

Perceived Quality of AI-Supported Recruitment: The Role of Explainability, Trust and Human Control

DOI: 10.12776/qip.v30i2.2374

Coi Tran, Radoslav Delina

Received: 12-06-2026 Accepted: 30-06-2026 Published: 31-08-2026

ABSTRACT

Purpose: This paper examines perceived quality of Artificial Intelligence (AI)-supported recruitment by testing how explanation, confidence information, practical AI experience, trust, willingness to use AI, perceived fairness and human control shape evaluations of AI candidate-ranking recommendations.

Methodology/Approach: The study uses an anonymous survey of Human Resources (HR)-related respondents (N = 145) and a randomised scenario comparing AI ranking without explanation with ranking supported by brief reasoning and confidence information. We use reliability checks, regression models, mediation and automation tests.

Findings: Practical AI experience significantly moderated the effect of explanation on trust and showed a weaker pattern for willingness to use AI. The moderation was not significant for perceived fairness. Intention to use AI was associated with trust, perceived usefulness, human control and prior AI experience.

Research Limitation/Implication: The non-probabilistic sample has a strong Central European component. The results measure perceptions rather than behavioural outcomes or audited fairness.

Originality/Value of paper: The paper shows that explanation and confidence strengthen trust mainly among AI-experienced respondents but do not significantly improve perceived fairness. It distinguishes between trusting AI as decision support and being convinced that AI-supported ranking is fair to candidates.

Category: Research paper

Keywords: artificial intelligence; recruitment; perceived quality; explainability; human control

Research Areas: Management of Technology and Innovation; Quality by Innovation

1 INTRODUCTION

Artificial intelligence (AI) is increasingly applied in recruitment and candidate selection processes, from vacancy drafting, sourcing and candidate communication to screening, ranking and interview support. Existing studies show that AI can improve not only the speed or scalability of recruitment, but also they emphasise ethical risks related to opacity, algorithmic bias, privacy and accountability (Hunkenschroer & Luetge, 2022; Will, Krpan, & Lordan, 2023; Fabris et al., 2025). For Human Resources (HR) managers, the central question is therefore not only whether AI can accelerate recruitment processes, but whether AI-supported recruitment can be implemented as a high quality, trustworthy and fair decision-support process.

This paper focuses on AI-supported recruitment as a quality and innovation management problem rather than only an HR technology issue. In recruitment, quality goes beyond speed or cost reduction because it also concerns the reliability of recommendations, transparency, human control or auditability and fairness. In this study, we perceive AI quality as a socio-technical evaluation of AI decision support, whether HR professionals perceive the recommendation as useful, understandable, trustworthy, controllable and fair. Quality 4.0 and human-centred quality-management literature provide a useful basis for this impact because they connect data-driven decision support with people, processes, governance and total organisational value (Liu, Liu, Gu, & Yang, 2023; Carvalho, Dias, Dias, & Sampaio, 2024; Sony, Antony, & Douglas, 2020). Based on Quality 4.0 and human-centred quality-management thinking approaches, this paper positions AI-supported recruitment as a sensitive decision-support process where quality is not limited to automation and its performance. It also depends on whether HR professionals perceive AI recommendations as understandable, trustworthy and fair.

The empirical contribution of this study is based on a survey of HR-related professionals with a randomised scenario-based module. The study tries to reveal whether brief reasoning and confidence information attached to an AI-supported recruitment recommendation improves trust, willingness to use AI and perceived fairness, and whether practical AI experience changes these effects. The paper contributes to the literature by separating three outcomes that are often implicitly connected: trust in the AI output, readiness to use the AI tool and perceived fairness of the AI-driven recruitment. This distinction is practically important because an HR user may understand and validate AI recruitment recommendation and may trust the recommendation as useful decision support, while still not being convinced that the ranking is fair, non-discriminatory and valid across candidate groups.

2 LITERATURE SURVEY

Quality and innovation management are increasingly influenced by the use of artificial intelligence in decision-making processes across the whole organisation. Quality 4.0 is commonly described as the alignment of quality-management practices with digital technologies, advanced analytics and data-driven decision support, while a more human-centred quality perspective emphasises the role of people, governance and social impact of technological esp. AI-driven decisions (Sony et al., 2020; Liu et al., 2023). Recent literature also highlights that digital quality is not a purely technical agenda. It requires data governance, leadership, training, culture, process integration and human interaction in quality planning, assurance, control and improvement (Liu et al., 2023; Carvalho et al., 2024). This perspective is relevant for AI-supported recruitment because recruitment is not only an HR technology issue. It is a quality-sensitive service and decision-support process in which AI may improve speed, scalability and consistency, but may also introduce opacity, bias, insufficient human control and fairness concerns.

AI-enabled recruitment has been studied as both an opportunity and a risk. Reviews and conceptual frameworks show that AI can support several stages of the recruitment pipeline. These include vacancy generation, sourcing, screening, candidate matching, ranking, assessment as well as coordination (Hunkenschroer & Luetge, 2022; Chen, 2023; Will et al., 2023). The HIRE framework suggests that AI can be equal or superior to human recruiters in efficiency and some performance-related outcomes, but perceptions of AI hiring are often worse than perceptions of human hiring, especially in relation to trust and perceived humaneness (Will et al., 2023). At the same time, algorithmic hiring is considered as a sensitive organisational decision-support process where historical data and opaque modelling choices can reproduce or activate discrimination (Raghavan, Barocas, Kleinberg, & Levy, 2020; Fabris et al., 2025). Ethical reviews therefore recommend transparency, human oversight, auditing, data governance and organisational accountability as necessary conditions for responsible AI-enabled recruitment (Hunkenschroer & Kriebitz, 2023; Kazim, Koshiyama, Hilliard, & Polle, 2021).

