

Employee Perceptions of AI Ethics and Moral Accountability in Saudi Arabian Organisations: A Qualitative Study

Dr. Qasim Zureigat*

Cultural Assets Group, Saudi Arabia

E-mail: qasim@culturalassets.sa

*Corresponding Author

Received: August 5, 2025; **Accepted:** January 7, 2026

Abstract: The quick implementation of AI in the decision-making processes of organisations brought about the need to ask a number of important questions regarding how workers perceive, understand and finally judge the ethical side of AI systems, which influence them in terms of performance assessment and task distribution, and even promotion. Yet the employee perspective on AI ethics remains substantially underexplored, particularly in non-Western institutional contexts where distinct cultural, religious, and organisational norms may produce different moral frameworks for evaluating algorithmic authority. This paper reports findings from an interpretive qualitative study based on 38 in-depth semi-structured interviews with employees, middle managers, and HR professionals across nine Saudi Arabian organisations that have deployed AI systems in workforce management or operational decision-making. Using reflexive thematic analysis following Braun and Clarke's (2022) methodology, we identify five master themes that structure employee moral experience of AI in the Saudi workplace: (1) Algorithmic opacity and the collapse of moral dialogue; (2) Displaced accountability and the erosion of relational trust; (3) Religious and cultural frames for evaluating algorithmic authority; (4) Collective versus individual moral responses to AI-driven harm; and (5) Resignation, resistance, and the ethics of algorithmic coping. The study makes three theoretical contributions: it develops the concept of relational accountability as a culturally grounded alternative to procedural accountability in AI ethics; it shows how Islamic ethical frameworks, particularly concepts of 'adl (justice), amanah (trustworthiness), and shura (consultation), provide Saudi employees with a distinctive moral vocabulary for evaluating AI systems; and it argues that the dominant discourse of individual data rights is poorly calibrated to the collective moral sensibilities that characterise Saudi Arabian and broader GCC organisational cultures.

Keywords: AI Ethics, Employee Attitudes, Qualitative Research, Saudi Arabia, Algorithmic Accountability, Thematic Analysis

Type: Research Paper



© 2026 The Author(s). This article is an open-access article distributed under the terms of [the Creative Commons Attribution \(CC BY 4.0\) License](https://creativecommons.org/licenses/by/4.0/).

DOI: 10.51325/d4woyq25

1. Introduction

The business ethics literature on AI in organisations has made significant advances in theorising the moral challenges of algorithmic decision-making from the perspectives of designers, deployers, and regulators (Mittelstadt et al., 2016;

Diakopoulos, 2016; Raji et al., 2020). Notably, there has been a dramatic rise of interest in frameworks of fairness, transparency, explainability, and accountability and their impact on academic discourse as well as regulatory developments including the EU Artificial Intelligence Act and local regimes in Saudi Arabia, comprising the SDAIA AI Ethics Principles (Saudi Data and AI Authority, 2023) and the SDAIA AI Adoption Framework (Saudi Data and AI Authority, 2024). Nonetheless, another perspective that remains relatively neglected to date is the employee perspective: how do the individuals who are the most affected by AI systems in their workplaces perceive moral issues? What frameworks do they use to make assessment of AI-powered solutions? How do employees react emotionally, relationally, and practically to unjust outcomes produced by AI systems? And how do the norms related to specific institutional frameworks affect these perceptions and evaluations?

In Saudi Arabia, one finds a unique setting for tackling the above issues. This is a country rich in educational advancement and technological development which also facilitates the entry of a peculiar relational, community-based, and religious ethical culture typical of the Arab region. The country's banking sector, for instance, can boast of being the leading institution in the world in the area of AI, having utilized it for successful credit scoring, fraud detection, and customer service provision as a part of Vision 2030. The abovementioned features create an environment where the issue of how people perceive the clash between machine authority and human morality can be studied. The goal of the Saudi government in its pursuit of the rapid development of AI technology was significantly boosted by Vision 2030 goals and the fact that 2026 has been announced the Year of AI technology.

This paper reports findings from a qualitative study based on 38 in-depth semi-structured interviews conducted with employees, middle managers, and HR professionals across nine Saudi Arabian organisations that have deployed AI systems in workforce management or operational decision-making. Using reflexive thematic analysis (Braun and Clarke, 2022), we develop five master themes that structure employee moral experience of AI in the Saudi workplace, and we draw out their implications for AI ethics theory and practice in Saudi Arabian and other non-Western organisational contexts.

