Responsible Artificial Intelligence in Talent Acquisition: The Role of Algorithmic Fairness in Recruitment Decision-Making and Organizational Performance

Kum. Ganjikunta Navya Sree[1], Dr. G. Kanuka Raju[2], Mrs*. P. Sujatha[3]*

[1]Independent researcher, gnavyasree29@gmail.com.

[2]Independent researcher, drgkr2022@gmail.com.

[3]Assistant Professor, Dept. of MBA, Sri Venkateswara Institute of Science and Technology, Kadapa, Kadapa District, Andhra Pradesh, E-mail: pasalasujatha20@gmail.com

Abstract

Placing artificial intelligence (AI) within talent acquisition is quickly taking root in organizations of all types of industries, though there is some anxiety of whether or not the actual practices of screening and evaluation of candidates through an algorithm can yield recruitment results that are both realistic and defensible. This paper, based on responsible AI and organizational information processing lenses, investigates the relationships between Responsible AI (RAI) practices in recruitment and perceived Algorithmic Fairness (AF), Recruitment Decision-Making quality (RDM), and Organizational Performance (OP). The survey information was gathered using data on 200 HR professionals and recruiters working in Indian organizations where AI-enabled recruitment tools have been implemented. The indicators of consistency were good (Cronbach alpha was between 0.86 to 0.88) and the problem of common method bias was not severe (Harman single-factor test = 36.31%). As mediated by regression analysis (with 5,000-resample bootstrapping revealed) showed, practices of RAI were a significant predictor of algorithmic fairness ($b = 0.47$, $p < .001$), which was a significant predictor of quality of recruiting decisions ($b = 0.44$, $p < .001$) and through this, organizational performance ($b = 0.40$, $p < .001$). The relationship between responsible AI practices and organizational performance was completely mediated by algorithmic fairness and the quality of decisions made by the majority of recruitment because there was no significant ($r = .098$, $p = .166$) direct relationship between RAI and OP. The serial indirect effect (RAI → AF → RDM → OP) was significant ($a \times b \times d = 0.084$, 95% CI [0.046, 0.131]). The results indicate that the advantages of AI-enabled recruitment applied to performance are achieved rather by fairness of the algorithms and quality of the decisions to which they provide information than by automation itself. Theoretical and practical implications of responsible AI governance of human resource management are presented.

Keywords: responsible AI, algorithmic fairness, decision-making in recruitment, talent acquisition, organizational performance, mediation analysis.

1. Introduction

Artificial intelligence (AI) has ceased to be an outsourced experiment to become a heart of talent acquisition. There are now resume-screening algorithms, automated video-interview scoring, chatbots used to engage with candidates, predictive fit-scoring models, and all of these features are implemented throughout the recruitment funnel of big and mid-size firms. Advocates believe that such tools save time in hiring candidates, ensure lower administrative expense, and create equal evaluation criteria among hiring outsourcers. Meanwhile, an expanding literature has voiced scepticism that algorithmic hiring systems may reproduce or amplify established patterns of disadvantage when conditioned on historical data, which are themselves based on bias human choices, and that the claims by vendors that they are fair is often hard to be checked by adopting organizations alone. It is against this backdrop that the discourse of Responsible Artificial Intelligence (RAI): a set of organizational practices, including transparency in model logic, human control of algorithmic recommendations, bias auditing, explainability, and accountability mechanisms, is emerging to help guarantee the deployment of AI systems in

manners that are equitable, secure, and aligned with organizational and social values. Although much has been said about RAI on the level of principle and policy, very little empirical evidence has been carried out on the way RAI practices are translated into the perceived fairness of these systems by the recruiters and the hiring managers who apply them on a daily basis, and whether perceived fairness reflects on the quality of the recruitment decision and downstream organizational performance. The paper will fill that void with survey data of 200 HR professionals and recruiters in the companies with operational operations in India and confirmed that adopted AI-enabled recruitment tools. In particular, it analyses a serial mediation model where the Responsible AI practices (RAI) are assumed to predict perceived Algorithmic Fairness (AF), which makes its predictions to Recruitment Decision-Making quality (RDM), which in its turn makes its predictions to Organizational Performance (OP). The study implicates the insights into responsible AI practice and algorithmic fairness perception measurement in the hiring setting because (a) it operationalizes and validates measures of responsible AI practice, and fairness perceptions and the decision quality; (b) it hypothesizes the channelling of the performance payoff of AI-based recruitment; (c) and provides organizations with a diagnostic framework to assess whether its AI recruitment investments are being translated into defensible and high-quality hiring decisions.