A second stream of research focuses on explainability, trust and general adoption. In HR, stakeholders frequently ask how AI recommendation, final ranking, rejection or prioritisation a potential candidate depends on perceived usefulness, transparency, controllability and trust in line with recruiters' acceptance of AI recommendations (Laurim, Arpacı, Prommegger, & Krcmar, 2021; Lacroux & Martin-Lacroux, 2022). Studies of human-AI and augmented hiring processes also show an interesting pattern in which reactions are more favourable when AI supports rather than replaces human decision-makers (Gonzalez et al., 2022). However, explanations according mentioned studies shouldn't be treated as a universally accepted trust-building mechanism. In human-AI decision support, explanations may increase fairness perceptions and reliance, but they do not necessarily help users identify correct versus incorrect recommendations or detect systematic bias (Schoeffer, De-Arteaga, & Kuehl, 2022). From a quality-

management perspective, this means that explainability should be understood as one quality pillar of AI decision support, not only as sufficient evidence that the AI-supported recommendation is valid, reliable or fair.

The literature also suggests that fairness of AI is not identical to user trust as it includes procedural justice, non-discrimination, validity, explainability and other features that outcomes do not disadvantage protected or vulnerable groups (Hunkenschroer & Kriebitz, 2023; Fabris et al., 2025). Within an audit-oriented research, studies stress that fairness assurance requires bias measurement, model assessment and mainly transparency of decision rules and governance mechanisms going beyond individual user perceptions (Kazim et al., 2021). Recent research in personnel selection also suggests that explanations can improve applicant's fairness perceptions under some conditions (Malin, Fleiß, Ortlieb, & Thalmann, 2025). However, this does not resolve whether HR professionals, as organisational users of AI decision support, interpret the same explanation and confidence cues as signals of trust, use-readiness and fairness in the same way. This gap opens an important research space between applicant-side fairness perception and HR-user-side perceived quality of AI-supported recruitment where this paper tries to contribute.

Existing studies have substantially advanced knowledge on different aspects of algorithmic hiring risks as well as recruiter's trust or applicant fairness perceptions, AI adoption and auditability. However, limited evidence is available on whether the same explainability-based quality cue affects different dimensions of perceived AI quality in the same way, esp. among respondents with different levels of practical AI experience. Experience level in other digital domains has been shown to be a very crucial factor and changemaker in innovative adoption processes. In the paper, we understand perceived AI quality as a socio-technical evaluation of AI-supported recruitment decision support – whether the AI output is perceived as useful, understandable, trustworthy, controllable, usable and fair. The study addresses this gap by testing whether reasoning and confidence information operate as a universal quality signal, or rather as an experience-contingent cue whose effect differs between trust in use, willingness to use and fairness confidence.

3 CONCEPTUAL MODEL AND HYPOTHESES

The conceptual model distinguishes between an explainability-based quality cue, user-level practical AI experience and broader perceived-quality constructs of AI-supported recruitment. The survey examines AI-supported recruitment broadly, while the randomised scenario focuses on one concrete and sensitive decision-support situation: an AI-generated ranking of candidates for the next stage. This scenario was selected because ranking and prioritisation are recruitment tasks where explainability is particularly important for evaluating whether the AI output is understandable, credible, usable and fair. Explanation is operationalised as the presence of brief reasoning and confidence information attached to the AI

recommendation. The design follows recruitment and Explainable Artificial Intelligence (XAI) studies that use scenario-based designs to examine reactions to AI-supported hiring decisions and explanations (Gonzalez et al., 2022; Malin et al., 2025). Practical AI experience is treated as a potential moderator because prior exposure and maturity may change whether explanation is interpreted as a useful quality signal, a source of reassurance or an indication of the complexity of AI-supported recruitment decisions.

A key conceptual distinction is made between trust in use and fairness confidence. Trust in use refers to whether HR users consider AI output credible and usable for their decision or not. On the other hand, fairness confidence refers to the extent to which the AI-supported process is perceived as justifiable and acceptable from a candidate-centred perspective and from non-discriminatory aspect. These two aspects are related but not identical dimensions of perceived AI quality. A recruiter may trust a recommendation enough to use it as support for his decision, but may still not be sufficiently convinced if it treats candidate fairly. This distinction is in line with different studies where some results show that explainability may support interpretability and user trust, but fairness assurance requires more validation, human control or auditing (Kazim et al., 2021; Hunkenschroer & Kriebitz, 2023; Fabris et al., 2025). The hypotheses of this study are summarised in Table 1.

Table 1 – Hypotheses and analytical approach

Hypothesis	Expected relationship	Analytical test
H1	Explanation and confidence information improve evaluations of AI-supported candidate-ranking recommendations in recruitment.	Direct scenario effect on D1 trust, D2 willingness to use and D3 perceived fairness; Welch/Mann–Whitney tests and OLS with HC3 robust SE
H2	Practical AI experience positively moderates the effect of explanation and confidence information on trust in AI outputs.	Interaction model for D1; OLS with HC3 robust SE and ordered logit robustness check
H3	Practical AI experience positively moderates the effect of explanation and confidence information on willingness to use AI-supported recruitment.	Interaction model for D2; OLS with HC3 robust SE and ordered logit robustness check
H4	Practical AI experience moderates the effect of explanation and confidence information on perceived fairness.	Interaction model for D3; OLS with HC3 robust SE and ordered logit robustness check
H5	Perceived usefulness, trust, explainability and human control are positively associated dimensions of perceived AI quality in AI-supported recruitment.	Cronbach’s alpha for scale reliability; Spearman correlations among constructs; OLS models for intention to use
H6	HR professionals are more likely to accept full automation without human confirmation in administrative or support recruitment tasks than in sensitive autonomous candidate-affecting decisions.	Cochran’s Q, pairwise McNemar tests and descriptive acceptance rates

4 METHODOLOGY

4.1 Research Design and Data Collection

The empirical design combines a cross-sectional questionnaire survey with a randomised scenario-based module. The questionnaire examined perceived quality of AI-supported recruitment more broadly, including usefulness, trust, explainability, human control, fairness, privacy, organisational support, resistance, intention to use and perceived outcomes. The randomised scenario module focused on a concrete recruitment decision-support situation in which respondents evaluated an AI-generated ranking of candidates for the next stage, either without explanation or with explanation and confidence information. Scenario and vignette designs are common in studies of AI-supported hiring because they allow respondents to evaluate controlled recruitment situations while keeping the decision context understandable and comparable (Gonzalez et al., 2022; Malin et al., 2025).