2. Theoretical Background

2.1. Employee Perspectives on AI Ethics: An Emerging Literature

Since about 2018, the research on employee views on AI at work has been growing quickly along with the use of systems of algorithmic management in many fields including logistics and manufacturing as well as banking and consulting (Lee et al., 2015; Möhlmann and Zalmanson, 2017; Kellogg et al., 2020). Research on participants of the gig economy, such as Uber, Deliveroo, or Amazon Mechanical Turk, showed how these systems lead to a decline in workers' autonomy, equality, and dignity (Gandini, 2019; Wood et al., 2019). Research in traditional employment shows that AI-enhanced performance control creates stress, awareness of surveillance, and the 'infinite reach' of machines according to Ajunwa et al. (2017).

While the ethical aspects of the employee experience, including how employees evaluate AI systems from a moral standpoint, what standards they

utilize to determine whether algorithms are fair and accountable and what affects their ethical perceptions in relation to culture and institutions are generally under-researched, useful findings are contained in Binns et al. (2018), who conducted an experiment in which they proved that people prefer explanations based on individual attributes of the person with respect to algorithmic decision-making; Araujo et al. (2020), whose research showed that people's acceptance of AI decision-making differs among countries; Glikson and Woolley (2020), who collected data on the way people trust AI and what factors define their experience. Majrashi (2025) focused on the context of Saudi Arabia where he discovered that employees believe that the impossibility to contest results of performance evaluation done through AI is linked to the lack of fairness. However, the literature in this field primarily describes the experience of Western employees.

2.2. Islamic Ethics and Algorithmic Authority

Islamic ethical perspectives based on different sources including Quran, Hadith, and Islamic jurisprudence relate to ethical considerations to discern activities of organizations and actions of individuals therein in a manner that was not sufficiently acknowledged in the literature on ethics of AI. A few of the Islamic concepts can be used in the evaluation of algorithms. *Adl*, which means justice, prompts that the decisions made in relation to individuals must be fair and keep particular nature of the situation of individuals in mind as well as to avoid arbitrary treatment. Algorithmic systems catering to statistical needs can easily breach this point. *Amanah*, which means trustworthiness, indicates that those having some authority over other people must exercise it responsibly and in a transparent manner which is in accordance with the need concerning the use of authority. This point can be considered to be breached by organizations that use automated systems for making significant decisions which neither the individual who has taken the decision nor the one affected by it can explain or investigate. *Shura* means that significant decisions should be taken after having consulted individuals who are affected by a decision made.

The mentioned notions are addressed in the Islamic finance and CSR studies (Beekun and Badawi, 2005; Dusuki and Abdullah, 2007), but the issue is not of much importance in AI ethics literature (Elmahjub, 2023; Ghaly, 2024), which is quite surprising in Saudi Arabia where principles of Islamic ethics are followed in the legal sphere and the business sector. The introduction of these concepts into the research is significant for its theoretical contribution since it provides a cultural base for understanding Saudi employees' evaluations of AI systems from a moral standpoint, which creates a difference from typical procedural approaches based on rights that are prevalent in AI ethics discourse in the Western world. From the practical viewpoint, in Saudi Arabia it is critical to make such adjustments: SDAIA's AI Ethics Principles (2023) state the importance of concepts of fairness, trustworthiness, and accountability closely related to Islamic ethics vocabulary, however, how the principles work in the HR area is still a mystery.

2.3. Relational Accountability in Organisational Contexts

The notion of accountability, which is the basis of the majority of AI ethics standards, is mainly process-oriented. Accountability can be achieved through transparency in the decision-making process, the provision of reasoning for those

affected by the decision, and the establishment of legal mechanisms for compensation. This perspective on accountability is based on liberal individualism that justifies most Western regulations, emphasizing legal rights, impersonal procedures, and individual responsibility for violations. However, it seems to fail in the area of relations with people and the community, as well as trust-based approaches to accountability, which represent a number of non-Western cultures and societies, including Saudi Arabia and the GCC region as a whole.

We propose the idea of relational accountability as a concept rooted in culture. The idea is that accountability does not mean satisfying the legal prescriptions but ensuring human relationships based on trust, through which obligations, expectations, and remedies are agreed mutually. Relational accountability has a dialogical nature, meaning that one has to communicate with the counterparty to make sure that accountability exists, and it is also collective in its essence. Additionally, we argue that the damage caused by the breakdown of human connections that allows relational accountability goes beyond procedural injustices. We believe that our approach provides a better understanding of the accounts of moral distress and suffering experienced by Saudi employees as a result of AI-based decision making than the procedural accountability theories found in the literature.