2. Literature Review and Hypotheses Development

2.1 Responsible AI in Talent Acquisition

Responsible AI in HR can be understood as a collection of governance practices followed by organizations to make sure that the algorithm systems applied in job hiring are transparent, explainable, auditable and subject to meaningful human action. According to Cappelli, Tambe, and Yakubovich (2019), there has long been a discrepancy between the promise and reality of data-driven HR decision-making due to the complexity of HR phenomena, the constraints of small and deceitful data sets and accountability issues regarding fairness and legal liability. In those organizations where responsible-use practices have been institutionalized as in written model logic and observed disparate impact prior to implementation and human retention of high-stakes recommendations, there is greater chance that recruiters will trust and reasonably tune their trust to the results of their algorithms.

2.2 Algorithmic Fairness in Recruitment

The concepts of algorithmic fairness related to the hiring process are usually interpreted on both procedural and distributive levels: are the steps employed to screen and rank candidates perceived as stable and expiable by all viable approaches, and are the final results perceived as fair and equal among the groups of candidates? Empirical reviews of commercial recruiting-evaluation providers have discovered that some vendors show significant differences in how they capture and or justify their bias-reduction efforts, with many giving little disclosure to the statistical foundation of their assertions (Raghavan et al., 2020; Bogen and Rieke, 2018). Since the actual users and subjects of algorithmic suggestions are recruiters and hiring managers who inter-rater perceive fairness (as opposed to fairness as a technical characteristic of the model), these are the users who determine the tool actually looks in practical terms (Barocas and Selbst, 2016).

2.3 Recruitment Decision-Making Quality and Organizational Performance

Recruitment decision-making quality can be defined as accuracy, consistency, and defend ability of hiring decision-making the degree to which hiring decisions are made based on job relevant factors, whether they are consistent across similar candidates and can be defended when faced with criticism. The hypothesized contribution of higher-quality recruitment decisions to organizational performance is enhanced performance due to an improved person-job fit, lower first-year turnover, and enhanced employer brand, in line with broader strategic human resource management logic that personnel staffing quality is valuable to an organization in terms of its output.

2.4 Hypotheses

According to the above argument, the hypotheses are presented below and tested:

H1: Responsible AI practices-RAI is positively linked to perceived Algorithmic Fairness (AF). H2: Quality of Recruitment Decision-Making (RDM) has a positive relationship with Algorithmic Fairness (AF).

H3: Decision-Making quality Recruitment Decision-Making (RDM) positively correlates with Organizational Performance (OP).

H4: There is a mediation between Responsible AI practices and the quality of Recruitment Decision-Making through Algorithmic Fairness.

H5: Recruitment Quality To test the hypothesis of mediation between Algorithmic Fairness and Organizational performance, the query is based on Recruitment Decision-Making quality. H6: Algorithmic fairness and the quality of Recruitment decision-making choices mediate between the Responsible AI practices and the Organizational Performance in a serial way.

2.5 Conceptual Framework

The proposed model will place Responsible AI practices as the most distant predictor, Algorithmic Fairness and Recruitment Decision-Making quality as chain of events in between, and Organizational Performance as the result: RAI KSF. The estimation of direct paths between RAI and RDM, and between AF and OP is also made to test the partial and the complete mediation.