The survey included questions on respondent, organisational and recruitment-context items, followed by branch-specific sections for respondents with and without practical AI experience to differentiate their perceptions and attitudes. In the scenario module, respondents were randomly assigned to one of two blocks. One group evaluated an AI-generated candidate ranking without explanation. The other group evaluated the same type of ranking with a brief AI-generated reasoning for the candidate order and confidence information about the recommendation. This design allows us to test whether explanation and confidence information operate as a perceived-quality cue and whether this effect differs based on respondent's practical AI experience.

The sample is interpreted as a purposive or convenience sample of HR-related professionals and it contained elements of a snowball survey approach. This approach has no ambition to achieve a statistically representative population and therefore limits generalisation. The survey was anonymous and did not systematically identify countries. Based on the distribution channels and response languages, the sample has a substantial Central European component (80+%). A condensed profile of the sample is provided in Table 2.

Table 2 – Condensed sample profile

Characteristic	Value	Characteristic	Value
Analytic sample	N = 145	Respondent role: Recruiter/ Talent Acquisition	31.0%
Start language: Slovak/Czech version	68.3%	Respondent role: Hiring manager	20.7%
Practical AI experience: Yes	54.5%	Respondent role: AI/IT support	13.1%
Scenario: explanation + confidence	55.9%	Respondent role: HR generalist	7.6%
Applicant Tracking System (ATS) use: Yes	46.9%	Respondent role: HR manager	12.4%

Characteristic	Value	Characteristic	Value
Formal AI policy: Yes/In preparation	49.7%	Respondent role: Other	15.2%
		Organisation size: 1–49	40.0%
		Organisation size: 50–249	8.3%
		Organisation size: 250–999	29.7%
		Organisation size: 1,000+	22.1%

4.2 Measures

The experimental treatment distinguished two AI candidate-ranking scenarios: (1) ranking without explanation and (2) ranking with brief AI-generated reasoning for the candidate order and confidence information about the recommendation. In this research, we measured three scenario outcomes: trust in the AI output, willingness to use such a tool and perceived fairness to candidates. This separation reflects prior literature indicating that explanation can influence perceived fairness and adoption, but that trust and willingness to rely on AI or fairness perceptions may be responded differently (Lacroux & Martin-Lacroux, 2022; Schoeffer et al., 2022; Malin et al., 2025).

The questionnaire generally included multi-item constructs measured on seven-point agreement Likert scales: perceived usefulness (PU), perceived ease of use (PEOU), trust and reliability (TR), explainability (EXPL), human control (CTRL), fairness/non-discrimination (FAIR), privacy/security (PRIV), organisational support (ORG), resistance/threat perception (RES), intention to use (INT), perceived outcomes (OUT) and satisfaction (SAT). Fairness and privacy combine perceived concerns with governance requirements related to non-discrimination, data protection and responsible AI use.

4.3 Analytical Procedure and Diagnostics

The analysis proceeded in six steps, where the first, sample characteristics and scenario allocation were described, and randomisation balance was assessed. Second, reliability was assessed using Cronbach's alpha and construct scores were computed as row means. Construct-level analyses used complete cases ($N = 141$), while the full analytic sample was $N = 145$. Third, Spearman correlations were used to examine construct relationships because ordinal character and not assumed to be normally distributed. Fourth, Welch tests, Mann-Whitney U tests and OLS models with HC3 robust standard errors were used to estimate scenario effects and moderation by practical AI experience. Fifth, ordered logit models were used as robustness checks for ordinal scenario outcomes, and variance inflation factors (VIF) were used to assess collinearity. Finally, automation acceptance was analysed using Cochran's Q/McNemar tests, complemented by exploratory segmentation trees and an exploratory predictor heatmap based on task-specific binary models. The diagnostic and robustness checks are reported in Table 3.

Table 3 – Diagnostic and robustness checks

Check	Result	Interpretation
Construct reliability	PU = 0.903; TR = 0.818; EXPL = 0.812; CTRL = 0.825; INT = 0.925	Core scales show good to excellent internal consistency.
Scenario model VIF	max VIF = 3.71	No critical multicollinearity in interaction models.
Construct model VIF	max VIF = 2.60	No critical multicollinearity in quality-predictor models.
Heteroskedasticity	Breusch-Pagan $p < 0.01$ in OLS models	HC3 robust standard errors used.
Ordinal robustness	Ordered logit significant for D1 interaction only	Trust moderation is the most robust scenario result.
Automation-choice dependence	Cochran Q = 170.8, $p < 0.001$; McNemar $p < 0.001$	Acceptance varies sharply by task; autonomous candidate rejection is the least accepted option.

5 RESULTS

First, we assessed measurement quality and construct relationships. Then, scenario effects and moderation by practical AI experience were tested, followed by a differentiated interpretation of trust, willingness and fairness. After reporting construct-level predictors and mediation results, automation boundaries and trust-building requirements are analysed. Finally, exploratory non-linear segmentation is used to illustrate configurations associated with high intention to use AI.

5.1 Measurement Quality and Construct Relationships

The empirical design combines a cross-sectional questionnaire survey with a randomised scenario-based between-subjects module. The questionnaire examined perceived quality of AI-supported recruitment in terms of usefulness, trust, explainability, human control, fairness, privacy, organisational support, resistance, intention to use and perceived outcomes.

As shown in Table 4, the multi-item constructs demonstrate satisfactory reliability. Perceived usefulness, trust and reliability, explainability, human control and intention to use AI achieved alpha values above the conventional 0.80 threshold. Lower-reliability or conceptually mixed constructs were retained for cautious exploratory interpretation while fairness and privacy we interpreted as conceptually distinct concern/governance components.