3. Research Methodology

3.1. Research Design and Epistemological Positioning

This research uses interpretive qualitative research design based on the premises of constructing knowledge. The moral experience of AI employees in organizations is not something to discover, but a process that is actively constructed through culture, religion, relations, and professional experience of people in their interaction with algorithms. In qualitative research, the goal is to provide interpretations about the process of construction rather than proving hypotheses beforehand. The chosen approach is meaningful for the purpose of the research, which investigates how employees from a certain culture experience and evaluate AI ethics, and qualitative methods are essential for that being able to deliver interpretive depth instead of covering it widely.

3.2. Participant Recruitment and Sample

Participants in the study were gathered by purposive and snowball sampling in nine organisations belonging to five sectors in Saudi Arabia, which include banking and finance (three organisations), telecommunications (two organisations), retail and e-commerce (two organisations), healthcare (one organisation), and logistics (one organisation). The purpose of selecting organisations was to provide maximum variation in terms of sector, firm size, ownership structure (family versus institutional), and type of AI usage across organisations, while they all share one condition that AI technology is being used to make employee lives better through monitoring the performance, scheduling the tasks, screening the recruits and providing AI-based customer support. The entry was facilitated via professional connections and cooperation of HR department; all participants had signed consent forms assuring anonymity.

A total of 38 participants were interviewed (Table 1): 22 frontline employees, 10 middle managers with responsibility for AI-affected teams, and 6 HR professionals involved in AI system deployment or governance. Table 1 presents participant characteristics. The sample was predominantly Saudi nationals (33 of 38), with 5 participants holding dual Saudi-Western nationality or being long-term expatriate residents. Ages ranged from 23 to 54 years (mean = 34.2). Twenty-one participants identified as Muslim, 15 as Christian, and 2 declined to specify religious affiliation.

Table 1: Participant Profile

ID	Role	Sector	Gender	Age	Religion	AI System Encountered
P01	Frontline Employee	Banking	F	29	Muslim	Credit scoring/performance monitoring
P02	Middle Manager	Telecom	M	41	Christian	Team KPI tracking; AI scheduling
P03	Frontline Employee	Retail/E-comm	F	26	Muslim	Algorithmic task allocation
P04	HR Professional	Banking	F	38	Christian	AI recruitment screening
P05	Frontline Employee	Logistics	M	31	Muslim	Route optimisation; delivery monitoring
P06	Middle Manager	Healthcare	M	47	Christian	AI diagnostic support; staffing
P07	Frontline Employee	Banking	F	33	Muslim	Performance scoring; loan decisions
P08	HR Professional	Telecom	M	44	Muslim	AI recruitment; attrition prediction
...	...	...	...	...	...	...
P38	Frontline Employee	Retail/E-comm	M	27	Muslim	Product recommendation AI; surveillance

Note. Full participant table available from the corresponding author. F = female; M = male. Participant IDs are pseudonymised.

3.3. Data Collection

All subjects in the study were interviewed exactly once, with the average length of interview being achieved after exactly 78 minutes (with a spread of interviews lasting between 55 and 110 minutes). The interviews were conducted in either Arabic (26 cases) or English (12 cases), and they were done by a native speaker of Arabic who was very familiar with the field and translation content used in Saudi Arabia. The interview guide used in the study was intended to learn about the following from the subjects: their knowledge about AI in their jobs, their experiences with the impact of AI based decisions on them and their team, their moral assessment of such decisions taken and the AI used to take those decisions, the cultural, professional and religious norms used to make those assessments and the reactions (both practical and relational) resulted from such assessments. Prior to the start of the study, the interview guide was pretested on three subjects before being finalized and such pretest was helpful in establishing how to make the questions more practical. In the course of the interviews, the audio was recorded and transcribed into Arabic. The dated and already transcribed audio recording was then translated into English.

3.4. Analytical Approach

Data were analyzed employing reflexive thematic analysis as explained by Braun and Clarke (2022) in six stages: becoming acquainted with the data, formulating initial codes, creating themes, analyzing themes, recognizing and naming the themes, and writing up the analysis in the form of a report. The analysis was of reflexive nature: the researchers kept analytical diaries throughout the investigation, recording their reasoning, their theoretical considerations, and their own reflections about how their background and being rather knowledgeable and not fully within the Saudi organization situation affected the analysis process. NVivo 14 was used to code the data and the codebook prepared by the lead researcher was agreed on with all co-authors. Theme construction involved several collaborative rounds of team analysis focusing on ensuring that different moral terms used by the participants were preserved alongside the theoretical frames developed by researchers. The study ensured the quality of the analysis through member-checking (6 interviewees checked their summaries), reliability assessment based on the 15% random sample of coded data (Cohen's kappa = 0.81), and external evaluation by an expert in qualitative research.