3. Research Methodology

3.1 Research Design and Sample

The research design and sample are described in 3.1. The survey design was a cross-sectional one. The respondents were HR professionals and recruiters who were employed in organizations that had implemented at least one AI-powered recruitment technology (resume screening, automated interview scoring, or predictive candidate matching). The structured questionnaire was used to collect data among the recruitment and talent-acquisition professionals working in five broad industry sectors (IT/ITES, manufacturing, BFSI, public sector, and other services). Having filtered out responses and attention failures due to incomplete responses, a final analysable sample N = 200 was retained

3.2 Measures

The scale of all constructs was measured on 5-point Likert scales (1 = strongly disagree, 5 = strongly agree), using multi-item scales based on multi-item adaptations of the measures of responsible AI, algorithmic fairness, and strategic HRM literature. Responsible AI practice (RAI) was assessed using 6 items representing transparency, human oversight, bias auditing and accountability mechanisms. The Algorithmic Fairness (AF) was assessed using 5 items that reflected procedural and distributive perceptions of fairness. Recruitment Decision-Making quality (RDM) was assessed using 5 items which focused on accuracy, consistency and defensibility of the hiring decisions. Organizational performance (OP) was assessed using 5 questions that included recruitment efficacy, quality of hire, and employer-brand results. The reliability analysis (Table 2) indicated that Cronbach alpha values were 0.869 (RAI), 0.881 (AF), 0.878 (RDM) and 0.863 (OP), which were higher than 0.70 the traditional threshold of acceptable internal consistency.

3.3 Common Method Bias

Since the measured constructs all relied on single respondent self-report measures, the single factor test of Harman was done by placing all 21 items in an unrotated principal components analysis. The former explained 36.31% of the total variance, which is less than 50 50% which is conventionally a marker of serious common method bias indicating that it is very unlikely that common method variance will explain the relationships.

3.4 Data Analysis Technique

Descriptive statistics, Pearson correlation, and ordinary least squares (OLS) regression-based mediation according to the Baron and Kenny causal-steps logic were used to analyse the data, along with Sobel tests and bias-corrected bootstrapping (5,000 resamples) indirect-effect confidence intervals, which is in line with modern mediation analysis practice (Hayes, 2018 approach implemented via custom OLS/bootstrap routines in Python).

4. Results

4.1 Demographic Profile of Respondents

Table 1 shows demographic and organizational profile of the 200 respondents.

Category	Group	n	%
Gender	Male	115	57.5
	Female	85	42.5
Age	25–34 years	63	31.5
	35–44 years	72	36.0
	45–54 years	43	21.5
	55 years and above	22	11.0
Recruitment Experience	Less than 5 years	47	23.5
	5–10 years	68	34.0
	11–15 years	58	29.0
	More than 15 years	27	13.5
Organization Size	Fewer than 500 employees	65	32.5
	500–2000 employees	80	40.0
	More than 2000 employees	55	27.5
Sector	IT / ITES	62	31.0
	BFSI	41	20.5
	Manufacturing	39	19.5
	Public Sector / Government	30	15.0
	Other Services	28	14.0
AI Recruitment Tool Tenure	Less than 1 year	49	24.5
	1–3 years	95	47.5
	More than 3 years	56	28.0

Table 1. Demographic and Organizational Profile of Respondents (N = 200)

4.2 Descriptive Statistics and Reliability

Construct	Items	M	SD	Min	Max	Cronbach's α
Responsible AI Practices (RAI)	6	3.57	0.59	1.83	5.00	0.869
Algorithmic Fairness (AF)	5	3.56	0.65	1.20	5.00	0.881
Recruitment Decision-Making (RDM)	5	3.66	0.62	2.00	5.00	0.878
Organizational Performance (OP)	5	3.59	0.61	2.00	5.00	0.863

Table 2. Descriptive Statistics and Internal Consistency Reliability

The totality of constructs fell at middle-range heavily between the scale midpoint (M range = 3.563.66), showing

generally positive but not quite homogenous perceptions of responsible AI practice, fairness, quality of decisions, and performance among the respondents.

4.3 Correlation Analysis

Construct	1	2	3	4
1. RAI	—			
2. AF	0.434**	—		
3. RDM	0.428**	0.563**	—	
4. OP	0.098	0.394**	0.500**	—

*Table 3. Pearson Correlation Matrix. **p < .01 (two-tailed). The non-significant zero-order correlation between RAI and OP (r = .098, p = .166) is consistent with full mediation via AF and RDM.*

4.4 Regression and Mediation Analysis

The results of the OLS regression coefficient of each path of the structure in the hypothesized serial mediation model can be seen in Table 4.

