Employees' Justice Perceptions and Trust in AI Systems for Performance Evaluation: Uncovering the Role of Gender and Culture

Tam Nguyen
Dublin City University
Lero - the Science Foundation Ireland
Research Centre for Software
tam.nguyen@dcu.ie

Regina Connolly
Dublin City University
Lero - the Science Foundation Ireland
Research Centre for Software
regina.connolly@dcu.ie

Abstract

Organizational use of Artificial Intelligence (AI) to assist with performance evaluation has increased in recent years. However, limited understanding exists about employees' perceptions regarding AI usage for this purpose, and how it reshapes justice perceptions, trust, and commitment towards the organization. Moreover, little attention has been paid to whether gender and cultural differences exist in relation to this issue. This study explores the intersectionality of gender and culture on employees' perceptions of AI adoption for performance evaluation, focusing on justice perceptions and trust in AI. Survey data collected from 291 participants across Eastern and Western countries was analyzed using partial least squares and multigroup analyses. The findings elucidate the relative impacts of distributive, procedural, and informational justice on trust in AI across cultures. Notably, gender differences in trust in AI formation pathways were observed in the Eastern sample, whereas no such distinction was found among Western participants.

Keywords: AI, organizational justice, trust, gender, culture

1. Introduction

In today's competitive business landscape, organizations are facing significant challenges in attracting and retaining talent. A global survey (PwC, 2023) found that one in four (26%) employees planned to change jobs in the next 12 months. The Great Resignation effect and recent trends like "quiet quitting" (Formica & Sfodera, 2022) underscores that traditional approaches to talent management are no longer sufficient. In light of this, organizations are increasingly turning to Artificial Intelligence (AI)—systems capable of performing tasks requiring human intelligence (Duan et al., 2019)—to assist with talent management for gaining a competitive advantage (Malik et al., 2023). According to the Global AI Adoption Index 2023, 42% of companies have integrated AI into their business operations, with an additional 40% actively exploring the technology (IBM and Morning Consult, 2024). Research indicates that AI can improve talent management in various areas, from streamlining recruitment processes, providing personalized training, enhancing performance evaluation, to promoting workforce diversity (Pereira et al., 2023).

Performance evaluation, which entails activities through which organizations seek to assess employees' performance, enhance competence, and distribute rewards (Fletcher, 2001), is a promising area for leveraging AI benefits. AI can improve employee performance evaluation by providing more accurate metrics, fostering real-time feedback, and assisting managers in making informed evaluation decisions (Malik et al., 2023). Moreover, it can help mitigate human biases and subjective judgements in decision-making, fostering fairer evaluations (Morse et al., 2022). Consequently, organizations are increasingly integrating AI into performance evaluation practices. For instance, IBM (2021) has used AI for providing salary recommendations. MetLife, a leading insurance company, employs AI for giving real-time performance feedback to customer service employees (Roose, 2019).

The growing use of AI in performance evaluation raises concerns about employees' perceived fairness of the system and its impact on their trust and commitment (Bankins et al., 2022). Perceived fairness (i.e., justice[1]) in performance evaluation is particularly important as it affects employees' job satisfaction, trust, commitment, and turnover intentions (Colquitt et al., 2013). When AI is integrated, fairness concerns

[1] We use the terms "fairness" and "justice" interchangeably in this paper.

URI: https://hdl.handle.net/10125/109657
978-0-9981331-8-8

intensify due to the system's inherent challenges. Specifically, although AI algorithms can theoretically remove bias, decisions made by AI are often seen as less fair than human decisions (Feldkamp et al., 2023; Lee, 2018). Additionally, research found that AI can discriminate against specific groups like women and ethnic minorities (Starke et al., 2022). For instance, Amazon recently discovered that its AI-powered recruitment engine exhibited bias against female applicants compared to male applicants (Dastin, 2018). Without addressing concerns of perceived fairness, organizations cannot fully leverage AI's benefits for performance evaluation.

In addition to perceived fairness, trust in AI is crucial for system acceptance and adoption (Glikson & Woolley, 2020). Trust in AI is defined as "the willingness to be vulnerable to the actions of an AI decision aid, based on the expectation that it will perform a decision-making task important to the trustor" (Solberg et al., 2022, p. 195). Prior studies indicate that individuals exhibit lower trust in AI decisions compared to human decisions, especially in tasks involving ethical and moral judgment (Bigman & Gray, 2018) or subjective evaluation (Lee, 2018). Thus, understanding how trust in AI can be developed is critically important.

Despite significant implications for employee satisfaction and retention, there is a deficit of research on how organizational use of AI influences employee's fairness perceptions and trust. Our research addresses this gap by providing an examination of fairness, trust, and their associated outcomes regarding the use of AI for performance evaluation. We further uncover the influence of culture and gender on these relationships. Accordingly, the study contributes to literature in a number of distinctive ways.

First, the study advances our understanding of trust in AI formation through social exchange (Cropanzano et al., 2017), psychological contract (Rousseau, 1989) and organizational justice (Colquitt et al., 2013) lenses. Previous studies have largely focused on trust formation via technical features and user characteristics (Bach et al., 2024), whereas there is a paucity of studies investigating the influence of organizational behavioral factors. In this study, we theorize that trust in AI can be developed through organizational justice perceptions. Moreover, we explore how these perceptions and trust in AI impact trust in organization, trust in managers, and affective commitment. This examination provides insights into the psychological processes employees engage in when AI is implemented in organizations.

Second, our study complements the existing literature by investigating fairness perceptions and trust in AI specifically within the context of performance evaluation. While research on fairness and trust in AI has grown rapidly, there remain limited studies in high-stakes decision-making contexts, such as employment or criminal justice (Langer et al., 2022). Additionally, existing studies that examine employment context predominantly focused on the AI usage in hiring and recruitment (e.g., Feldkamp et al., 2023; Yu et al., 2023). While these studies offer valuable insights, their findings are constrained to those contexts. Our study expands this focus to AI-based performance evaluation, a domain with significant implications for individual and organizational growth (Denisi & Murphy, 2017).