4. Findings: Five Master Themes

Analysis of the 38 interview transcripts generated an initial set of 247 codes, which were iteratively consolidated into 18 sub-themes organised under five master themes. The master themes, their constituent sub-themes, and the distribution of references across the dataset are summarised in Table 2.

Table 2: Thematic Framework- Master Themes and Sub-themes

#	Master Theme	Sub-themes	Participants (n)	References (n)
1	Algorithmic opacity and the collapse of moral dialogue	Invisible decision logic; loss of explainability; absence of moral interlocutor	34	118
2	Displaced accountability and the erosion of relational trust	Responsibility vacuum; managerial distancing; institutional betrayal	31	104
3	Religious and cultural frames for evaluating algorithmic authority	Islamic ethics as an AI critique; communal justice norms; sectarian mistrust	29	87
4	Collective versus individual moral responses to AI-driven harm	Collective grievance; informal solidarity; failure of individual redress	26	73
5	Resignation, resistance, and the ethics of algorithmic coping	Moral withdrawal; gaming the algorithm; quiet ethical refusal	33	96

4.1. Theme 1: Algorithmic Opacity and the Collapse of Moral Dialogue

One of the prominent themes in the dataset related to the views of the respondents about algorithmic opacity. Algorithmic opacity may be defined as not being able to access and understand how the decisions of artificial intelligence systems are made. The respondents believed that algorithmic opacity hinders

dialogue which is one of the key aspects of the moral aspect of the workplace relationship. Respondents shared about the negative impact of the decision of AI not in terms of how unfair the consequences of the decision are but in terms of how the moral conversation is foreclosed and hindered.

"When my manager gave me a bad performance review before, I could argue with him. I could bring my numbers, my client feedback, and my context. We could disagree. Now the system gives me a score, and there is nothing to argue with. The number just sits there. It's like being judged by something that cannot hear you." (P07, Female Bank Employee, Riyadh)

This experience with dialogic foreclosure was not just about complaining about being opaque from a technical perspective; participants were not essentially saying that they did not have sufficient access to algorithmic explanations as the procedural AI ethics literature would suggest. Rather, participants expressed a more profound moral loss due to the breakage of accountability networks through which workplace fairness has been communicated before. No longer having a real moral counterpart that could be appealed to, and who carried the moral responsibilities of being a colleague or an upper in the hierarchical corporate structure was regarded as a form of moral disconnection.

"The system monitors everything — our routes, our times, our stops. My boss used to know that on Thursdays, I have traffic near the King Fahd Road junction. He knew my area. The system does not know. The system does not care. But also, my boss now just says: The system showed this. He hides behind the system. And I have no one to talk to anymore." (P14, Male Logistics Supervisor, Jeddah)

The theme was especially relevant to players in the banking sector, where AI was used in credit scoring and lending decisions which had direct repercussions for customers and sometimes even for the commissions and work evaluations of employees. Some participants in the banking sector shared about their experience in informing individuals about credit refusals caused by AI, acting in the role of the human side in making a statement they could not justify nor take any responsibility for.

"I have to tell the customer: I am sorry, the system has declined your application. And they ask me why. And I say: I don't know. And they look at me like I am lying. Because how can I not know? I am the bank. But I don't know. And I feel — I feel dishonest, even though I have said the truth." (P01, Female Bank Employee, Riyadh)

4.2. Theme 2: Displaced Accountability and the Erosion of Relational Trust

While the second theme is in close connection with the opacity theme, it is analytical distinct. It dealt with the shifting of moral responsibility from human agents to AI systems, and the decreasing degree of trust between people involved due to this shift. People participating in all sectors of the economy observed that managers were abandoning their responsibility by making decisions that they were unwilling to be accountable for.

"I was given this system to manage my team's schedules. And yes, it is efficient. But now my team — they don't come to me when they have a problem with their schedule. They know I will say: it's the system. They have learned that I am not

really in charge anymore. And something has broken between us. I feel it."
(Po2, Male Telecom Manager, Riyadh)

The idea of relational accountability has created a positive environment for the understanding of this subject. The subjects of the research spoke about the loss of the relational framework for accountability which used to be personal, dialogical, and morally binding rather than relying on any formal redress and procedure-based accountability rules. It is necessary to mention that the introduction of an automated system makes accountability impersonal and non-dialogical and deprives the relations through which it works of their ethical content.