Table 4 – Reliability and descriptive statistics for core constructs

Construct	Items	N	Cronbach's alpha	M	SD
PU	3	141	0.903	5.72	1.14
PEOU	2	141	0.777	5.59	1.05
TR	3	141	0.818	5.04	1.03
EXPL	2	141	0.812	5.46	1.01
CTRL	2	141	0.825	5.98	0.98
RES	2	141	0.616	3.91	1.37
INT	2	141	0.925	5.77	1.12
OUT	2	141	0.716	5.28	1.16
SAT	1	141	n/a	5.36	1.06

Note: Cronbach's alpha reports internal consistency; M = mean; SD = standard deviation; n/a = not applicable for a single-item measure.

The main correlations among the perceived-quality constructs are summarised in Table 5.

Table 5 – Selected Spearman correlations among perceived-quality constructs

Relationship	Spearman rho
PU – OUT	0.751***
PU – INT	0.648***
TR – SAT	0.758***
TR – INT	0.627***
EXPL – TR	0.580***
EXPL – CTRL	0.593***
CTRL – INT	0.431***
RES – TR	-0.385***

5.2 Scenario Effects and Moderation Analysis

The moderation results are reported in Table 6 and show that the direct comparison between the two scenarios did not support a simple universal explanation effect. Mean trust was slightly higher in the explanation scenario, while willingness and perceived fairness were slightly lower. However, none of these direct differences was statistically significant. The stronger statistical evidence appears in the moderation models, where practical AI experience changes how explanation and confidence information are interpreted.

Table 6 – Moderation by practical AI experience: explanation x AI experience

Outcome	Interaction coefficient (b)	SE	95% CI low	95% CI high	p	R ²
Trust in AI outputs	1.18	0.487	0.225	2.13	0.015	0.141
Willingness to use	0.930	0.427	0.093	1.77	0.029	0.161
Perceived fairness	0.675	0.499	-0.304	1.65	0.176	0.048

Note: Regression coefficients are unstandardised coefficients. Robust p-values are based on HC3 standard errors. R² denotes the coefficient of determination.

Figure 1 illustrates the scenario-response patterns by practical AI experience. The interaction was strongest for trust in AI outputs. Explanation and confidence information work as a credibility cue mainly for respondents with practical AI experience. This is consistent with research suggesting that trust in AI decision support is shaped not only by the presence of an algorithmic recommendation but also by users' ability to interpret and calibrate the recommendation (Lacroux & Martin-Lacroux, 2022; Schoeffer et al., 2022).

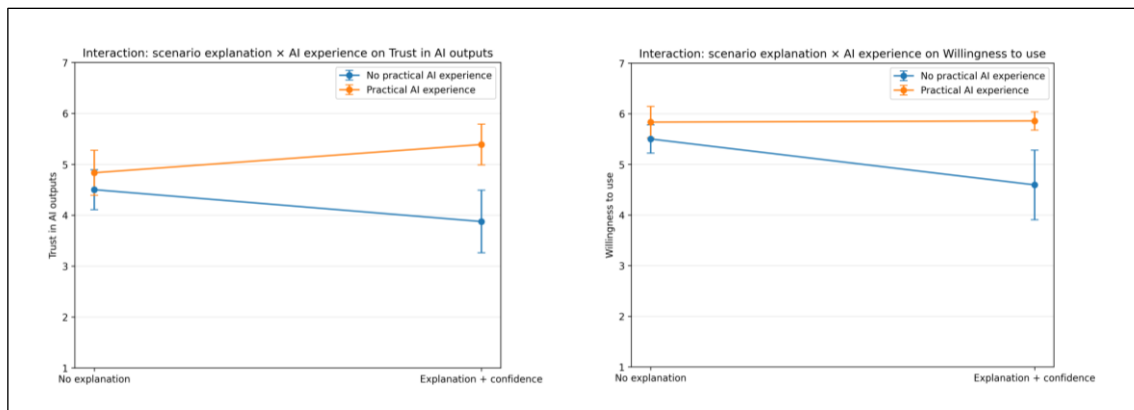


Figure 1 – Scenario response patterns for trust and willingness by practical AI experience

For respondents without practical AI experience, explanation may not yet function as reassurance; it may instead highlight the complexity and uncertainty of AI-supported ranking. A differentiated interpretation of these scenario outcomes is provided in Table 7.

Table 7 – Differentiated interpretation of scenario outcomes

Outcome	Interpretation	b	p	Ordinal p	Managerial meaning
Trust in AI outputs (D1)	Trust in use / credibility of AI output	1.18	0.015	0.007	Experienced users interpret explanation and confidence as a credibility cue that makes the ranking more usable as decision support.
Willingness to use (D2)	Adoption readiness	0.930	0.029	0.349	A similar pattern appears, but robustness is weaker under ordinal modelling; interpreted as a managerial pattern rather than the strongest statistical result.
Perceived fairness (D3)	Fairness confidence / perceived non-discrimination	0.675	0.176	0.442	The same explanation cue does not statistically convince respondents that the ranking is fair; fairness remains a separate assurance issue.

The pattern is managerially important. An experienced HR professional may say: “I understand why AI ranked this candidate higher. The explanation and confidence score help me trust the recommendation enough to use it as decision support.” However, the same professional may still add: “This explanation does not convince me that the ranking is fair to all candidate groups.” The first statement concerns trust in use and operational understanding; the second concerns fairness confidence. The non-significant fairness moderation therefore does not simply mean that fairness is an uninteresting outcome. It suggests a stronger implementation warning: explanation and confidence information may make an AI recommendation more usable without making it perceived as fair. Fairness therefore requires independent validation, bias testing and auditability.

5.3 Construct-level Regression and Mediation Analysis

The multivariable regression results are presented in Table 8.