Path	b	SE	p	R^2
RAI → AF (path a)	0.473	0.070	< .001	0.188
RAI → RDM, total effect (path c)	0.443	0.067	< .001	—
AF → RDM, controlling RAI (path b)	0.442	0.060	< .001	0.359
RAI → RDM, controlling AF (path c′, direct)	0.234	0.066	< .001	—
RDM → OP, controlling AF (path d)	0.402	0.073	< .001	0.268
AF → OP, controlling RDM (path e)	0.156	0.069	= .026	—

Table 4. Regression Coefficients for Structural Paths

The hypothesis 1 was accepted: The prediction of AF was significant (b = 0.473, SE = 0.070, p <.001), and RAI explained 18.8-percent of the AF variance. Hypothesis 2 had proven true: AF was significantly predictive of RDM, even controlling for RAI (b = 0.442, p <. 001); even model (RAI + AF) only explained 35.9% of the variance in RDM. Hypothesis 3 was approved: The effect of RDM on OP was significant after the removal of AF (b = 0.402, p <.001); the complete model (AF + RDM) explained the 1/4th of the variance in OP. The AF (a x b) of RAI to RDM was estimated as 0.209, Sobel z = 4.979 with a p <.001 with [0.136, 0.290] bias corrected bootstrap confidence interval (excluding zero) supporting Hypothesis 4. Since the direct effect of RAI on RDM was significant in the presence of AF (c prime= 0.234, p = .001) but the combined effect is c = 0.443, this pattern suggests that AF partially mediates the relationship between RAI and RDM. In Hypothesis 5, the zero-order correlation between RAI and OP was not significant statistically (r =.098, p =.166), but the correlations between AF and RDM with OP were significant. This trend, namely, a nonsignificant total effect between the distal predictor and the outcome that is not standardized by significant indirect route is congruent with the complete (indirect-only) mediation of the RAIOP association by AF and RDM. Hypothesis 6 was a test of the full serial indirect effect RAI = AF = RDM = OP (a x b d). The magnitude of this effect was 0.084 with bootstrap 95% interval [0.046, 0.131] and not equal to zero, which is in favor of Hypothesis 6. When considered collectively, the findings suggest that responsible AI practices can affect organizational performance indirectly by chaining in which anticipated algorithmic fairness is first created by these practices, which can then increase the quality of recruitment decisions, subsequently boosting organizational performance.

4.5 Summary of Hypotheses Testing

Hypothesis	Path	Result
H1	RAI → AF	Supported

H2	AF → RDM	Supported
H3	RDM → OP	Supported
H4	AF mediates RAI → RDM	Supported (partial mediation)
H5	RDM mediates AF → OP	Supported (full mediation of RAI–OP link)
H6	AF & RDM serially mediate RAI → OP	Supported

Table 5. Summary of Hypotheses Testing

5. Discussion

The results prove that the organizational performance rationale behind AI-based recruitment should be treated as a sequence of mediators where the positive effects of the applied technology are not intrinsically linked to its use, but rather conditional on perceptions of the underlying algorithms as being fair, and the perception of these perceptions of being fair turning into quality recruiting decisions. This aligns with fears expressed in the literature on algorithmic fairness that the existence of an AI hiring tool itself is not a guarantee of fair and defendable results; rather it is the responsible-AI scaffolding that envelops the tool, such as transparency, human directing, and auditing, that is the scaffolding perceived as neutral by professionals utilizing it (Raghavan et al., 2020; Bogen and Rieke, 2018). The non-significant zero-order association between RAI and OP with the significant indirect relationships is a theoretically meaningful trend. It posits that companies engaging in responsible AI practices without minding how these practices can be received and implemented by frontline recruiters can fail to realize performance effects. In their turn, companies that manage to establish authentic perceived fairness around their AI recruitment tools seem poised to translate the perceived fairness into improved decision-making and, eventually, improved organizational performance. This result that the RAI-RDM relationship is partially mediated by responsible AI practices (H4) demonstrates that responsible AI practices not only enhance the quality of a decision-making process but can facilitate quality improvement through other means than the perceived fairness only, e.g., by providing better-quality information to recruiters themselves or by equipping decision-making standards with greater uniformity across all perceptions of equity.