Third, this study employs an intersectionality approach to explore the effects of gender and culture on AI perceptions. Specifically, we first compare the interrelationships among perceived fairness, trust and related outcomes in the context of AI-based performance evaluation between Eastern and Western cultures, then examine gender differences within each culture. Understanding gender and cultural variances is critical to organizations seeking to implement AI, as it can help them tailor their efforts to better align with diverse beliefs, values, and expectations of their employees, thereby enhancing acceptance, trust, and the overall effectiveness of technological integration (Leidner & Kayworth, 2006). Indeed, gender has been shown to influence technology adoption (Trauth, 2011) and trust in AI (Kaplan et al., 2021). Similarly, culture—the shared basic assumptions, values, and artifacts within a social group such as a country, region, or organization—has emerged as a critical factor that shapes user attitudes toward technology implementation (Leidner & Kayworth, 2006). Researchers also found that cultural background can affect trust in general technology (Vance et al., 2008) and trust in AI (Kaplan et al., 2021). In this study, we seek to explore the differences in AI perceptions between Eastern and Western cultures because their cultural characteristics, including power distance, uncertainty avoidance, individualism-collectivism, and masculinity-femininity, are significantly different (Hofstede, 2011). Furthermore, prior research presented empirical evidence highlighting the differences between Eastern and Western countries in the adoption of various technologies, such as mobile recommender system (Choi et al., 2014), enterprise system (Vos & Boonstra, 2022), and AI assistants (Alnefaie & Kang, 2024). By employing the intersectionality approach to understand the influence of gender and culture, we aim to uncover the nuanced variances in employees' perceptions of AI use (Joshi, 2022).

This study is structured as follows. First, we outline the theoretical background and hypotheses. Then, we describe the research methodology, followed by a description of the findings. Finally, we discuss theoretical and practical implications, study limitations, and future research directions.

2. Theoretical background

This study examines employees' perceptions of AI in performance evaluation, focusing on the relationships among fairness perceptions, trust, and affective commitment. To that end, we employ social exchange, organizational justice, and psychological contract theories, which are among the most widely used conceptual perspectives to explain employees' behaviors (Colquitt et al., 2013; Cropanzano et al., 2017; Rousseau et al., 2018). These theories provide a theoretical foundation for understanding how AI integration into performance evaluation influences employee attitudes and behaviors.

2.1 Social exchange theory

Social exchange theory (SET) is a multidisciplinary paradigm that conceptualizes social behavior as an exchange process in which individuals assess relationships in terms of their benefits and risks (Cropanzano et al., 2017). The underlying process of social exchange relationships relies on the norm of reciprocity, whereby individuals tend to reciprocate the positive or negative behaviors of another party (Gouldner, 1960). For example, in the organizational context, SET predicts if the organization provides organizational support or fair treatment, employees will reciprocate with positive behaviors, such as increased trust and commitment (Cropanzano et al., 2007). SET therefore provides a theoretical framework for exploring the relationships among justice perceptions, trust, and affective commitment for this study. More specifically, SET suggests that when employees perceive the performance evaluation using AI as fair, they tend to reciprocate with increased trust and commitment to the organization.

2.2 Organizational justice theory

Organizational justice theory is leveraged to examine the impact of fairness perceptions on trust in AI. Organizational justice refers to employees' perceptions of fairness within the organization, and is commonly divided into three distinct types, namely, distributive, procedural, and interactional justice (Colquitt, 2001). *Distributive justice* is concerned with the perceived fairness of outcomes or final decisions. It is evaluated based on the perceived balance between evaluation outcomes (e.g., performance rating, allocation of compensation, rewards and benefits, or promotion decisions) and the employee's estimated contributions. *Procedural justice* refers to the perceived fairness of the processes leading to outcomes. It is assessed through (a) the adherence to fair process criteria, such as consistency, bias suppression, correctability, representation, accuracy, and ethicality, and (b) the provision of opportunity for employees' voice and input over the decision-making processes and outcomes. *Interactional justice* reflects the perceived fairness of interpersonal treatment that employees receive during the performance evaluation process. It can be further divided into two components: *interpersonal* and *informational justice*. Whereas interpersonal justice refers to the degree of treating employees with respect and dignity, informational justice refers to the provision of information and explanations during the procedure.

Research indicates that all three types of justice—distributive, procedural, and interactional—positively influence employees' satisfaction with performance evaluation (Thurston & McNall, 2010) and important behavioral outcomes, including trust, commitment, job satisfaction, organizational citizenship behavior, and turnover intentions (Cropanzano et al., 2007). While these justice types are linked to similar outcomes, they are conceptually distinct and exert differential impacts on the outcomes (Cohen-Charash & Spector, 2001). Accordingly, we investigate the effects of distributive, procedural, and informational justice perceptions separately in this study. Interpersonal justice was not included because it emphasizes interpersonal interactions, which is not applicable in the context of AI usage for performance evaluation.

2.3 Psychological contract theory

The psychological contract is defined as "an individual's belief in mutual obligations between that person and another party such as an employer" (Rousseau & Tijoriwala, 1998, p. 679). It entails the evaluation of the inducements (e.g., supervisory support, compensation, or rewards) promised by an employer compared to those delivered (Lambert et al., 2003). This theory is deeply influenced by social exchange theory (SET), emphasizing the norm of reciprocity (Coyle-Shapiro & Conway, 2005). However, psychological contract theory places a greater emphasis on promise fulfillment, whereas SET regards reciprocation as unspecified obligations, without requiring the equivalence of return (Blau, 1964). In a psychological contract, consistency between what is promised (or understood) and what is received is a critical issue (Shore & Tetrick, 1994).