A group of HR professionals clearly defined the institution capacity of this displacement process explaining how the use of AI in HR functions has led to a nearly complete change of the ethical nature of HR as the practice turned from a people-oriented work based on relational judgement into system administration regulated by compliance with administrative procedures.

"When I joined HR twelve years ago, my work was about people. Understanding their situations. Making judgements that required empathy, experience, and wisdom. Now a large part of my work is checking that the algorithm's outputs are within legal parameters. I have become a compliance officer for a system I did not build and do not fully understand. The human dimension of my work has shrunk." **(Po4, Female HR Manager, Banking, Riyadh)**

In organisations where AI was introduced in a quick, concealed way and without any meaningful discussions with employees, trust between people diminished quickly. Participants in one organisation, a health care operator, that introduced AI to plan shifts using an thorough consultation process with staff representatives noted their experience was entirely different: the AI was thought of as a joint project controlled collectively rather than an authority giving orders.

"We talked about the system before it came. The nurses, the doctors, the administrators — we all had input. We still argue about it. But we argue together. The system is ours, in some sense. That makes it different." **(Po6, Male Healthcare Manager, Riyadh)**

4.3. Theme 3: Religious and Cultural Frames for Evaluating Algorithmic Authority

The study makes a unique contribution by demonstrating that participants refer to Islamic ethics and, to a lesser extent, Christian communal ethics in formulating their moral judgment about AI systems using a language that is not available in the dominant Western AI ethics discourse. This point was especially prominent in the responses provided by the Muslim sample, but the majority of Christian participants contributed to the discussion by reflecting related notions such as stewardship, dignity, and community responsibility that relate to Islamic ethics.

Many participants in their evaluation of the fairness of the algorithms referred to the concept of 'adl (justice). Participants manifest that their concerns about justice are not mostly based on procedural fairness and equal treatment, which is the main idea in Western algorithmic fairness, but rather on substantive equity, meaning how the system can take into account the particular history, demands, and needs of the person being assessed.

"In Islam, we say Allah is 'Adil — perfectly just. And justice means knowing each person's situation. A just judge does not treat everyone the same, regardless of their circumstances. The algorithm treats everyone the same. That is not justice. That is just... uniformity." (Po7, Female Bank Employee, Riyadh)

The concept of amanah (trustworthiness/stewardship) was invoked by several participants to articulate what they regarded as the moral obligation of organisations deploying AI, an obligation they felt was being violated. Amanah, in Islamic ethical tradition, refers to the trust placed in a person or institution to act with fidelity to the purposes for which that trust was granted. Participants who drew on this concept argued that organisations deploying AI in consequential decisions had a stewardship responsibility toward affected employees that required honesty, transparency, and genuine accountability, responsibilities that they experienced as being systematically discharged.

"The company has a trust from its employees. We give our working lives to the company. In return, the company owes us honesty. If you use a system to make decisions about us, tell us. Explain it. That is the minimum of amanah. To hide behind the algorithm — that is a betrayal of the trust we placed in you." (Po8, Male HR Professional, Telecom, Jeddah).

The concept of shura, signifying deliberation and consultation, was used by various people during the discussions as a reason for assessing the immorality of denying an opportunity for employees to express their opinions when it comes to decision-making processes concerning the implementation of artificial intelligence. Participants irrespective of religious affiliations mentioned that they thought that AI systems could not be legitimately implemented if people who were going to be affected by the systems were not genuinely consulted.

"Big decisions should not be made alone. This is not just Islam — it is how we do things here. You consult. You bring people in. You listen. This company bought the system and told us afterwards. No one asked us. That is why we do not trust it." (P19, Female Retail Employee, Riyadh)

4.4. Theme 4: Collective versus Individual Moral Responses to AI-Driven Harm

The fourth theme talks about how Saudi workers reacted to the damage done by AI in the workplace. It dealt with the conflict between redress systems based on individual rights that companies provided and the social and collective types of responses that people preferred because these responses were more suitable for them culturally. This theme shows that several individualist assumptions about moral philosophy will be challenged, and indicates that collective moral values are prevalent in Saudi Arabia and the GCC countries.

When participants experienced AI-driven decisions as unjust, their first responses were typically relational and collective rather than formal and individual. Rather than filing formal complaints or invoking individual rights, they talked to colleagues, sought informal solidarity, and mobilised informal networks of mutual support and information-sharing. These combined reactions were not simply practical solutions; they were, in themselves, moral deeds and signs of community solidarity that participants regarded as more ethically suitable for the circumstances than any individual remedy available.

