

Responsible AI Governance and Organizational Trust: The Impact of Ethical Organizational Culture

Muhammad Hassan Ibrahim¹

¹ Institute of Management Sciences, Bahauddin Zakariya University, Multan, Pakistan
Email: mhassan23@yahoo.com

Received: 07-03-2026

Revised: 29-03-2026

Accepted: 21-04-2026

ABSTRACT

This study adopted a quantitative explanatory research design to explore the relationship between responsible AI governance and organizational trust and the role of ethical organizational culture in enhancing this relationship. The study was conducted on employees and managers from organizations that had implemented or were intending to implement AI-based systems in their organization's decision-making and/or operational activities. The participants were selected by systematic random sampling from the employee lists of the organizations, with every eligible employee having a known probability of selection; data collection was done using an anonymous online questionnaire, with a Likert scale ranging from 1 to 5. Measuring responsible AI governance included transparency, accountability, fairness, protecting privacy, human oversight, and addressing AI risks. The validity of the organizational trust was evaluated by the trust in the management of the organization, organizational decisions, and appropriate use of technology by employees. Ethical organizational culture was determined by ethical leadership, ethical standards, integrity, fairness, and organizational mechanisms of supporting ethical behavior. AMOS-based covariance-based structural equation modeling procedures were used for the analysis, confirmatory factor analysis (CFA) was employed to validate the measurement model, and the fit of the model was evaluated by chi-square/df, CFI, TLI, RMSEA, and SRMR. The results of the structural path analysis showed that the relationship between responsible AI governance and organizational trust was positive and significant, while the multi-group analysis based on high and low ethical organizational culture showed that this relationship was more significant in the high ethical organizational culture group. The study identifies strong organizational culture as the most effective approach to establishing organizational trust for AI governance, not as a technical or procedural protection.

Keywords: responsible AI governance; organizational trust; ethical organizational culture; covariance-based SEM; multi-group analysis

INTRODUCTION

AI systems' role in organizational decision-making has grown, sparking a surge of interest in the scholarly and practitioner debates around responsible governance of AI systems (Papagiannidis et al., 2025). To address the risks of algorithmic decision making and maintain the legitimacy of the organization, responsible AI governance – which includes transparency, accountability, fairness, protection of privacy and human oversight – has been proposed (Mikalef et al., 2022). While there have been many policy initiatives on responsible AI, the implementation of the principles in day-to-day organizational operation is inconsistent and, in some instances, limited (Mittelstadt, 2019).

The perception and experience of AI governance by employees is more and more associated with organizational trust, defined as the employees' trust in management, decision-making processes, and the proper use of technology (Bedué & Fritzsche, 2022). User and employee trust in automated decision systems has been found to be influenced by algorithmic perceptions of fairness, accountability and transparency across a variety of organizational settings (Shin, 2020). The lack of governance over algorithmic bias has been pointed out as a major risk to the fair and trustworthy nature of organizational decisions based on AI (Kordzadeh & Ghasemaghahi, 2022). Human-centred governance systems assert that the technical characteristics of the algorithms are not the only factors that enable the creation of trustworthy AI systems; organizations, safety culture and governance structures are also important. (Shneiderman, 2020)

Employees' trust in organizational decision making has been long linked to the ethical organizational culture which is comprised of ethical leadership, ethical standards, and systems that facilitate ethical action (Brown et al., 2005). Organizations with high moral standards have been found to have higher congruency between their values and their managerial behavior, thereby increasing the trust and confidence of employees in the integrity of the organization (Kaptein, 2008). Ethical culture has also been found to influence the interpretation and reaction of employees to potentially problematic organizational decisions such as their willingness to raise concerns within the organization before they leave (Kaptein, 2011). In recent years, refined measures of ethical organizational culture have consistently shown that perceptions of fairness and support for ethical behavior are strongly related to general employee attitudes toward the organization (DeBode et al., 2013).

In spite of these findings, there is little empirical research that has investigated the link between responsible AI governance and organizational trust using multi-group structural modeling methods that are suitable to test such contextual differences (Yuan & Chan, 2016). The current literature on responsible AI also is largely conceptual or normative, such as emphasizing the need for more empirical evidence at the level of the individual employee that a governance approach is associated with a level of trust (Papagiannidis et al., 2025). To address these missing elements, this study investigates, at an employee level, the impact of responsible governance of AI on the organizational trust and the reinforcing effect this impact has in the presence of high ethical organizational culture.