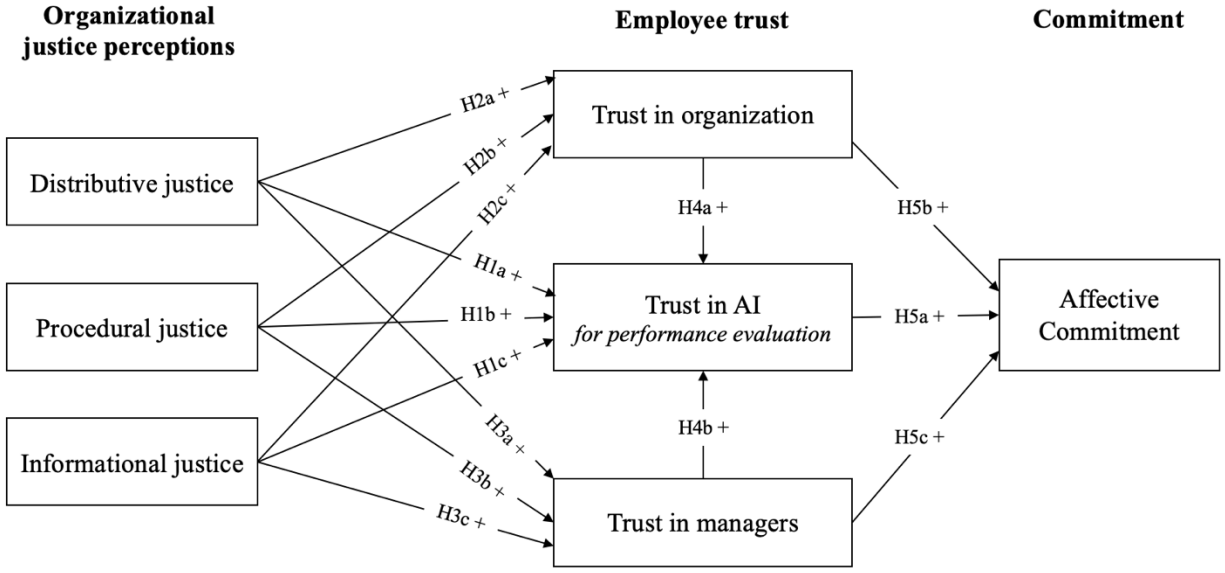

Figure 1. Proposed research model

The psychological contract theory has been extensively utilized to examine the reciprocation of organizational inducements in fostering employees' trust and commitment (Lambert et al., 2020). It suggests when employees perceive the comparison between promised and delivered inducements as favorable, they increase trust and commitment. We anticipate this relationship to hold true in the context of employees' perceptions regarding the AI usage for performance evaluation.

3. Proposed model and hypotheses

The research model is shown in Figure 1. It proposes that perceptions of distributive justice (i.e., fairness of outcomes), procedural justice (i.e., fairness of procedure), and informational justice (i.e., fairness in providing relevant information) influence employee trust in AI for performance evaluation. Furthermore, trust in AI is influenced by trust in organization and trust in manager, both of which are also affected by three types of justice perceptions. Those trust-related constructs subsequently impact affective commitment.

3.1 Organizational justice perceptions

Social exchange, organizational justice, and psychological contract theories consistently suggest that when employees perceive fair treatment from the organization, they tend to reciprocate with greater trust in their employers (Cropanzano et al., 2017). Previous studies provide substantial empirical evidence on the relationships between justice perceptions and trust (e.g., Jiang et al., 2017; Tremblay et al., 2010). Based on these findings, we expect that employees' perceptions of distributive, procedural, and informational justice influence trust in AI. It is thus proposed:

H1: In the AI-based performance evaluation context, distributive justice (H1a), procedural justice (H1b) and informational justice (H1c) perceptions are positively related to employee trust in AI.

Consistent with previous research, we also posit that organizational justice perceptions regarding AI-based performance evaluation influence trust in organization and trust in managers. Hence, we propose:

H2: In the AI-based performance evaluation context, distributive justice (H2a), procedural justice (H2b) and informational justice (H2c) perceptions are positively related to trust in organization.

H3: In the AI-based performance evaluation context, distributive justice (H3a), procedural justice (H3b) and informational justice (H3c) perceptions are positively related to trust in managers.

3.2 Trust in AI, trust in organization, and trust in managers

As AI for performance evaluation serves as a decision support for managers within the organization, we explore whether trust in organization and managers influences trust in AI. The relationship between trust in technology and trust in the employing organization has not been thoroughly explored in the existing literature. Only a number of studies on e-government adoption suggests trust in government affects trust in

the systems (e.g., Alzahrani et al., 2017). In the context of AI, van der Werff et al. (2021) found that employee trust in AI is influenced by the perception of the organization's trustworthiness. This study posits that employee trust in AI for performance evaluation is intertwined with trust in the organization and managers. Employees who possess higher levels of trust in their organization and managers may be more inclined to trust the AI system for performance evaluation. The following hypothesis is thus presented:

H4: In the AI-based performance evaluation context, employee trust in organization (H4a) and trust in managers (H4b) are positively related to trust in AI.

3.3 Affective commitment

Commitment is defined as a psychological state that "has implications for the decision to continue or discontinue membership in the organization" (Meyer & Allen, 1991, p. 67). Affective commitment specifically refers to the employee's emotional attachment or identification with their organization (Meyer & Allen, 1991). Social exchange and psychological contract theories suggest that, in exchange for fair treatment received from their organization, an employee may reciprocate with increased trust, which, in turn, leads to higher affective commitment. Prior research on the link between trust and affective commitment has yielded consistent results (Aryee et al., 2002; Jiang et al., 2017). Accordingly, the following hypothesis is presented:

H5: In the AI-based performance evaluation context, employee trust in AI (H5a), trust in organization (H5b), and trust in managers (H5c) are positively related to affective commitment.

4. Methodology

4.1 Data collection

This study employed a quantitative survey to test the proposed model. The survey was distributed between October 2023 and February 2024. As the study aims to capture individuals' perspectives from diverse backgrounds across Eastern and Western countries, various recruitment strategies were utilized. First, the survey was advertised on social media platforms including LinkedIn, Facebook, and X. Additionally, email invitations were sent to individuals working in different industries through the researchers' contacts. As the research team includes members from both Asian and European countries, these approaches helped yield 310 responses from both regions. For the Eastern sample, additional data were collected during a member's research visits to Thailand and Vietnam. In total, 345 responses were received. After eliminating those with a high level of missing data, 291 valid responses remained. The Eastern sample includes 140 participants from six countries, including India, China, Vietnam, Thailand, the Philippines, and Singapore. The Western sample includes 151 participants from 12 countries: Ireland, the United Kingdom, the Netherlands, Germany, Italy, France, Poland, Finland, Spain, Turkey, the United States and Canada.