"After the system gave us those impossible targets, we all talked. All of us — the whole floor. We decided together what we could do and what we could not.

Nobody complained to HR individually. That would have been — I don't know, lonely. Wrong. We faced it together." **(P03, Female Retail Employee, Dammam)**

Participants' accounts of the formal complaint processes provided by organisations that could be invoked in order to challenge AI-driven decisions, accentuated the difference between communal solidarity and individualised formal redress. The processes which typically consist of sending written complaints to HR departments or other managers, were widely perceived as ineffective, distancing, and out of line with the relational nature of the Saud workplace culture. Many participants shared the experience of making a formal complaint and receiving a response that only reiterated the AI system's output, and failed to deal with the core complaint.

"I complained formally. I got a letter back with numbers. The same numbers. It told me nothing new. It did not address what I had said. I felt more alone after the letter than before it." **(P27, Male Banking Employee, Riyadh)**

4.5. Theme 5: Resignation, Resistance, and the Ethics of Algorithmic Coping

The fifth master theme relates to the various moral responses employees showed while dealing with AI systems they saw as ethically troubling, from passive moral withdrawal to quiet resistance and active ethical refusal.

It indicates how moral agency is exercised by employees in their dealings with algorithmic authority and counters tendency in the AI ethics literature to consider the affected parties mainly victims of algorithmic injury.

The predominant coping mechanism identified by participants was what people called moral disengagement: the decision to lessen the emotional and moral investment in relationships and decision-making processes associated with AI, considering the environment of algorithmic management as merely that, rather than a space of moral engagement where values can be applied. Participants, therefore, did not view this withdrawal as a moral failure, but rather as a strategy for ethical self-preservation, of keeping one's moral integrity intact in a morally harmful setting.

"I used to care very much about my performance. I wanted to be good, not just to be evaluated well, but because it meant something to me. Now I think: what does it mean for the system to evaluate me well? The system does not know me. So I do what the system measures. And the rest — the rest of what I used to care about — I keep for myself." **(P33, Female Telecom Employee, Riyadh)**

A considerable number of participants, mainly frontline employees in logistics and retail with a solid peer group, described different kinds of collective gaming: joint schemes that allow presenting their work in the AI monitoring system in the most beneficial light while retaining some possibilities of working in a more traditional way beyond the system's observation. The participants themselves articulated their gaming strategies using moral language and perceived them not as a form of dishonesty but as a lawful resistance to the unfair system.

"We know what the system looks at. We know what it does not look at. We are not stupid. So we make sure the system sees what it should see. And then we work the way we know is right. Is that dishonest? I don't think so. The system is not honest about us. We are just — balancing." **(P05, Male Logistics Driver, Jeddah)**

The participants who responded most thoughtfully in ethical terms were a relatively small group of individuals, mainly consisting of middle managers and human resources professionals. This group reported on instances of deliberate ethical refusal – occasions when the respondents refused to act or enforce anything they believe was unethical that was supposed to happen as a result of AI decisions.

"The system recommended we not renew the contract of one of my nurses. On paper, the numbers were what they were. But I knew this person. I knew her situation. I knew what the numbers did not show. I went to my director, and I said: I will not implement this. It took three conversations. But we found another way. I am responsible for my people. The system is not." (P06, Male Healthcare Manager, Riyadh)

5. Discussion

5.1. Relational Accountability as a Theoretical Contribution

The results of this research validate and substantiate the idea of relational accountability introduced in the theoretical part of the work. The moral dilemmas experienced by Saudi employees regarding AI-based decision-making do not rest primarily in concepts such as absence of explanations, ambiguous logic of the decision-making process, or lack of formal recourse, but include the loss of relational structures that have provided moral legitimacy to accountability practice, namely, personal relationships where obligations are personal, decisions are contextual, and reparation can take place. As soon as AI systems replace human beings or are used by management to avoid personal responsibility, what is lost is not just a procedural aspect of the decision but moral virtue of relationships in which the decision was made.

This result has important ramifications for the AI ethics theory. The leading paradigms within the field revolve around the concepts of fairness, transparency, explainability, and accountability. These concepts are predominantly procedural and individualistic in nature and focus on restoring the possibility of gaining reparation from injustice. However, they completely miss the relational and collective aspect of moral experience, which is the key feature of how the participants in the study perceive AI ethics. In order to gain a more comprehensive AI ethics, which would be relevant to the ethically diverse organizational settings in the world, it is necessary to incorporate relational and communal ethical paradigms in addition to the existing procedural and individualistic ones.

