing features such as explanation interfaces, justification for decisions, and simplified descriptions of the AI models used. Such design features can positively affect users' confidence in AI systems and increase the reliance of users on the outputs of AI systems. Second, the negative effects of perceived bias highlight the need for continuous auditing for fairness, the use of tools designed to detect bias, and explicit communication to users regarding the collection of data and algorithms, so as to minimize the likelihood of perceived favouritism with regard to decisions made using AI systems. Third, it is important for managers to understand that trust with respect to using AI does not come automatically, and therefore managers need to make a conscious effort to create conditions that encourage users to build confidence in AI. Training programs, live demonstrations and hands-on experiences which help educate the end-users of how the AI systems operate may lead to greater levels of trust (and lower risk) on the part of end-users toward AI. In addition to the technical capabilities of AI, organizations seeking to increase the extent to which they use AI need to consider the views of the users who will employ AI systems. Those individual perspectives greatly impact the outcomes of the adoption of AI systems. Organizations need to view responsible AI practices as a strategy to enable long term adoption of new technologies, as opposed to just viewing them as obligations required under regulations or standards.

Limitations and Future Research Directions

Although this study has provided considerable evidence for its conclusions through extensive empirical findings, there are some limitations that can provide direction for future research endeavours. One major limitation in this study is the cross-sectional nature of the survey methodology employed in this research.

This prevents researchers from drawing causal relationships among concepts. Longitudinal studies would represent a stronger methodological approach to track changes over time in user's attitudes toward trust and bias in AI systems. Another major limitation of this study is that all the data collected were derived from a single organization and thus limits the generalizability of the proposed model. It is expected that subsequent studies employing the proposed model will test it within different industries (i.e., healthcare, finance, education), since these industries have unique forms of manifesting biases in AI. Future research efforts can also expand the proposed model to include other moderators (such as organizational culture, government support of regulation and AI literacy) that can help clarify the contextual factors that affect individuals' behaviour when adopting AI systems.

Conclusion

Given the increased roles of AI in enabling customer interaction with organizations and manager's decision-making, building trust based on transparency and equity with respect to AI systems will become increasingly critical for organizations to successfully implement AI. This research empirically connects users' perception of transparency and bias in AI systems to their usage/acceptance behavior of AI systems, provides theoretical insight into why organizations wish to use AI responsibly, ethically and profitably, and presents practical guidance for organizations wishing to successfully implement AI. Empirical support for the proposed conceptual model was obtained using PLS-SEM methods applied to survey data. The empirical results indicated that users' perceptions of transparency in AI systems reduce their perceptions of bias in AI systems, enhance their trust in AI systems and ultimately facilitate increased adoption of AI systems. Conversely, users' perceptions of bias in AI systems negatively affect users' trust in AI systems, suggesting that fairness and ethical design of AI technologies is a vital requirement. Users' trust in AI systems was determined to be the strongest predictor of users' adoption of AI systems, reinforcing the significance of users' trust in models of technology acceptance. Overall, this research suggests that developing ethically-designed AI systems is not solely a technological problem, but a critical managerial responsibility affecting whether users accept and adopt AI systems, organizations' reputations and organizations' legal legitimacy. The research contributes to the ongoing discussion regarding the responsible and sustainable utilization of AI by illustrating that transparency in designing and communicating about AI is more than simply overcoming technical challenges; it is a strategic management tool capable of influencing users' perceptions, organizations' credibility and long-term trends of adopting AI systems.

References

Abulail, R. N., Badran, O. N., Shkoukani, M. A., & Omeish, F. (2025). Exploring the Factors Influencing AI Adoption Intentions in Higher Education: An Integrated Model of DOI, TOE, and TAM. *Computers, 14*(6), 230. https://doi.org/10.3390/computers14060230

Afroogh, S., Akbari, A., Malone, E., Kargar, M., & Alambeigi, H. (2024). Trust in AI: progress, challenges, and future directions. *Humanities and Social Sciences Communications, 11*(1). https://doi.org/10.1057/s41599-024-04044-8

Alkhwaldi, A. F., Abdulmuhsin, A. A., Abu-AlSondos, I. A., Yaseen, S. G., & Ajouz, M. (2025). Beyond the algorithm: exploring user intentions and anthropomorphic perceptions of AI recruitment tools. In *Studies in systems, decision and control* (pp. 295–312). https://doi.org/10.1007/978-3-031-90271-0_22

Alshammari, S. H., Alkhwaldi, A. F., Alshammari, A. E. A., Alshammari, M. H., & Alrashidi, M. E.