Demographic	Profile	Eastern (%)	Western (%)
Gender	Female	58	58.9
	Male	42	41.1
Age (years)	Under 20	2.4	3.3
	21-30	46.8	49
	31-40	34.5	28.5
	41-50	10.6	12.6
	51-60	4.3	5.3
	Above 60	1.4	1.3
Education	Second-level education	3.6	6
	Undergraduate Degree	25.7	28.5
	Master's Degree	57.9	52.3
	Doctoral/PhD Degree	9.3	11.9
	Others	3.6	1.3
Sector	Public	34.3	40.3
	Private	62.9	56.4
	Charitable/Not for profit	2.9	3.4
Industry	Technology	23.4	22.8
	Education	18.9	22.8
	Professional Services	9.5	5.4
	Healthcare/Medical	9.5	4
	Financial Services/Banking	7.3	8.1
	Marketing	6.6	8.1
	Retail/Consumer Goods	3.6	7.4
	Others	21.2	21.4
Job position	Entry-level employees	14.8	14.1
	Junior associate	23.7	18.8
	Senior associate	16.3	16.1
	Managers/Directors	14.7	16.8
	Self-employed	2.2	1.3
	Student	17	20.1
	Others	11.3	12.8

Table 1. Demographic characteristics

4.2 Sample characteristics

Table 1 presents the demographic profiles of both Eastern and Western samples. The two samples share similarities in the distribution of gender, age, education, and occupational profiles. Specifically, both samples have a comparable gender distribution, with females comprising 58% of the Eastern sample and 58.9% of the Western sample. Almost half of the participants were aged between 21-30 years. The majority (above 92%) of participants were educated to at least college level. Concerning employment profiles, over half of the participants worked in the private sector. The participants represented diverse industries, with Technology and Education being the most prominent. About one-third of the participants reported being at entry and junior levels, while the remaining participants held higher-level positions as senior associates and managers or directors. These demographic similarities provide support for the comparability of the Eastern and Western samples.

4.3 Measures

We measured procedural, distributive, and informational justice regarding the organizational use of AI for performance evaluation by adapting items from Colquitt (2001). *Distributive justice* was assessed using a three-item scale, including items such as "The outcomes of using AI to evaluate performance would be fair". *Procedural justice* was measured with seven-item scale, including items such as "The procedures for using AI in performance evaluation would be free from bias". *Informational justice* was assessed using a four-item scale, including items such as "My employer would provide clear information regarding their use of AI for performance evaluation". *Trust in AI* was measured using four items adapted from Höddinghaus et al. (2021) and Vance et al. (2008). Example items include "AI can be trusted to provide accurate information to guide decisions about performance evaluation", and "I believe that AI is effective in supporting performance evaluation decisions". The items for *trust in organization* and *trust in managers* were drawn from Robinson (1996), including items such as "My organisation has high integrity towards its members", and "I trust senior management of my organisation". *Affective commitment* was measured using five items based on Meyer & Allen (1991), including items such as "I feel emotionally attached to my current organisation". All measures were assessed with a five-point Likert scale ranging from 1 = "strongly disagree" to 5 = "strongly agree." We also include gender, age, and education as control variables. The survey was pretested with expert researchers and a convenience sample of participants to ensure appropriateness and relevance.

4.4 Data analysis

The research model was tested for each sample (i.e., Eastern and Western), using partial least squares structural equation modelling (PLS-SEM) as implemented in SmartPLS software version 4. Following the research model assessment, we conducted a multigroup analysis (MGA) for each sample to discover the effect of gender on the proposed relationships within each culture. Power analysis using G*Power 3.1 software (Faul et al., 2009) indicated that the current sample size of both Eastern (n=140) and Western (n=151) samples was sufficient to detect a medium effect size of $f = 0.25$ (Cohen, 1988) with 80% power and a significance level of 5 % for each subgroup.

5. Results

5.1 Measurement model

Factor loadings for most construct items were higher than 0.70 in both samples. Several items were dropped due to low loadings, including PJ1, PJ2, and PJ3 from procedural justice, and AC5 from affective commitment. After removing these items, all items loaded at greater than 0.708 on their intended constructs, which is considered excellent (Hair et al., 2019). We then examine reliability, convergent validity and discriminant validity of constructs in both samples. As shown in Table 2, Cronbach's alpha (CA) and composite reliability (CR) values were all above 0.70, indicating satisfactory internal consistency reliability (Hair et al., 2019). The average variance extracted (AVE) of all constructs exceeded the suggested value of 0.5 minimum, confirming convergent validity (Fornell & Larcker, 1981). For discriminant validity, the heterotrait–monotrait (HTMT) ratio of correlation values of all the constructs were below the recommended threshold of 0.85 (Henseler et al., 2015). Variance inflation factor (VIF) values were all below 3.3, meaning multicollinearity did not exert a biasing influence on the results (Hair et al., 2019).

	Construct	CA	CR	AVE	1	2	3	4	5	6	7
Eastern (n=140)	1. Distributive justice	0.703	0.725	0.631							
	2. Informational justice	0.913	0.914	0.794	0.360						
	3. Procedural justice	0.778	0.797	0.603	0.804	0.603					
	4. Trust in AI	0.838	0.845	0.675	0.871	0.427	0.799				
	5. Trust in managers	0.902	0.904	0.773	0.178	0.503	0.273	0.268			
	6. Trust in organization	0.870	0.870	0.794	0.108	0.483	0.261	0.289	0.757		
	7. Affective commitment	0.868	0.914	0.716	0.232	0.333	0.179	0.315	0.713	0.531	

	Construct	CA	CR	AVE	1	2	3	4	5	6	7
Western (n=151)	1. Distributive justice	0.812	0.839	0.729							
	2. Informational justice	0.935	0.937	0.836	0.448						
	3. Procedural justice	0.855	0.869	0.697	0.813	0.560					
	4. Trust in AI	0.909	0.909	0.786	0.816	0.431	0.687				
	5. Trust in managers	0.884	0.884	0.742	0.202	0.625	0.386	0.220			
	6. Trust in organization	0.851	0.858	0.770	0.291	0.633	0.408	0.244	0.813		
	7. Affective commitment	0.814	0.879	0.638	0.281	0.399	0.342	0.256	0.655	0.644	

Note: CA: Cronbach's alpha; CR: composite reliability; AVE: average variance extracted; HTMT: heterotrait-monotrait ratio of the correlations.

Table 2. Reliability, Convergent Validity, and Discriminant Validity (using HTMT)

Since the study collected data with a single instrument within a single timeframe, there are concerns about common method bias (CMB) (Podsakoff et al., 2003). To address this concern, we employed both procedural and statistical remedies, using Harman's single-factor test (Podsakoff et al., 2003) and lateral multicollinearity assessment (Kock & Lynn, 2012). The results confirm that CMB was not an issue in this study.