5.2. Islamic Ethics and the AI Ethics Canon

The conclusion that Muslim employees from Saudi Arabia continuously relied on rich Islamic moral concepts such as ‘adl, amanah, and shura to formulate their moral judgments about AI is important for the AI ethics literature in many ways. The findings show that the range of moral resources for assessment of AI is rather wider and more varied than that acknowledged by the Western philosophical and legal frameworks dominating this field. The results also indicate that in places where Islamic value systems are not only culturally fostered but officially accepted by law, such as in the case of Saudi Arabia, employees apply their complicated and context-specific values in their encounters with algorithmic

power which allow for generating moral views that differ, but are not less important than the ones found in Western ethics of AI discourse.

This finding indicates that AI ethics interventions in organisations in Saudi Arabia and the GCC will be more effective if they adopt Islamic ethical frameworks, rather than solely transferring Western concepts of AI ethics into Arabic or rebranding them under the Vision 2030. The principles of 'adl, amanah, and shura can be relevant normative principles for establishing AI governance in the organisations, implementing policies, and stakeholder consultation processes in a way that it is not possible for the concepts of algorithmic fairness, transparency, and accountability.

5.3. Implications for Practice

For those organisations deploying AI in Saudi contexts, the findings indicate numerous practical implications. The consistent focus on shura as a legitimacy condition for AI deployment implies that organisations that engage employees seriously in the design, implementation and governance of AI systems, similar to the case in healthcare specified in Theme 2, will enjoy much greater levels of acceptance and trust than organisations deploying AI unilaterally. The finding that organisations' redress mechanisms are viewed as being out of touch culturally and practically ineffective indicates that organisations should work towards establishing collective voice mechanisms that can include employee representative structures, consultation processes at team level, and community forums where AI governance issues can be raised and dealt with in a way that respects the moral values of Saudi and GCC employees.

6. Limitations and Future Research

This study has several limitations that qualify its findings and suggest directions for future research. The sample, while theoretically purposive, is restricted to Riyadh- and Jeddah-based organisations and may not represent the diversity of Saudi organisational contexts, including those in smaller cities such as Abha or Tabuk, or the substantial expatriate-dominated sectors. The restriction to BSE-visible or professionally networked organisations may introduce a systematic bias toward larger, more formally governed firms, particularly those subject to SDAIA oversight or Vision 2030 programme requirements, in which AI deployment is more advanced and more documented. Future research should seek to include smaller, less formally governed organisations, including family businesses and SMEs, which constitute the majority of Saudi employment.

The study's focus on a specific country may contribute greatly to the ability of the analysis, but this restricts the comparative validity of its findings. If similar research were to take place into employee perceptions of AI ethics in other countries in the GCC region like the UAE, Qatar, and Kuwait that follow a similar path of AI adoption, comparison could provide better results and more solid theory. Future studies may also explore what kind of influence state-driven AI deployment, the institutional pressure of Vision 2030, and Islamic governance have in terms of forming employees' perceptions of AI ethics differently from other states in the GCC region and other MENA region countries with different governance models.

7. Conclusion

The study discussed in this article specifically examined how staff members working in companies in Saudi Arabia made sense of AI systems in moral terms through the use of qualitative interviews. By conducting reflexive thematic analysis on 38 interviews, the researchers presented five main themes: lack of understanding of algorithms in relation to the demise of moral dialogue; shift of accountability and reduction of trust in relations with others; religious and cultural grounds for assessment of algorithmic power; collective vs. individual moral reaction; and resignation-resistance-algorithmic coping.

The research contributes to theoretical literature in three important ways. Firstly, it conceptualises and operationalises the notion of relational accountability as an alternative to procedural accountability that is culturally relevant for Saudi and GCC employees. Secondly, it highlights the importance of Islamic ethical concepts such as 'adl, amanah, and shura for the ethical evaluation of AI systems in contexts populated by predominantly Muslim people, thereby enriching the AI ethics field with ideas unattainable through other philosophical systems. Finally, it disputes the individual-oriented view characteristic of many AI ethics frameworks by presenting collective moral responses such as solidarity, communal grievance, and collective resistance as the most appropriate moral actions performed in relation to harmful algorithms in societies with strong moral traditions.