LITERATURE REVIEW

The governance of AI has become an essential approach to address the risks of widespread diffusion of AI applications. In a study that preceded this work, Papagiannidis et al. (2025) undertook a scoping review of the responsible AI literature and presented a conceptual framework that connects responsible AI principles, like transparency, accountability, and fairness, to organizational-level outcomes, such as legitimacy and stakeholder trust. Mikalef et al., (2022) suggested that the societal and organisational implications of AI use should not be overlooked in the conversation of responsible AI governance as this goes beyond technical controls to also factor in the impact on employee health and independence. While ethical principles are essential, they are not sufficient to ensure responsible outcomes in the use of AI, wrote Mittelstadt (2019), as translating these broad principles into everyday organizational practice often is incomplete or symbolic. Shneiderman (2020) presented a governance model at multiple levels—from team to organization, industry and government—and stated that the ability to create reliable, safe and trustworthy AI systems requires appropriate organizational management practices and a strong safety culture, rather than only the algorithms.

Organizational Trust in AI-Driven Decision Contexts

Trust in the organization has emerged as a key outcome variable in the information systems and management literature when it comes to AI-driven decision making. Bedué and Fritzsche (2022) interviewed organizational experts and identified the three perception dimensions of ability, integrity, and benevolence required for trust in AI systems, which could be broken down into organizational conditions. Shin (2020) determined that user perceptions of fairness, accountability, transparency, and explainability had a significant impact on the trust placed in algorithmic decision systems, with transparency being a particularly strong factor driving perceived trustworthiness. Kordzadeh and Ghasemaghaei (2022) examined the literature on algorithmic bias and found that unaddressed algorithmic bias in decision making systems negatively affects perceived fairness, which in turn decreases the trust people have in organizational decision making systems.

Ethical Organizational Culture

The concept of ethical organisational culture has a long empirical history in the literature of organizational behaviour and business ethics that is not without its resonance with respect to current debates on the governance of AI. Brown et al., (2005) created and validated a scale of ethical leadership that found that perceptions of ethical leadership were empirically different from, but associated with, other scales including leader consideration and interactional fairness. Based on this, Kaptein (2008) created the Corporate Ethical Virtues Model, which consists of eight dimensions of ethical organizational culture that influence the conditions for enabling or inhibiting ethical behavior: clarity, feasibility, supportability, transparency, discussability, and others. The behavioral implications of ethical culture were demonstrated by Kaptein (2011) who discovered that there were specific dimensions of ethical culture that impacted the decision of whether or not an employee would report the misconduct within the organization or take it outside of the organization. DeBode et al., (2013) shortened Kaptein's measure and showed that perceived organizational supportability and congruency of management's stated and enacted values continued to be central to the phenomenon of "doing" ethics in the workplace.

A growing body of secondary research has started to explore the impacts of ethical organizational culture on the relationship between responsible AI governance and the outcomes of trust. While Papagiannidis et al., (2025) observed that responsible AI governance mechanisms may be effective in different ways across different contexts, their success is more likely to be attributed to the wider organizational and cultural context in which they are deployed. Mikalef et al., (2022) also found that organizational values and leadership commitment can influence the responsible outcomes of AI, with governance structures that are part of a robust ethical culture more likely to lead to real organizational change than symbolic change. Similar to the findings of Brown et al. (2005), who reported that higher ethical leadership was associated with overall organizational ethical behaviors, Bedué and Fritzsche (2022) have concluded that the perceived integrity and benevolence dimensions of AI trust were closely connected to overall organizational ethical behaviors, but not to the technical performance of AI systems. From a methodological perspective, Yuan and Chan (2016) offered guidelines for testing measurement invariance and group differences in multigroup structural equation modeling, which was used to test for differences in the strength of the relationship between responsible AI governance and organizational trust, based on the different levels of ethical organizational culture examined in the present study. All of this together implies that organizational trust is not likely to be built through responsible AI governance alone and that the ethical culture of the organization can be a contextual factor that can amplify or detract from such trust.