5.2 Hypotheses testing

The Eastern and Western samples demonstrate consistent findings and provide partial support for our hypotheses (see Table 3). First, employee trust in AI is influenced by perceptions of distributive justice ($\beta_{East} = 0.46$, $\beta_{West} = 0.53$; $p < 0.001$) and procedural justice ($\beta_{East} = 0.36$, $\beta_{West} = 0.23$; $p < 0.001$) but not informational justice ($\beta_{East} = -0.01$, $\beta_{West} = 0.09$; n.s), thus supporting H1a and H1b but not H1c. Second,

informational justice perception positively influences employee trust in organization (β_{East} = 0.44, β_{West} = 0.53; $p < 0.001$) and trust in managers (β_{East} = 0.46, β_{West} = 0.55; $p < 0.001$). Meanwhile, no significant effects were observed for the relationships between distributive justice and trust in organization (β_{East} = -0.06, β_{West} = -0.04; n.s), distributive justice and trust in managers (β_{East} = -0.01, β_{West} = -0.15; n.s), procedural justice and trust in organization (β_{East} = 0.03, β_{West} = 0.11; n.s), and procedural justice and trust in managers (β_{East} = 0.002, β_{West} = 0.17; n.s). Accordingly, both samples confirmed H2c and H3c while rejecting H2a, H2b, H3a and H3b. No significant relationships were found between trust in organization, trust in managers and trust in AI, thus rejecting H4a and H4b. Finally, affective commitment is influenced by trust in AI (β_{East} = 0.12, β_{West} = 0.14; $p < 0.01$) and trust in managers (β_{East} = 0.59, β_{West} = 0.39; $p < 0.001$), but not trust in organization (β_{East} = 0.05, β_{West} = 0.23; n.s), thereby supporting H5a and H5c but not H5b.

Hypothesis	Eastern	Western	Supported?
H1a. Distributive Justice -> Trust in AI	0.46***	0.53***	Yes
H1b. Procedural Justice -> Trust in AI	0.36***	0.23**	Yes
H1c. Informational Justice -> Trust in AI	-0.01 n.s	0.09 n.s	No
H2a. Distributive Justice -> Trust in Organisation	-0.06 n.s	-0.04 n.s	No
H2b. Procedural Justice -> Trust in Organisation	0.03 n.s	0.11 n.s	No
H2c. Informational Justice -> Trust in Organisation	0.44***	0.53***	Yes
H3a. Distributive Justice -> Trust in Managers	-0.01 n.s	-0.15 n.s	No
H3b. Procedural Justice -> Trust in Managers	0.002 n.s	0.17 n.s	Yes
H3c. Informational Justice -> Trust in Managers	0.46***	0.55***	Yes
H4a. Trust in Organisation -> Trust in AI	0.12 n.s	-0.08 n.s	No
H4b. Trust in Managers -> Trust in AI	0.02 n.s	0.04 n.s	No
H5a. Trust in AI -> Affective Commitment	0.12**	0.14**	Yes
H5b. Trust in Organisation -> Affective Commitment	0.05 n.s	0.23 n.s	No
H5c. Trust in Managers -> Affective Commitment	0.59***	0.39***	Yes

Note: Significances: *** 0.1%; ** 1%; * 5%.; n.s not significant

Table 3. Hypotheses testing results

In summary, our findings suggest that employee trust in AI for performance evaluation is influenced by perceptions of distributive and procedural justice. Informational justice perception positively affects trust in organization and managers but does not exert influence on trust in AI. Trust in AI and trust in managers positively influence affective commitment.

5.3 Multigroup analysis for the gender effect

To explore the impact of gender on employees' perceptions regarding AI usage for performance evaluation within each culture, we employed a multigroup analysis (MGA). Before conducting MGA, we assessed the measurement invariance of composite models (MICOM) and found full measurement invariance for both samples, allowing for path coefficients comparison between men and women in each sample (Henseler et al., 2016). Table 4 presents the MGA results.

For both genders within each culture, the results show that distributive justice perception has a positive impact on employee trust in AI. Informational justice perception does not influence trust in AI, but exerts significant effects on trust in organization and trust in managers. Affective commitment was found to be dependent on trust in AI and trust in managers.

An interesting gender distinction emerges regarding the impact of procedural justice on trust in AI. In the Eastern sample, the relationship between procedural justice and trust in AI is significant for women but non-significant for men ($\beta_{East\text{-}women}$ = 0.55, $p < 0.001$; $\Delta\ \beta_{East\text{-}women}, \beta_{East\text{-}men}$ = 0.42, $p < 0.05$). Meanwhile, the Western sample reveals that the impact of procedural justice on trust in AI is significant for both genders ($\beta_{West\text{-}women}$ = 0.21, $\beta_{West\text{-}men}$ = 0.31; $p < 0.05$). Moreover, procedural justice exerts effects on both trust in organization ($\beta_{West\text{-}men}$ = 0.37, $p < 0.05$) and trust in managers ($\beta_{West\text{-}men}$ = 0.35, $p < 0.05$) for men, whereas no such effects were observed for women. Although the significant value for path coefficient differences is only at $p < 0.10$ (trust in organization: $\Delta\beta_{East\text{-}women}, \beta_{East\text{-}men}$ = -0.38; trust in managers: $\Delta\beta_{East\text{-}women}, \beta_{East\text{-}men}$ = -0.28), the findings speak to the marginal role of procedural justice perception in influencing employee trust in organization and managers for men respondents in the Western sample.

6. Discussion

This study provides a cross-cultural examination of employees' justice perceptions, trust in AI and their associated outcomes, as well as the gender impact on these relationships in the context of AI adoption for performance evaluation. Data collected from Western and Eastern countries largely supported our hypotheses, revealing both gender similarities and differences.

6.1 Theoretical contributions

This study contributes to the existing literature in three ways. *First*, it advances our insights into the relative effects of distributive, procedural, and informational justice perceptions on trust in AI. Our findings suggest that, regardless of cultural background, perceptions of distributive and procedural justice positively affect employee trust in AI, with distributive justice exerting a stronger influence. This finding aligns with previous evidence that individuals put more weight on distributive justice than procedural justice assessment in the context of AI usage (Morse et al., 2022). Moreover, studies on justice perceptions in performance evaluation indicate that distributive justice takes precedence over procedural justice because employees who perceive their outcomes as fair are likely to regard the procedures as fair and just (Gupta & Kumar, 2013).