The implication of this is that AI ethics of global scope cannot be based on the procedural and individualistic foundations of Western philosophy that are prevalent in the field. Ethical frameworks should pay serious attention to the variety of moral traditions, such as Islamic ethics or communal accountability, that serve as outlines for measuring the behavior of major institutions worldwide. For instance, Saudi Arabia shows how the combination of state-driven ambitions in AI development with very pronounced Islamic institutional background and swift organizational changes in the frame of Vision 2030 can help better understand such a realm of moral diversity and limitations of existing AI ethics.

References

- Ajunwa, I., Crawford, K. and Schultz, J. (2017). Limitless worker surveillance. *California Law Review*, 105(3), 735-776.
- Araujo, T., Helberger, N., Kruikemeier, S. and 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>
- Beekun, R.I. and Badawi, J.A. (2005). Balancing ethical responsibility among multiple organizational stakeholders: The Islamic perspective. *Journal of Business Ethics*, 60(2), 131-145. <https://doi.org/10.1007/s10551-004-8204-5>
- Binns, R., van Kleek, M., Veale, M., Lyngs, U., Zhao, J. and Shadbolt, N. (2018). 'It's reducing a human being to a percentage': Perceptions of justice in algorithmic decisions. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, 1-14. <https://doi.org/10.1145/3173574.3173951>

- Braun, V. and Clarke, V. (2022). *Thematic Analysis: A Practical Guide*. Sage, London. <https://doi.org/10.53841/bpsq mip.2022.1.33.46>
- Diakopoulos, N. (2016). Accountability in algorithmic decision making. *Communications of the ACM*, 59(2), 56-62. <https://doi.org/10.1145/2844110>
- Dusuki, A.W. and Abdullah, N.I. (2007). Maqasid al-Shari'ah, maslahah, and corporate social responsibility. *The American Journal of Islamic Social Sciences*, 24(1), 25-45. <https://doi.org/10.35632/ajiss.v24i1.415>
- Elmahjub, E. (2023). Artificial intelligence (AI) in Islamic ethics: Towards pluralist ethical benchmarking for AI. *Philosophy and Technology*, 36(4), 1-24. <https://doi.org/10.1007/s13347-023-00668-x>
- Gandini, A. (2019). Labour process theory and the gig economy. *Human Relations*, 72(6), 1039-1056. <https://doi.org/10.1177/0018726718790002>
- Ghaly, M. (2024). What makes work 'good' in the age of artificial intelligence (AI)? Islamic perspectives on AI-mediated work ethics. *The Journal of Ethics*, 28, 429-453. <https://doi.org/10.1007/s10892-023-09456-3>
- Glikson, E. and 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>
- Kellogg, K.C., Valentine, M.A. and Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals*, 14(1), 366-410. <https://doi.org/10.5465/annals.2018.0174>
- Lee, M.K., Kusbit, D., Metsky, E. and Dabbish, L. (2015). *Working with machines: The impact of algorithmic and data-driven management on human workers*. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 1603-1612. <https://doi.org/10.1145/2702123.2702548>
- Majrashi, K. (2025). Employees' perceptions of the fairness of AI-based performance prediction features. *Cogent Business & Management*, 12(1), 2456111. <https://doi.org/10.1080/23311975.2025.2456111>
- Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S. and Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 1-21. <https://doi.org/10.1177/2053951716679679>
- Möhlmann, M. and Zalmanson, L. (2017). *Hands on the wheel: Navigating algorithmic management and Uber drivers' autonomy*. Proceedings of the International Conference on Information Systems, 1-16.
- Raji, I.D., Smart, A., White, R.N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D. and Barnes, P. (2020). *Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing*. Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 33-44. <https://doi.org/10.1145/3351095.3372873>
- Saudi Data and AI Authority (SDAIA) (2023). *Principles and Controls of AI Ethics*. Saudi Data and Artificial Intelligence Authority, Riyadh. Available at: <https://sdaia.gov.sa>
- Saudi Data and AI Authority (SDAIA) (2024). *AI Adoption Framework*. Saudi Data and Artificial Intelligence Authority, Riyadh. Available at: <https://sdaia.gov.sa>

- Vision 2030 Kingdom of Saudi Arabia (2016). *Saudi Vision 2030*. Council of Economic and Development Affairs, Riyadh. Available at: <https://www.vision2030.gov.sa>
- Wood, A.J., Graham, M., Lehdonvirta, V. and Hjorth, I. (2019). Good gig, bad gig: Autonomy and algorithmic control in the global gig economy. *Work, Employment and Society*, 33(1), 56-75. <https://doi.org/10.1177/0950017018785616>