Table 8 – Construct-level regression models with HC3 robust standard errors

Model	Predictor	b	SE	p	R ²	N
TR model	EXPL	0.350	0.095	0.000	0.475	141
TR model	PU	0.184	0.068	0.007	0.475	141
TR model	PEOU	0.279	0.110	0.011	0.475	141
TR model	ai_experience	0.178	0.134	0.183	0.475	141
INT model	TR	0.472	0.105	0.000	0.636	141
INT model	PU	0.255	0.088	0.004	0.636	141
INT model	CTRL	0.227	0.080	0.004	0.636	141
INT model	ai_experience	0.547	0.138	0.000	0.636	141

Model	Predictor	b	SE	p	R ²	N
INT extended	TR	0.481	0.107	0.000	0.641	141
INT extended	PU	0.269	0.083	0.001	0.641	141
INT extended	CTRL	0.242	0.075	0.001	0.641	141
INT extended	RES	0.058	0.043	0.175	0.641	141
INT extended	ai_experience	0.510	0.136	0.000	0.641	141

Note: Regression coefficients are unstandardised coefficients. Robust p-values are based on HC3 standard errors. R² denotes the coefficient of determination.

Construct-level models also support the multi-dimensional interpretation of perceived AI quality. Trust and reliability were predicted by parameters of explainability, perceived usefulness and ease of use. The parameter intention to use AI was predicted by trust, perceived usefulness, human control and practical AI experience. Bootstrap mediation, as presented in Table 9, indicated that trust really partially mediated the association between explainability and intention to use AI, controlling for perceived usefulness and AI experience.

A similar indirect association was found also for human control through trust. These findings support the interpretation that trust is an important transmission mechanism in this decision making area, but it is not the only requirement for quality-oriented implementation.

Table 9 – Bootstrap mediation results

Path	N	Indirect effect	95% CI low	95% CI high	Bootstrap p
EXPL → TR → INT	141	0.187	0.094	0.294	0.000
CTRL → TR → INT	141	0.130	0.042	0.195	0.008

5.4 Automation Acceptance and Implementation Requirements

The practical module shows a sharp boundary between accepted automation and rejected autonomous decision-making. Full automation without human confirmation was most often accepted for job-posting text generation, candidate outreach and interview scheduling, but was rarely accepted for ranking without approval and almost never accepted for automatic candidate rejection. This is aligned with existing literature on augmented hiring, where these studies show more favourable reactions of users when AI assists humans rather than replaces them in sensitive candidate-affecting selection decisions through full automation (Gonzalez et al., 2022; Hunkenschroer & Kriebitz, 2023).

Figure 2 summarises automation acceptance and the main trust-building requirements.

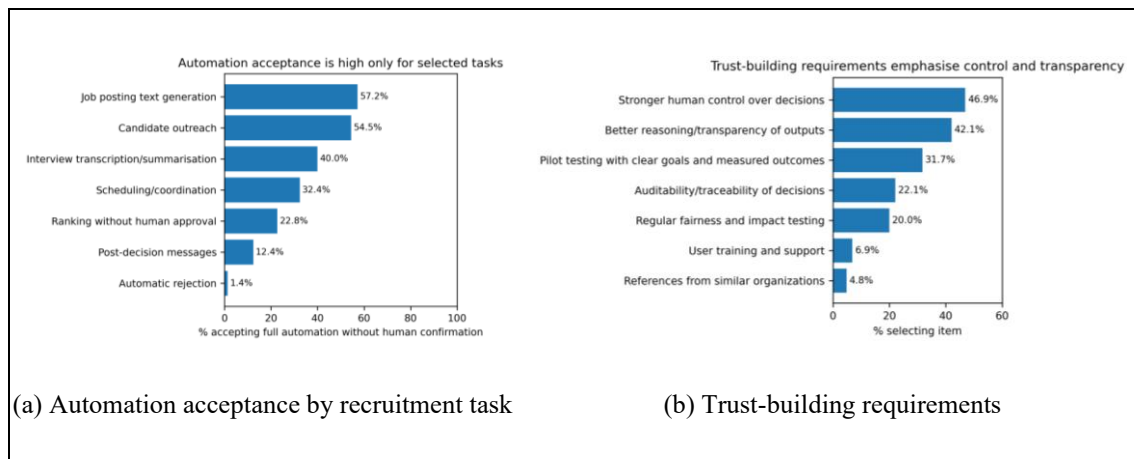


Figure 2 – Automation acceptance and trust-building requirements

Using Cochran's Q we confirmed that acceptance differed significantly across tasks. The McNemar comparison between job-posting automation and automatic candidate rejection was highly significant. The most frequently selected trust-building requirement was stronger human control, followed by reasoning/transparency and pilot testing with measured outcomes.

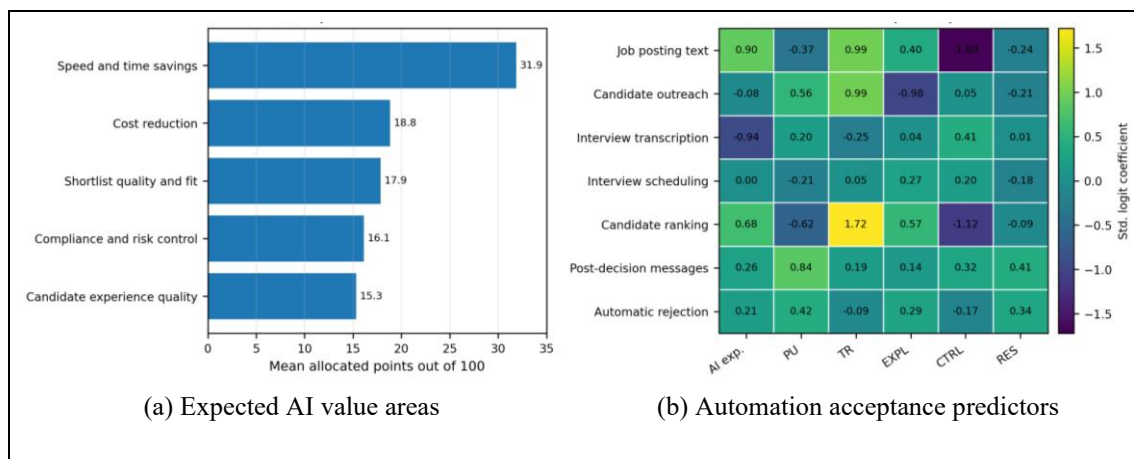


Figure 3 – Expected value allocation and exploratory predictors of automation acceptance

These tests support general findings in a human-centred quality-management interpretation in which AI is acceptable when it augments human work and remains controllable, but less acceptable when it replaces accountability in sensitive candidate-affecting decisions.