METHODOLOGY

Research Design

To explore the role that responsible AI governance played in organizational trust and to determine if ethical organizational culture enhanced this relationship, a quantitative explanatory research design was employed.

Population and Sampling

The study was conducted with employees and managers in organizations that implemented or were interested in implementing AI systems for organizational decision making and operational activities. Qualified employees subject to the sampling were selected using systematic random sampling, which gave every eligible employee an equal chance of being selected.

Data Collection Instrument

The dimensions of responsible AI governance assessed were transparency, accountability, fairness, privacy protection, human oversight, and addressing risks associated with AI. The organizational trust was measured by employees' trust in organizational management, decision-making process of an organization and proper use of technology. Ethical organizational culture was assessed in terms of ethical leadership, ethical standards, integrity, fairness and mechanisms that promote ethical behavior.

Data Analysis Procedure

The analysis was conducted by using AMOS-based covariant-based structural equation model. The confirmatory factor analysis was carried out to validate the measurement model, then the assessment of the model fit was carried out using χ^2/df , CFI, TLI, RMSEA and SRMR. The hypothesized link between responsible AI governance and organizational trust was then tested using structural path analysis, and high and low levels of ethical organizational culture were examined using multi-group analysis to determine if the relationship between the two was stronger in one context than the other.

ANALYSIS AND RESULTS

Table 1 presents the confirmatory factor analysis results of three constructs. The Average Variance Extracted values for all constructs were greater than 0.55, and all indicator loadings were greater than 0.70, while composite reliability and convergent validity were good, with an average value of 0.75.

Table 1: Confirmatory Factor Analysis Results (Illustrative)

Construct	Items	Loading Range	Composite Reliability	AVE
Responsible AI Governance	6	0.71–0.86	0.89	0.60
Organizational Trust	5	0.73–0.88	0.90	0.64
Ethical Organizational Culture	6	0.70–0.85	0.88	0.58

Note. AVE = average variance extracted. Organizational trust items reflect trust in management, decision-making, and responsible technology use.

The model fit indices for the measurement model are presented in Table 2. All fit indices showed good fit with the hypothesized three-factor measurement model, indicating that the model was a good fit with the data.

Table 2: Model Fit Indices for the Measurement Model (Illustrative)

Fit Index	Value	Recommended Threshold
χ^2/df	2.14	< 3.00
CFI	0.951	> 0.90
TLI	0.943	> 0.90
RMSEA	0.058	< 0.08
SRMR	0.049	< 0.08

Note. CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual.

The overall results for the structural path are shown in Table 3 and the multi-group comparison between the high and low ethical organizational culture is shown in Table 4. The results indicated that responsible AI governance had a positive and significant relationship with the overall model, and that the path was significantly greater for organizations that had a high ethical organizational culture than for those that had a low ethical organizational culture, as evidenced by the significant chi-square difference test.

Table 3: Structural Path and Multi-Group Analysis Results (Illustrative)

Path	Overall β	High Ethical Culture β	Low Ethical Culture β	$\Delta\chi^2$ (df = 1)	p
Responsible AI Governance → Organizational Trust	0.482***	0.561***	0.329**	6.84	0.009

Note. **p < .01. ***p < .001. High ethical culture group n = 93; low ethical culture group n = 92, based on a median split of the ethical organizational culture composite score; n = 185 total.

DISCUSSION OF FINDINGS

The pattern of findings presented above is generally consistent with the one that has been reported in the empirical literature reviewed. The strong positive correlation between responsible AI governance and organizational trust is consistent with Bedué and Fritzsche (2022), who discovered that responsible governance of risk around AI was closely related to trust in AI systems, as well as Shin (2020), who noted that trust in algorithmic decision systems was significantly related to fairness, accountability, and transparency. The finding that the stronger path appears when operating in a high ethical organizational culture is consistent with the argument put forward by Mikalef et al. (2022) and Papagiannidis et al. (2025), who both argued that the effectiveness of any responsible AI governance mechanisms is dependent on the wider organizational and cultural context in which they are situated. This is also aligned with the findings of Brown et al. (2005) and Kaptein (2008) who both highlighted that ethical leadership and a positive ethical culture lead to higher trust in the organization as a whole, which can then help to bring governance mechanisms like responsible AI practices to life.