(2026). Extending the UTAUT Model: The role of cognitive flexibility in AI adoption in higher education. *Acta Psychologica*, 263, 106363. https://doi.org/10.1016/j.actpsy.2026.106363

Araujo, T., Helberger, N., Kruikemeier, S., & De Vreese, C. H. (2020). In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI & society, 35*(3), 611-623. https://doi.org/10.1007/s00146-019-00931-w

Bansal, C. (2025). AI ethics and sustainability: Accelerating paradigm shifts toward sustainable development. *Journal of Strategy & Innovation, 36*(1), 200537. https://doi.org/10.1016/j.jsinno.2025.200537

Bansal, C., Kumar, A., Dogra, N., AlZohbi, G., & Prakash, C. (2024). AI chatbots in fintech sector: A study towards technological convergence. *Recent Advances in Computer Science and Communications, 18*(5). https://doi.org/10.2174/0126662558325422240926103718

Bansal, C., Kumar, K., Goel, R., & Sharma, A. (2025). A PLS-SEM mediation analysis of factors influencing the adoption intention of banking chatbots. *Issues in Information Systems, 26*(3). https://doi.org/10.48009/3_iis_2025_2025_116

Bansal, C., Pandey, K. K., Goel, R., Sharma, A., & Jangirala, S. (2023). Artificial intelligence (AI) bias impacts: Classification framework for effective mitigation. *Issues in Information Systems, 24*(4), 367-389. https://doi.org/10.48009/4_iis_2023_128

Bedué, P., & Fritzsche, A. (2022). Can we trust AI? An empirical investigation of trust requirements and guide to successful AI adoption. *Journal of Enterprise Information Management, 35*(2), 530-549. https://doi.org/10.1108/jeim-06-2020-0233

Charles, V., Emrouznejad, A., & Kunz, W. H. (2025). Advancements in Artificial Intelligence-based prescriptive and cognitive analytics for business performance: a special issue editorial. *Journal of business research*, 200, 115576. https://doi.org/10.1016/j.jbusres.2025.115576

Choudhury, A., & Shamszare, H. (2023). Investigating the impact of user trust on the adoption and use of ChatGPT: survey analysis. *Journal of Medical Internet Research*, 25, e47184. https://doi.org/10.2196/47184

Choung, H., David, P., & Ross, A. (2023). Trust in AI and its role in the acceptance of AI technologies. *International Journal of Human–Computer Interaction, 39*(9), 1727-1739. https://doi.org/10.1080/10447318.2022.2050543

Daly, S. J., Wiewiora, A., & Hearn, G. (2025). Shifting attitudes and trust in AI: Influences on organizational AI adoption. *Technological Forecasting and Social Change*, 215, 124108. https://doi.org/10.1016/j.techfore.2025.124108

Daneshjou, R., Smith, M. P., Sun, M. D., Rotemberg, V., & Zou, J. (2021). Lack of transparency and potential bias in artificial intelligence data sets and algorithms: a scoping review. *JAMA*

dermatology, 157(11), 1362-1369. https://doi.org/10.1001/jamadermatol.2021.3129

De Cremer, D., Narayanan, D., Nagpal, M., McGuire, J., & Schweitzer, S. (2024). AI Fairness in Action: A human-computer perspective on AI fairness in organizations and society. *International Journal of Human–Computer Interaction, 40*(1), 1-3. https://doi.org/10.1080/10447318.2023.2273673

Fornell, C., & Larcker, David, F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research, 18*(1), 39–50. https://doi.org/10.1177/002224378101800104

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping: an integrated model. *MIS Quarterly, 27*(1), 51–90.

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of management annals, 14*(2), 627-660. https://doi.org/10.5465/annals.2018.0057

Gudis, D. A., McCoul, E. D., Marino, M. J., & Patel, Z. M. (2023). Avoiding bias in artificial intelligence. In *International Forum of Allergy & Rhinology , 13*(3), 193–195. https://doi.org/10.1002/alr.23129

Gursoy, D., Chi, O. H., Lu, L., & Nunkoo, R. (2019). Consumers acceptance of artificially intelligent (AI) device use in service delivery. *International journal of information management*, 49, 157-169. https://doi.org/10.1016/j.ijinfomgt.2019.03.008

Ha, T., & Kim, S. (2024). Improving trust in AI with mitigating confirmation bias: Effects of explanation type and debiasing strategy for decision-making with explainable AI. *International journal of human–computer interaction, 40*(24), 8562-8573. https://doi.org/10.1080/10447318.2023.2285640

Hair, J. F., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2021). A primer on partial least squares structural equation modeling (PLS-SEM) (3rd ed.). *Sage*.

Harman, H. H. (1976). Modern Factor Analysis (3rd ed.). *University of Chicago Press.*

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

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).