When analysing an expected value allocation (Figure 3a), we have found that respondents see AI value primarily in speed and time savings, but not only in efficiency. Shortlist quality, cost reduction, compliance/risk control and candidate experience also received substantial allocations. The exploratory predictor heatmap in Figure 3b summarises standardised coefficients from task-specific regularised logistic models of full-automation acceptance. It illustrates how

automation preferences are associated with perceived usefulness, trust, explainability, human control, resistance and practical AI experience across recruitment tasks.

5.5 Exploratory segmentation analysis

A parsimonious classification tree, presented in Figure 4, was estimated as an exploratory visual explanation of high intention to use AI ($INT \geq 6$). The tree should be interpreted cautiously due to the sample size, but it provides a useful managerial segmentation. The first-strongest split was human control. High intention was concentrated among respondents who perceived both strong human control and high usefulness. In the intermediate group with high control but lower usefulness, the explanation-and-confidence scenario was associated with a higher share of high intention, suggesting that explanation may help only after minimum control conditions.

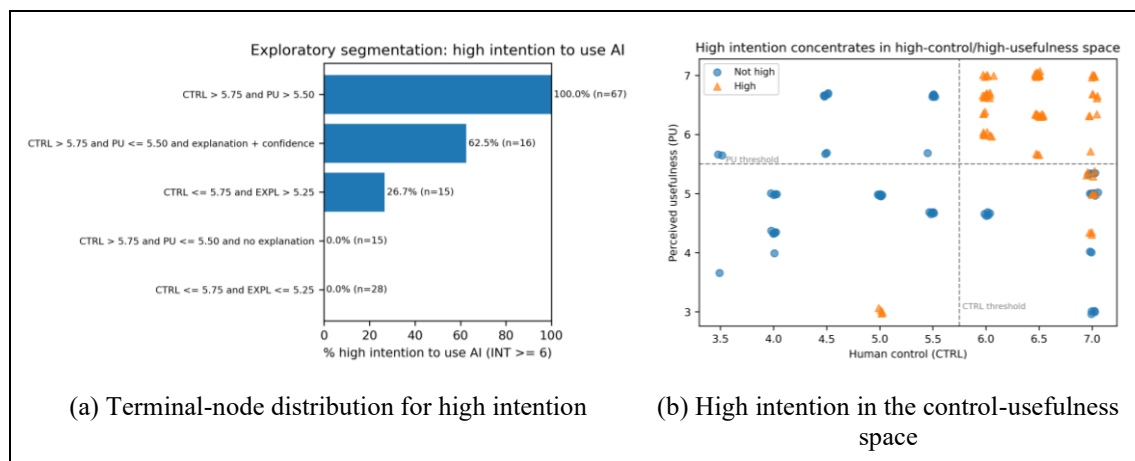


Figure 4 – Exploratory segmentation of high intention to use AI

6 DISCUSSION

Before interpreting the findings in managerial terms, Figure 5 summarises the empirically supported and unsupported relationships. The scheme is presented as an ex-post synthesis of the results, not as an a priori structural model.

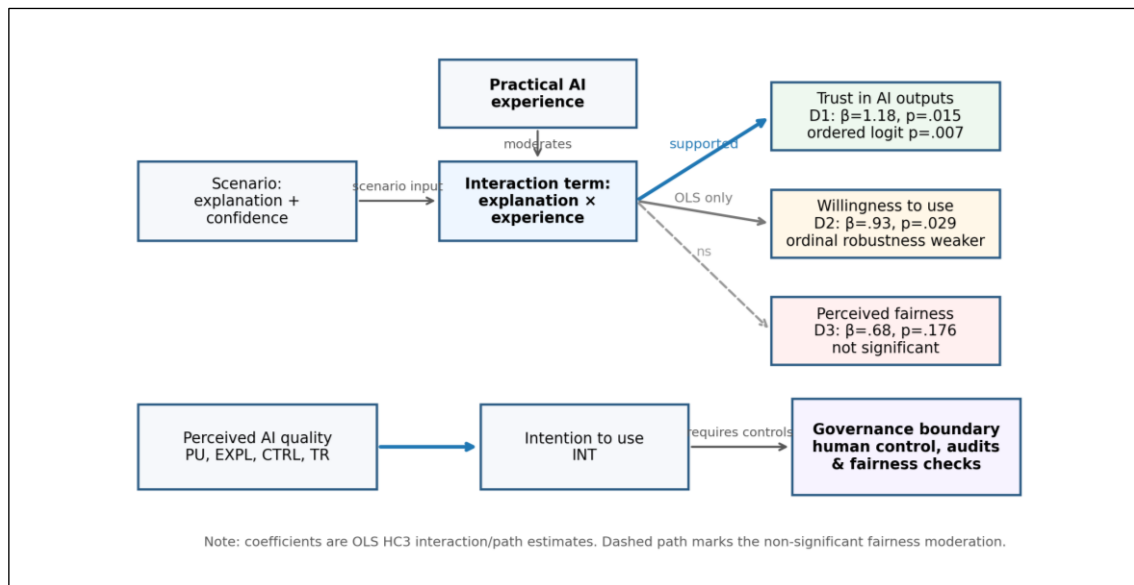


Figure 5 – Empirical synthesis of supported and unsupported relationships

6.1 Interpretation of Moderation Effects

The results challenge the assumption that adding explanation and confidence information automatically increases perceived quality of AI-supported recruitment. The direct scenario effect was not significant. The key statistical result is the moderation pattern. It supports the argument, that practical AI experience changes how explanation is interpreted, especially for trust in AI outputs. It means, explanation is not a universal quality signal. It becomes a quality cue only when users can relate it to their own mental model, domain knowledge and calibration of AI limitations (Schoeffer et al., 2022; Lacroux & Martin-Lacroux, 2022).