Eastern sample (n=140)								
	Trust in AI		Trust in organization		Trust in managers		Affective commitment	
	Women	Men	Women	Men	Women	Men	Women	Men
Distributive justice	0.34**	0.55***	-0.19	0.07	-0.08	0.07		
Procedural justice	0.55***	0.11	0.12	0.00	0.12	-0.12		
Informational justice	0.03	0.06	0.47***	0.39**	0.43***	0.57***		
Trust in organization	0.07	0.17					0.11	-0.03
Trust in managers	0.01	0.01					0.59***	0.57***
Trust in AI							0.1*	0.22**
R^2	66.9%	54.0%	24.6%	16.9%	21.5%	26.5%	47.2%	42.7%
Western sample (n=151)								
	Trust in AI		Trust in organization		Trust in managers		Affective commitment	
	Women	Men	Women	Men	Women	Men	Women	Men
Distributive justice	0.54***	0.50***	0.09	-0.29	-0.05	-0.31		
Procedural justice	0.21*	0.31*	-0.01	0.37*	0.06	0.35*		
Informational justice	0.16	0.07	0.49***	0.56***	0.53***	0.53***		
Trust in organization	-0.12	-0.23					0.29	0.19
Trust in managers	0.15	0.010					0.32*	0.45**
Trust in AI							0.09*	0.19**
R^2	61.5%	55.1%	28.4%	50.9%	29.6%	43.0%	38.2%	39.5%

Note: Eastern sample: n(women) = 80; n(men) = 60; Western sample: n(women) = 87; n(men) = 64.
Significances: *** 0.1%; ** 1%; * 5%. Shaded areas detach the supported hypotheses.

Table 4. Effect of gender on the hypotheses

While distributive justice perception was found to be more strongly related to trust in AI compared to other justice dimensions, procedural justice remains essential. Procedural justice is particularly important in situations where information about the decision-maker's trustworthiness is uncertain, which is often the case for AI systems (Glikson & Woolley, 2020). Indeed, researchers are increasingly recognizing the importance of procedural justice perceptions in the context of AI usage for personnel and managerial decisions (Lee et al., 2019; Nagtegaal, 2021).

The findings also show that perceived informational justice serves as a primary driver of trust in organization and managers, but does not directly influence employee trust in AI. A possible explanation for this result is that informational justice involves the employees' assessment of information and explanations provided by their organization and managers during the performance evaluation process (Thurston & McNall, 2010), rather than originating from the AI system. Therefore, it is more strongly associated with employees' trust in their organization and managers in this context.

Second, our study contributes to the literature on the intersectionality of gender and culture pertaining to trust in AI. The findings demonstrate gender variances across Eastern and Western cultures regarding AI trust formation via justice perceptions. Specifically, in both cultures, perceived distributive justice plays a significant role in enhancing trust in AI for both men and women. However, the effect of procedural justice perception on trust in AI was found to be more salient for women than men in the Eastern sample, while it held equal significance for both genders in the Western sample. The gender distinction in procedural justice perception has been substantiated in prior research. Existing literature indicates that women tend to place greater emphasis on process-oriented issues than men do (Lee & Farh, 1999; Sweeney & McFarlin, 1997). Choi (2011) also found that women exhibit higher levels of trust in managers than men when they perceive procedural justice. One possible explanation may lie in the fact that women often have to rely more heavily on formal procedures as a means to ensure fair treatment and equitable opportunities within the organization due to historical discrimination and gender-role stereotyping which excluded them from decision-making processes (Choi, 2011; Sweeney & McFarlin, 1997). This phenomenon may be heightened in the Eastern countries examined in this study, primarily represented by participants from Vietnam, China, and India, which are characterized by collectivism and high power distance (Hofstede, 2011) in the workplace. The collectivist culture, which emphasizes group harmony and conformity, can inadvertently reinforce traditional gender roles, making it more difficult for women to challenge the status quo. Additionally, high power

distance in these cultures means that authority and decision-making are concentrated at the top levels of the organizational hierarchy, which are predominantly occupied by men. This structural barrier necessitates that women seek fair treatment and opportunities through formalized processes that are less susceptible to bias and favoritism. These findings lay a foundation for further theorizing the role of gender on individuals' justice perceptions and trust in AI.

Third, the study provides novel insights into the interplay between employee trust in AI systems and organizational behavior outcomes, including trust in organization and managers and affective commitment. It is interesting that employees' trust in organization and managers do not directly influence their trust in AI for both cultures. One plausible explanation for this finding is that the presence of other influential factors within this sample, such as organizational justice perceptions, may inhibit the impact of trust in the organization and managers. Our study also found that trust in AI for performance evaluation positively impacts affective commitment, a relationship that has not been thoroughly explored in prior research. This finding highlights the role of AI trust in fostering employees' emotional attachment to the organization. Overall, our findings shed light on the nuanced ways in which trust in AI can develop and influence employee attitudes across cultures and genders.

6.2 Practical implications

The findings offer practical insights for organizations aiming to enhance employees' fairness perception and trust in AI systems used for people management practices, such as performance evaluation. The findings underscore the significance of fairness perceptions in shaping trust in AI and employees' attitudes towards the organisation. They further indicate that distributive and procedural justice are most relevant to developing trust in AI. Accordingly, organizations should ensure the fairness and equity of outcomes in AI-assisted management practices. To enhance perceived procedural fairness, organizations should establish standardized processes, provide clear ethical guidelines, as well as develop mechanisms for employees' input and feedback. Finally, it is imperative to consider gender and cultural differences while introducing AI at the workplace. Acknowledging these variations allows organizations to tailor AI implementation interventions to better address diverse beliefs, values, and expectations of their employees.

6.3 Limitations and future research

This study is subject to certain limitations, which presents avenues for future research. *First*, our study design was cross-sectional and thus the ability to establish causality between variables was limited. *Second*, survey responses related to the use of AI for performance evaluation were mainly hypothetical, as not all organizations currently implementing AI for this purpose. Future research should conduct studies in real-world settings to obtain a more detailed understanding. A *third* limitation concerns the generalizability of results. While our sample includes respondents from six Eastern and twelve Western countries, it is not representative of these regions. Future investigations should seek to replicate and validate findings across other national cultures. *Finally*, there exist avenues for future research on cultural and gender effects. For instance, while our study categorized participants as Eastern or Western based on nationality, future research could delve deeper into other cultural dimensions, such as those proposed by Hofstede (2011), including power distance, uncertainty avoidance, masculinity-femininity, and individualism-collectivism. Regarding gender, our study measures gender as a biological construct, and thus future research could explore the effects of social, psychological, or cultural dimensions of gender to obtain a better understanding of the gender impact on AI perceptions.