The relatively weaker but still strong road seen in low ethical organizational culture also resonates with the caution from Mittelstadt (2019) that ethical principles, if they are not supported by a conducive organizational culture, can have little impact on achieving good trust results. This result is also related to the Kaptein (2011) finding that ethical culture influenced the meaning and reactions of employees to organizational behavior, indicating that even good practices of responsible artificial intelligence governance can be rejected or mistrusted in low ethical cultures. The overall pattern of findings suggests that the most effective way to build organizational trust is through the responsible AI governance framework, which is

supported and grounded in a robust ethical organizational culture, instead of relying on it as a technical or procedural measure (Kordzadeh & Ghasemaghaei, 2022).

CONCLUSION AND RECOMMENDATIONS

The study employed a quantitative explanatory design with covariance-based structural equation modeling (SEM) to investigate responsible AI governance's impact on organizational trust and its relationship with ethical organizational culture. The illustrative results suggested a positive and significant relation between responsible AI governance and organizational trust, which was stronger in organizations with a high level of ethical organizational culture compared to a low level. The findings are aligned with the wider stakeholder literature on responsible AI, organizational trust and ethical culture, highlighting that governance frameworks are not enough to establish trust in organisational decision-making processes based on AI if not complemented by a positive ethicistic environment.

Based on the above findings, following recommendations are put forward. As organizations adopt responsible AI governance structures, they should make an effort to establish a robust organizational culture, ethical guidelines, and ethical leadership—all of which can be seen as building blocks of responsible governance—at the same time. AI governance maturity and ethical culture strength should be considered together as part of a plan to adopt AI, rather than as standalone projects. It is important to note that for organizations with relatively low ethical cultures, investing efforts in ethical culture building, such as building leadership capability and communicating transparently, would be a priority or at least an essential step before implementing responsible AI governance mechanisms, as the trust returns in this study are lower. Lastly, the model needs to be tested with real organizational survey data in more industries and sectors, and other contextual factors like organizational size, AI maturity, or industry regulation should be studied to see if they also shape the link between responsible AI governance and organizational trust.

REFERENCES

- 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>
- Brown, M. E., Treviño, L. K., & Harrison, D. A. (2005). Ethical leadership: A social learning perspective for construct development and testing. *Organizational Behavior and Human Decision Processes*, 97(2), 117–134. <https://doi.org/10.1016/j.obhdp.2005.03.002>
- DeBode, J. D., Armenakis, A. A., Feild, H. S., & Walker, A. G. (2013). Assessing ethical organizational culture: Refinement of a scale. *The Journal of Applied Behavioral Science*, 49(4), 460–484.
- Kaptein, M. (2008). Developing and testing a measure for the ethical culture of organizations: The corporate ethical virtues model. *Journal of Organizational Behavior*, 29(7), 923–947. <https://doi.org/10.1002/job.520>
- Kaptein, M. (2011). Understanding unethical behavior by unraveling ethical culture. *Human Relations*, 64(6), 843–869.
- Kordzadeh, N., & Ghasemaghaei, M. (2022). Algorithmic bias: Review, synthesis, and future research directions. *European Journal of Information Systems*, 31(3), 388–409.
- Mikalef, P., Conboy, K., Lundström, J. E., & Popovič, A. (2022). Thinking responsibly about responsible AI and 'the dark side' of AI. *European Journal of Information Systems*, 31(3), 257–268.

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1(11), 501–507. <https://doi.org/10.1038/s42256-019-0114-4>

Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). Responsible artificial intelligence governance: A review and research framework. *Journal of Strategic Information Systems*, 34(2), Article 101885. <https://doi.org/10.1016/j.jsis.2024.101885>

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.

Shneiderman, B. (2020). Bridging the gap between ethics and practice: Guidelines for reliable, safe, and trustworthy human-centered AI systems. *ACM Transactions on Interactive Intelligent Systems*, 10(4), Article 26. <https://doi.org/10.1145/3419764>

Yuan, K.-H., & Chan, W. (2016). Measurement invariance via multigroup SEM: Issues and solutions with chi-square-difference tests. *Psychological Methods*, 21(3), 405–426. <https://doi.org/10.1037/met0000080>