This finding has practical implications for AI implementation. Explainability should not be deployed as a stand-alone feature or as a pro-forma transparency layer. It should be embedded in training, examples, calibration exercises and pilot feedback so that users can learn when to follow AI recommendations, when to challenge them and finally how to interpret confidence information. This corresponds to quality-oriented implementation logic from different studies, where AI-enabled decision support requires people capability, leadership, process integration and data governance, not only some technical functionality (Sony et al., 2020; Carvalho et al., 2024).

6.2 Fairness as a Distinct Assurance Dimension

The most important conceptual contribution concerns the difference between trust in use and fairness confidence. Trust and willingness to use are user-side adoption perceptions: they indicate whether HR professionals find AI recommendations credible, useful and operationally acceptable. Fairness is a harder assurance dimension: it concerns whether the system is perceived as valid, non-discriminatory and procedurally just from the candidate's point of view. The

empirical result does not prove that respondents believed the AI was unfair. It shows a more precise and practically relevant point: the explanation and confidence cue did not convince them that the AI-supported ranking was fair. Algorithmic hiring research similarly shows that fairness cannot be inferred from local explanation alone; it requires auditing, bias testing, governance and evidence across groups and cases (Kazim et al., 2021; Fabris et al., 2025).

A practical example illustrates the point. An experienced HR user may say: “I understand why AI ranked this candidate higher. The explanation and confidence score help me trust the recommendation enough to use it as decision support.” However, the same user may not be willing to say: “I am convinced that this ranking is fair to all candidate groups.” This difference is not only a knowledge gap. It is a fairness-confidence gap. It captures the empirical pattern of the study: explanation can support experienced-user trust, while perceived fairness remains insufficiently improved by the same explanation cue.

An AI system may explain that a candidate was ranked higher because of relevant experience, skill overlap and role fit. This may help an experienced HR user understand the ranking and accept it as a decision-support input in the decision making process. However, we have to be careful also in this aspect, as this doesn’t demonstrate whether years of experience penalise career breaks, whether prior employer prestige proxies social background or whether model errors differ across groups. This should be examined in further research on larger samples to identify more complex patterns. In quality-management terms, AI explanations improve local interpretability of an output, whereas fairness confidence requires a system-level quality control and governance process.

6.3 Human Control and Acceptable Automation Boundaries

The results on full automation opportunities show that respondents do not generally reject AI automation. They differentiate between tasks with lower decision-finality and tasks that directly affect candidate outcomes and are more sensitive. Acceptance was highest for job-posting text generation and candidate outreach, but the interpretation of job-posting automation must remain cautious. If respondents understood the task as drafting or stylistic support based on human-defined requirements, the risk is lower. If they understood it as AI autonomously defining role requirements and publishing vacancies, the risk would be much higher. The current item cannot fully separate these interpretations.

The almost complete rejection of automatic candidate rejection without human approval is one of the clearest managerial results. It indicates that HR professionals draw a boundary around sensitive autonomous organisational decisions. This boundary is consistent with the trust-building responses.

6.4 Implications for Quality-oriented AI Implementation in HR

We can find four recommendations for the world of HR practice. AI implementation in HR should begin with tasks where perceived value is high and decision-finality is lower. Organisations should clearly distinguish between drafting support and autonomous publication or rejection to increase trust and willingness of AI users. AI explanations or reasoning features should be also designed differently for different levels of user maturity. Experienced users may benefit from concise reasoning and confidence indicators, while less experienced users may require training, use cases or examples and calibration feedback to improve their awareness and readiness to work with AI tools efficiently.

Together, we consider as very interesting and critical result in governance implication area. Fairness assurance must be separated from trust-building. Explanations and confidence scores may make AI recommendations more understandable and usable for experienced HR users, but they do not by themselves demonstrate that the underlying ranking is fair. This creates a practical risk that organisations may become comfortable relying on AI-supported recommendations before they have sufficient evidence that the systems are fairness-assured. Therefore, explainability should be complemented by independent bias testing, auditability and fairness governance. Impact of these results can be significant, as it can develop bias in AI hype era and can bring high risks for whole society and in AI adoption area generally.

Finally, AI-supported recruitment should be managed as a quality-controlled organisational process in which efficiency, decision quality, human control, candidate experience, auditability and social responsibility are treated as interconnected quality requirements.

7 CONCLUSION

This paper examined perceived quality of AI-supported recruitment through the lens of explainability, trust and human control. Based on relevant literature, the study shows that explanation and confidence information do not create a uniform improvement in evaluations of AI candidate ranking. Instead, they operate as explainability-based quality cues whose effect depends on practical AI experience, with the strongest and most robust effect observed for trust in AI outputs.

The research further shows that intention to use AI is associated with trust, perceived usefulness, human control and prior AI experience. Regarding automation acceptance, the results show that HR professionals accept AI support mainly in selected non-sensitive support tasks and reject autonomous candidate rejection because of the risk of discriminatory or unfair results. These results support a configurational view of AI quality in recruitment. Explainability matters, but it matters together with user experience, usefulness, control and governance. The distinctive finding is that explanation may strengthen trust in use without

increasing fairness confidence. Fairness remains a distinct assurance problem that requires independent validation beyond explanation.

We encourage researchers in future studies to test these relationships in larger and more representative samples, distinguish different levels of AI support in job-ad generation and connect perceived quality with objective indicators of model performance, fairness, bias and candidate outcomes. Specific attention should be paid to the trust–fairness gap identified in this study, especially to whether explanation-induced usability can encourage premature organisational reliance on AI-supported decisions before fairness and group-level validity have been independently audited.