6. Theoretical and Practical Implications

6.1 Theoretical Implications

The paper contributes to the literature on responsible AI, which has thus far been more theoretical or auditory (Raghavan et al., 2020; Bogen and Rieke, 2018), by modelling empirically how the practices of responsible AI translate into perceived fairness on the part of the recruiter, and into the final decision and performance outcomes. It also adds to the strategic human resource management theory by defining algorithmic fairness and the quality of decisions as sequential processes linking a technology-governance antecedent with firm-level performance, but not by considering the adoption of AI as a driver of performance in its own right.

6.2 Practical Implications

Two recommendations to the HR leaders and organisations who use AI in talent acquisition would be to invest in responsible-AI governance tools, such as transparency of the algorithm, bias audits, and retained human oversights, in the same breath follow up by attempting to explain these practices to recruiters and hiring managers, as it is their sense of equity and not the technical presence of safeguards that may result in better decision quality. The validated measures in this study can also be utilized by recruitment technology vendors and internal AI governance teams, as a diagnostic tool to benchmark how much trust is placed by recruiters in deployed systems at any point in time.

7. Limitations and Future Research

There are a number of limitations to this study. First, cross-sectional design does not allow powerful causal inferences; longitudinal/experimental design would reinforce the arguments on the direction that the hypothesized paths have an understanding of. Second, each construct was measured through self-report of recruitment professionals and, even though single-factor testing by Harman did not suggest a situation of grave common method bias, multi-source data (e.g., matching the perception held by recruiters with objective data on hiring outcomes) could have enhanced the design. Third, the sample was restricted to Indian organizations within a small sample of sectors; generalization to other institutions and regulatory environments ought to be determined in future studies. Future research may also focus on the perception of fairness of candidates, any boundary conditions (e.g., AI literacy or organizational AI maturity) and how regulatory setting can help condition the relationship between RAI and fairness.

8. Conclusion

This research paper exhibits that the crucial aspect of AI in the talent acquisition is responsible AI practices, which are positively correlated with perceived algorithmic fairness facilitating quality recruitment decisions, which, followed by the better quality of recruitment decisions, organizational performance. The indirect, but seemingly automatic, benefits of AI-enabled recruitment seem to be going through the fairness perceptions of recruiters and the quality of decision-making flowing through them. Responsible-AI governance practices of transparency, auditability, and human oversight should therefore be promoted by organizations that are interested in achieving the said efficiency and performance gains offered by AI in hiring employees, as they develop a very real sense of fairness among the professionals upon which such systems run.

Declarations

Funding: This study was not funded by any particular grant through funding organizations in the public, commercial, and not-for-profit sectors. Conflict of interest: The author states that he has no conflicts of interest. Ethical Approval: The research was conducted in accordance with normal ethical requirements of survey research, participation was voluntary and anonymity and confidentiality of the respondents were ensured during data collection and data analysis. Data Availability: The data created and manipulated in the definition of the current study is provided through the respective author on an adequate notice.

References

1. Ajunwa, I. (2020). The paradox of automation as anti-bias intervention. Cardozo Law Review, 41(5), 1671–1742.
2. Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. California Law Review, 104(3), 671–732.
3. Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. Proceedings of the 1st Conference on Fairness, Accountability and Transparency, 149–159.
4. Bogen, M., & Rieke, A. (2018). Help wanted: An examination of hiring algorithms, equity, and bias. Upturn.
5. Cappelli, P., Tambe, P., & Yakubovich, V. (2019). Artificial intelligence in human resources management: Challenges and a path forward. California Management Review, 61(4), 15–42.
6. Dastin, J. (2018, October 10). Amazon scraps secret AI recruiting tool that showed bias against women. Reuters.
7. Hayes, A. F. (2018). Introduction to mediation, moderation, and conditional process analysis: A regression-based approach (2nd ed.). Guilford Press.
8. Köchling, A., & Wehner, M. C. (2020). Discriminated by an algorithm: A systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development. Business Research, 13(3), 795–848.

9. Kim, P. T. (2017). Data-driven discrimination at work. William & Mary Law Review, 58(3), 857–936.

10. 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.

11. Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2020). Mitigating bias in algorithmic hiring: Evaluating claims and practices. Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 469–481.

12. Sánchez-Monedero, J., Dencik, L., & Edwards, L. (2020). What does it mean to 'solve' the problem of discrimination in hiring? Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 458–468.