7. Acknowledgements

This research was supported by Lero – the Science Foundation Ireland Research Centre for Software (www.lero.ie) PhD scholarship.

8. References

Acikgoz, Y., Davison, K. H., Compagnone, M., & Laske, M. (2020). Justice perceptions of artificial intelligence in selection. *International Journal of Selection and Assessment*, *28*(4), 399–416.

Alnefaie, A., & Kang, K. (2024). Cultural Transformations of E-commerce Consumer Behavior and Intention Toward Using Artificial Intelligence (AI) Assistants. *Proceedings of the 57th Hawaii International Conference on System Sciences*, 2423-2432.

Alzahrani, L., Al-Karaghouli, W., & Weerakkody, V. (2017). Analysing the critical factors influencing trust in e-government adoption from citizens' perspective: A systematic review and a conceptual framework. *International Business Review*, *26*(1), 164–175.

Aryee, S., Budhwar, P. S., & Chen, Z. X. (2002). Trust as a mediator of the relationship between organizational justice and work outcomes: Test of a social exchange model. *Journal of Organizational Behavior*, *23*(3), 267–285.

Bach, T. A., Khan, A., Hallock, H., Beltrão, G., & Sousa, S. (2024). A Systematic Literature Review of User Trust in AI-Enabled Systems: An HCI Perspective. *International Journal of Human-Computer Interaction*, *40*(5), 1251–1266.

Bankins, S., Formosa, P., Griep, Y., & Richards, D. (2022).

AI Decision Making with Dignity? Contrasting Workers' Justice Perceptions of Human and AI Decision Making in a Human Resource Management Context. *Information Systems Frontiers*, *24*(3), 857–875.
Bigman, Y. E., & Gray, K. (2018). People are averse to machines making moral decisions. *Cognition*, *181*, 21–34
Blau, P. (1964). *Exchange and Power in Social Life*. Wiley.
Choi, S. (2011). Organizational justice and employee work attitudes: The federal case. *American Review of Public Administration*, *41*(2), 185–204.
Choi, J., Lee, H. J., Sajjad, F., & Lee, H. (2014). The influence of national culture on the attitude towards mobile recommender systems. *Technological Forecasting and Social Change*, *86*, 65-79.
Cohen-Charash, Y., & Spector, P. E. (2001). The role of justice in organizations: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *86*(2), 278–321.
Colquitt, J. A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology*, *86*(3), 386–400.
Colquitt, J. A., Scott, B. A., Rodell, J. B., Long, D. M., Zapata, C. P., Conlon, D. E., & Wesson, M. J. (2013). Justice at the millennium, a decade later: A meta-analytic test of social exchange and affect-based perspectives. *Journal of Applied Psychology*, *98*(2), 199–236.
Coyle-Shapiro, J. A. M., & Conway, N. (2005). Exchange relationships: Examining psychological contracts and perceived organizational support. *Journal of Applied Psychology*, *90*(4), 774–781.
Cropanzano, R., Anthony, E. L., Daniels, S. R., & Hall, A. V. (2017). Social exchange theory: A critical review with theoretical remedies. *Academy of Management Annals*, *11*(1), 479–516.
Cropanzano, R., Bowen, D. E., & Gilliland, S. W. (2007). The Management of Organizational Justice. *Academy of Management Perspectives*, *21*(4), 34–48.
Denisi, A. S., & Murphy, K. R. (2017). Performance appraisal and performance management: 100 years of progress? *Journal of Applied Psychology*, *102*(3), 421–433
Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. *International Journal of Information Management*, *48*, 63–71.
Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160.
Feldkamp, T., Langer, M., Wies, L., & König, C. J. (2023). Justice, trust, and moral judgements when personnel selection is supported by algorithms. *European Journal of Work and Organizational Psychology*, *00*(00), 1–16.
Figueroa-Armijos, M., Clark, B. B., & da Motta Veiga, S. P. (2023). Ethical Perceptions of AI in Hiring and Organizational Trust: The Role of Performance Expectancy and Social Influence. *Journal of Business Ethics*, *186*(1), 179–197.
Fletcher, C. (2001). Performance appraisal and management: The developing research agenda. *Journal of Occupational and Organizational Psychology*, *74*(4), 473–487.
Formica, S., & Sfodera, F. (2022). The Great Resignation and Quiet Quitting paradigm shifts: An overview of current situation and future research directions. *Journal of Hospitality Marketing and Management*, *31*(8), 899–907.
Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, *18*(1).
Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, *14*(2), 627–660.
Gouldner, A. W. (1960). The Norm of Reciprocity : A Preliminary Statement. *American Sociological Review*, *25*(2), 161–178.
Gupta, V., & Kumar, S. (2013). Impact of performance appraisal justice on employee engagement: A study of Indian professionals. *Employee Relations*, *35*(1), 61–78.
Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2–24.
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.
Henseler, J., Ringle, C. M., & Sarstedt, M. (2016). Testing measurement invariance of composites using partial least squares. *International Marketing Review*, *33*(3), 405–431
Höddinghaus, M., Sondern, D., & Hertel, G. (2021). The automation of leadership functions: Would people trust decision algorithms? *Computers in Human Behavior*, *116*.
Hofstede, G. (2011). Dimensionalizing Cultures: The Hofstede Model in Context. *Online Readings in Psychology and Culture*, *2*(1), 1–26.
IBM (2021). *The business case for AI in HR*.
IBM & Morning Consult (2024). *IBM Global AI Adoption Index 2023*.
Jiang, Z., Gollan, P. J., & Brooks, G. (2017). Relationships between organizational justice, organizational trust and organizational commitment: a cross-cultural study of China, South Korea and Australia. *International Journal of Human Resource Management*, *28*(7), 973–1004.
Joshi, K. D. (2022). Conceptualizing the Marginalized Context in Information Systems Research: An Intersectionality View. *Data Base for Advances in Information Systems*, *53*(2), 7–10.
Kaplan, A. D., Kessler, T. T., Brill, J. C., & Hancock, P. A. (2021). Trust in Artificial Intelligence: Meta-Analytic Findings. *Human Factors*, 1–25.
Kock, N., & Lynn, G. S. (2012). Lateral collinearity and misleading results in variance-based SEM: An illustration and recommendations. *Journal of the Association for Information Systems*, *13*(7), 546–580.
Lambert, L. S., Bingham, J. B., & Zabinski, A. (2020). Affective commitment, trust, and the psychological contract: contributions matter, too! *European Journal of Work and Organizational Psychology*, *29*(2), 294–314.
Lambert, L. S., Edwards, J. R., & Cable, D. M. (2003). Breach and fulfillment of the psychological contract: A

comparison of traditional and expanded views. *Personnel Psychology*, *56*(4), 895–934.