Korzyński, P., Silva, S. C. E., Górska, A. M., & Mazurek, G. (2024). Trust in AI and top management support in generative-AI adoption. *Journal of Computer Information Systems*, 1-15. https://doi.org/10.1080/08874417.2024.2401986

Kovács, E., & Horváth, I. (2025). Trust, transparency, and AI adoption in business. *Acta Polytechnica Hungarica, 22*(6), 81–100. https://doi.org/10.12700/aph.22.6.2025.6.6

Lai, S., Kim, J., Kunievsky, N., Potter, Y., & Evans, J. (2025). Biased AI improves human decision-making but reduces trust. *arXiv preprint arXiv*:2508.09297.

Lankton, N., McKnight, D. H., & Thatcher, J. (2015). Incorporating trust-in-technology into expectation disconfirmation theory. *Journal of Strategic Information Systems, 24*(1), 45–58.

Lee, C., & Cha, K. (2025). Toward the dynamic relationship between AI transparency and trust in AI: A case study on ChatGPT. *International Journal of Human–Computer Interaction, 41*(13), 8086-8103. https://doi.org/10.1080/10447318.2024.2405266

Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big data & society, 5*(1). https://doi.org/10.1177/2053951718756684

Liang, L. Y., & Hongtao, L. (2025). The factors influencing the adoption of AI in e-commerce by SMEs in Shandong province. International *Journal of Research and Innovation in Applied Science, 10*(6), 1268-1288. https://doi.org/10.51584/ijrias.2025.10060096

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research, 46*(4), 629–650. https://doi.org/10.1093/jcr/ucz013

Matta, S. S., & Bolli, M. (2023). Trustworthy ai: Explainability & fairness in large-scale decision systems. *Review of Applied Science and Technology, 2*(04), 54-93.

Mehendale, P. (2023). Mitigating Bias in AI Models through Ethics and Transparency. *Journal of Artificial Intelligence & Cloud Computing, 2*(4), 1-4. https://doi.org/10.47363/jaicc/2023(2)373

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM computing surveys (CSUR), 54*(6), 1-35.

Miao, S., & Photchanachan, S. (2025). The factors influencing artificial intelligence adoption intention in enterprise management. *Journal of Information Systems Engineering and Management, 10*(22s), 1–9. https://doi.org/10.52783/jisem.v10i22s.3466.

Murire, O. T. (2024). Artificial intelligence and its role in shaping organizational work practices and culture. *Administrative Sciences, 14*(12), 316. https://doi.org/10.3390/admsci14120316

Nguyen, T. L., NGUYEN, V. P., & DANG, T. V. D. (2022). Critical factors affecting the adoption of artificial intelligence: An empirical study in Vietnam. *The Journal of Asian Finance, Economics and Business, 9*(5), 225-237.

Shin, D. (2020). User perceptions of algorithmic decisions in the personalized AI system: Perceptual evaluation of fairness, accountability, transparency, and explainability. *Journal of Broadcasting & Electronic Media, 64*(4), 541–565. https://doi.org/10.1080/08838151.2020.1843357

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

Tursunbayeva, A., & Gal, H. C. B. (2024). Adoption of artificial intelligence: A TOP framework-based checklist for digital leaders. *Business Horizons, 67*(4), 357-368. https://doi.org/10.1016/j.bushor.2024.04.006

Venkatesh, V. (2022). Adoption and use of AI tools: a research agenda grounded in UTAUT. *Annals of operations research, 308*(1), 641-652. https://doi.org/10.1007/s10479-020-03918-9

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology (UTAUT2). *MIS Quarterly, 36*(1), 157–178.

Wang, Y., Liu, C., & Tu, Y. F. (2021). Factors affecting the adoption of AI-based applications in higher education. *Educational technology & society, 24*(3), 116-129.

Wanner, J., Herm, L. V., Heinrich, K., & Janiesch, C. (2022). The effect of transparency and trust on intelligent system acceptance: Evidence from a user-based study. *Electronic Markets, 32*(4), 2079-2102. https://doi.org/10.1007/s12525-022-00593-5

Yang, J., Blount, Y., & Amrollahi, A. (2024). Artificial intelligence adoption in a professional service industry: A multiple case study. *Technological Forecasting and Social Change*, 201, 123251. https://doi.org/10.1016/j.techfore.2024.123251

Ye, T., Xue, J., He, M., Gu, J., Lin, H., Xu, B., & Cheng, Y. (2019). Psychosocial factors affecting artificial intelligence adoption in health care in China: cross-sectional study. *Journal of medical Internet research, 21*(10), e14316. https://doi.org/10.2196/14316