ACKNOWLEDGEMENTS

No funding acknowledgement is reported in the double-blind review version.

REFERENCES

- Carvalho, A. M., Dias, A. R., Dias, A. M., & Sampaio, P. (2024). The Quality 4.0 roadmap: Designing a capability roadmap toward quality management in Industry 4.0. *Quality Management Journal*, 31(2), 117–137. <https://doi.org/10.1080/10686967.2024.2317478>
- Chen, Z. (2023). Collaboration among recruiters and artificial intelligence: Removing human prejudices in employment. *Cognition, Technology & Work*, 25(1), 135–149. <https://doi.org/10.1007/s10111-022-00716-0>
- Fabris, A., Baranowska, N., Dennis, M. J., Graus, D., Hacker, P., Saldivar, J., Zuiderveen Borgesius, F., & Biega, A. J. (2025). Fairness and bias in algorithmic hiring: A multidisciplinary survey. *ACM Transactions on Intelligent Systems and Technology*, 16(1), Article 16. <https://doi.org/10.1145/3696457>
- Gonzalez, M. F., Liu, W., Shirase, L., Tomczak, D. L., Lobbe, C. E., Justenhoven, R., & Martin, N. R. (2022). Allying with AI? Reactions toward human-based, AI/ML-based, and augmented hiring processes. *Computers in Human Behavior*, 130, 107179. <https://doi.org/10.1016/j.chb.2022.107179>
- Hunkenschroer, A. L., & Kriebitz, A. (2023). Is AI recruiting (un)ethical? A human rights perspective on the use of AI for hiring. *AI and Ethics*, 3(1), 199–213. <https://doi.org/10.1007/s43681-022-00166-4>
- Hunkenschroer, A. L., & Luetge, C. (2022). Ethics of AI-enabled recruiting and selection: A review and research agenda. *Journal of Business Ethics*, 178(4), 977–1007. <https://doi.org/10.1007/s10551-022-05049-6>
- Kazim, E., Koshiyama, A. S., Hilliard, A., & Polle, R. (2021). Systematizing audit in algorithmic recruitment. *Journal of Intelligence*, 9(3), 46. <https://doi.org/10.3390/jintelligence9030046>

- Lacroux, A., & Martin-Lacroux, C. (2022). Should I trust the artificial intelligence to recruit? Recruiters' perceptions and behavior when faced with algorithm-based recommendation systems during resume screening. *Frontiers in Psychology*, 13, 895997. <https://doi.org/10.3389/fpsyg.2022.895997>
- Laurim, V., Arpaci, S., Prommegger, B., & Krcmar, H. (2021). Computer, whom should I hire? Acceptance criteria for artificial intelligence in the recruitment process. In *Proceedings of the 54th Hawaii International Conference on System Sciences* (pp. 5495–5504). <https://doi.org/10.24251/HICSS.2021.668>
- Liu, H. C., Liu, R., Gu, X., & Yang, M. (2023). From total quality management to Quality 4.0: A systematic literature review and future research agenda. *Frontiers of Engineering Management*, 10(2), 191–205. <https://doi.org/10.1007/s42524-022-0243-z>
- Malin, C., Fleiß, J., Ortlieb, R., & Thalmann, S. (2025). Rejected by an AI? Comparing job applicants' fairness perceptions of artificial intelligence and humans in personnel selection. *Frontiers in Artificial Intelligence*, 8, 1671997. <https://doi.org/10.3389/frai.2025.1671997>
- Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2020). Mitigating bias in algorithmic hiring: Evaluating claims and practices. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 469–481). <https://doi.org/10.1145/3351095.3372828>
- Schoeffer, J., De-Arteaga, M., & Kuehl, N. (2022). *Explanations, fairness, and appropriate reliance in human-AI decision-making*. Retrieved from arXiv database. <https://doi.org/10.48550/arXiv.2209.11812>
- Sony, M., Antony, J., & Douglas, J. A. (2020). Essential ingredients for the implementation of Quality 4.0. *The TQM Journal*, 32(4), 779–793. <https://doi.org/10.1108/TQM-12-2019-0275>
- Will, P., Krpan, D., & Lordan, G. (2023). People versus machines: Introducing the HIRE framework. *Artificial Intelligence Review*, 56(2), 1071–1100. <https://doi.org/10.1007/s10462-022-10193-6>

ABOUT AUTHORS

Coi Tran ORCID: 0000-0002-7947-0907 (C.T.) – Doctoral researcher, his research focuses on applications of AI in HR, Department of Banking and Investment, Faculty of Economics, Technical University of Kosice, Košice, Slovakia, e-mail: coit@fpt.com.

Radoslav Delina ORCID: 0000-0003-0412-9840 (R.D.) – Prof., his research focuses on complex decision automation and intelligence in business processes, Department of Banking and Investment, Faculty of Economics, Technical University of Kosice, Košice, Slovakia, e-mail: radoslav.delina@tuke.sk.

AUTHOR CONTRIBUTIONS

Conceptualisation, R.D.; Methodology, R.D.; Formal analysis, C.T.; Investigation, C.T.; Original draft preparation, C.T.; Review and editing, R.D. and C.T.; Supervision, R.D.; Project administration, R.D.

CONFLICTS OF INTEREST

The authors declare that they have no conflict of interest. The funders had no role in the design of the study, in the collection, analysis, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

DISCLOSURE OF ARTIFICIAL INTELLIGENCE ASSISTANCE

The authors acknowledge the use of artificial intelligence tools in the preparation of this manuscript. Specifically, Asta.Allen.ai and ChatGPT by OpenAI was utilised to support linguistic refinement and summarisation of background literature. All AI-assisted content was carefully reviewed, edited, and validated by the authors to ensure scholarly integrity, originality, and compliance with ethical standards. The authors affirm full responsibility for the content of the manuscript, including any material generated or supported by artificial intelligence.



© 2026 by the authors. Submitted for possible open-access publication under the Terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).