Langer, M., König, C. J., Back, C., & Hemsing, V. (2022). Trust in Artificial Intelligence: Comparing Trust Processes Between Human and Automated Trustees in Light of Unfair Bias. *Journal of Business and Psychology*.

Langer, M., & Landers, R. N. (2021). The future of artificial intelligence at work: A review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers. *Computers in Human Behavior*, *123*(May), 106878.

Lee, C., & Farh, J.-L. (1999). The effects of gender in organizational justice perception. *Journal of Organizational Behavior*, *20*(1), 133–143.

Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data and Society*, *5*(1), 1–16.

Lee, M. K., Jain, A., Cha, H. J. I. N., Ojha, S., & Kusbit, D. (2019). Procedural justice in algorithmic fairness: Leveraging transparency and outcome control for fair algorithmic mediation. *Proceedings of the ACM on Human-Computer Interaction*, *3*.

Leidner, D. E. L., & Kayworth, T. (2006). A Review of Culture in Information Systems Research: Toward a Theory of Information Technology Culture Conflict. *MIS Quarterly*, *30*(2), 357–399.

Malik, A., Budhwar, P., & Kazmi, B. A. (2023). Artificial intelligence (AI)-assisted HRM: Towards an extended strategic framework. *Human Resource Management Review*, *33*.

Meyer, J. P., & Allen, N. J. (1991). A Three-Component Conceptualization of Organizational Commitment. In *Human Resource Management Review*, *1*(1), 61–89.

Mirowska, A., & Mesnet, L. (2022). Preferring the devil you know: Potential applicant reactions to artificial intelligence evaluation of interviews. *Human Resource Management Journal*, *32*(2), 364–383.

Morse, L., Teodorescu, M. H. M., Awwad, Y., & Kane, G. C. (2022). Do the Ends Justify the Means? Variation in the Distributive and Procedural Fairness of Machine Learning Algorithms. *Journal of Business Ethics*, *181*(4), 1083–1095.

Nagtegaal, R. (2021). The impact of using algorithms for managerial decisions on public employees' procedural justice. *Government Information Quarterly*, *38*(1).

Pereira, V., Hadjielias, E., Christofi, M., & Vrontis, D. (2023). A systematic literature review on the impact of artificial intelligence on workplace outcomes: A multi-process perspective. *Human Resource Management Review*, *33*(1).

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology*, *88*(5), 879–903.

PwC (2023). *Global Workforce Hopes & Fears Survey 2023*.

Robinson, S. L. (1996). Trust and breach of the psychological contract. *Administrative Science Quarterly*, *41*(4), 574–599.

Roose, K. (2019, June 23). A Machine May Not Take Your Job, but One Could Become Your Boss. *The New York Times*.https://www.nytimes.com/2019/06/23/technology/artificial-intelligence-ai-workplace.html

Rousseau, D. M. (1989). Psychological and Implied Contracts in Organizations. *Employee Responsibilities and Rights Journal*, *2*(2), 121–139.

Rousseau, D. M., Hansen, S. D., & Tomprou, M. (2018). A dynamic phase model of psychological contract processes. *Journal of Organizational Behavior*, *39*(9).

Rousseau, D. M., & Tijoriwala, S. A. (1998). Assessing psychological contracts: Issues, alternatives and measures. *Journal of Organizational Behavior*, *19*, 679–695.

Solberg, E., Kaarstad, M., Eitrheim, M. H. R., Bisio, R., Reegård, K., & Bloch, M. (2022). A Conceptual Model of Trust, Perceived Risk, and Reliance on AI Decision Aids. *Group & Organization Management*, *47*(2), 187–222.

Starke, C., Baleis, J., Keller, B., & Marcinkowski, F. (2022). Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature. *Big Data and Society*, *9*(2).

Sweeney, P. D., & McFarlin, D. B. (1997). Process and outcome: Gender differences in the assessment of justice. *Journal of Organizational Behavior*, *18*(1), 83–98.

Thurston, P. W., & McNall, L. (2010). Justice perceptions of performance appraisal practices. *Journal of Managerial Psychology*, *25*(3), 201–228.

Trauth, E. M. (2011). Rethinking Gender and MIS for the Twenty-First Century. In R. D. Galliers & W. L. Currrie (Eds.), *The Oxford Handbook of Management Information Systems: Critical Perspectives and New Directions* (pp. 560-585). Oxford University Press.

Tremblay, M., Cloutier, J., Simard, G., Chênevert, D., & Vandenberghe, C. (2010). The role of HRM practices, procedural justice, organizational support and trust in organizational commitment and in-role and extra-role performance. *International Journal of Human Resource Management*, *21*(3), 405–433.

van der Werff, L., Blomqvist, K., & Koskinen, S. (2021). Trust Cues in Artificial Intelligence: A Multilevel Case Study in a Service Organization. In *Understanding Trust in Organizations: A Multilevel Perspective* (pp.307–333).

Vance, A., Elie-dit-cosaque, C., & Straub, D. W. (2008). Examining trust in information technology artifacts: The effects of system quality and culture. *Journal of Management Information Systems*, *24*(4), 73–100.

Vos, J. F., & Boonstra, A. (2022). The influence of cultural values on Enterprise System adoption, towards a culture–Enterprise System alignment theory. *International Journal of Information Management*, *63*, 102453.

Yu, J., Ma, Z., & Zhu, L. (2023). The configurational effects of artificial intelligence-based hiring decisions on applicants' justice perception and organisational commitment. *Information Technology and People*.