

Harnessing Artificial Intelligence (AI) to Overcome Cognitive Bias and Bounded Rationality: Advancing Human Resource Management through Co-evolution

Yukiko Nakagawa,
Aoyama Gakuin University

ABSTRACT

This study examines how human-AI (Artificial Intelligence) collaboration can address the challenges posed by cognitive biases and bounded rationality in Human Resource Management (HRM). In today's complex environment, human judgment is influenced by factors such as cognitive bias (systematic errors in thinking) and bounded rationality, which is the tendency to make satisfactory rather than optimal decisions due to cognitive limitations. These issues can compromise the fairness and efficiency of HRM processes such as recruitment and performance evaluation.

AI offers a solution by enabling objective, data-driven decision-making. However, AI systems can also reinforce biases present in their training data. The paper proposes a framework of "dual dynamics" where AI acts as a cognitive partner, not just a tool, to enhance human judgment by identifying bias patterns and simulating alternative scenarios. This approach integrates AI's analytical power with human empathy and ethical reasoning to foster more equitable and rational decision-making.

Case studies in human resources analytics illustrate this partnership. In recruitment, companies like SoftBank and Unilever use AI to reduce bias and increase efficiency in tasks such as resume screening and video interviews, while human oversight remains for final decisions. In talent development,

Coca-Cola uses AI to recommend career paths and learning opportunities, helping employees navigate complex decisions. The paper concludes that AI's transformative potential in HRM hinges on the co-evolution of humans and AI, with each leveraging the other's strengths and offsetting their respective weaknesses. Future research should explore ways to improve AI's understanding of emotional cues and to strengthen human trust in AI-driven decision-making.

Key words:

Cognitive Bias, Bounded Rationality, Artificial Intelligence, Human Resource Management, Dual Dynamics

1. Introduction

Contemporary society is characterized by an overwhelming abundance of information and rapid external competitive environmental shifts, resulting in unprecedented complexity in both individual decision-making and organizational governance. Within this milieu, psychological constructs such as cognitive bias and bounded rationality exert profound influences on human judgment. Cognitive bias refers to systematic deviations in information processing, exemplified by phenomena such as confirmation bias and loss aversion. Bounded rationality, as conceptualized by Herbert Simon, denotes a decision-making framework that acknowledges the limitations of human cognition, whereby individuals often settle for satisfactory rather than optimal solutions when confronted with excessive information and alternatives.

These psychological constraints are particularly salient in the domain of Human Resource Management (HRM), where they may compromise the fairness and efficiency of decisions related to recruitment, performance evaluation, and talent development. Concurrently, the rapid advancement of Artificial Intelligence (AI) presents a promising avenue for mitigating such biases by enabling data-driven, objective decision-making. However, AI systems are not immune to the biases embedded within their training data, and may inadvertently reinforce existing inequities (Hall et al., 2022). Thus, the uncritical **conceptual framework in which AI functions as a cognitive partner rather than a mere analytical tool risks obscuring these limitations. The hypothesis suggests that AI systems—when designed with contextual awareness, transparency, and cultural sensitivity—can enhance human judgment by identifying**

bias patterns, broadening decision frames, and simulating alternative scenarios. Within HRM, such capabilities show promise for improving talent acquisition, skill development, and organizational sound checks. The study further explores implications for leadership, governance, and cross-cultural communication, advocating inclusive and adaptive approaches to AI integration.

The deployment of AI may engender novel challenges including concerns about data privacy and the potential for increased social inequality.

This study explores the potential and limitations of human-AI co-evolution in addressing these issues. Drawing upon Nonaka's (2024) concept of **"dual dynamics"**, which advocates integrative rather than dichotomous thinking, the paper examines how the synthesis of AI's analytical capabilities with human empathy and ethical reasoning may foster more equitable and rational decision-making within HRM.

2. Review of Prior Research

2.1 Theoretical Foundations of Cognitive Bias and Bounded Rationality

The constructs of cognitive bias and bounded rationality have been extensively examined across various disciplines, including psychology, behavioral economics, and decision theory. Foundational contributions by Kahneman and Tversky (1979), notably their Prospect Theory, elucidated deviations from expected utility theory by highlighting systematic biases such as loss aversion and anchoring. Their work provided a robust framework for understanding how individuals make judgments under uncertainty, often relying on heuristics that lead to

predictable errors.

Simon's (1955) theory of bounded rationality challenged the classical economic assumption of perfect rationality, positing that individuals operate under cognitive constraints and thus pursue "satisficing" solutions. This paradigm has become instrumental in explaining decision-making behavior in complex, real-world contexts.

These theoretical insights have significant implications for HRM. For instance, **representativeness bias** may lead interviewers to overestimate candidates' competencies based on superficial attributes such as appearance or speech patterns. Tversky et al. (1974) systematically categorized such heuristics, demonstrating that individuals often rely on intuitive judgments rather than probabilistic reasoning. Similarly, **availability bias**, wherein salient memories disproportionately influence evaluations, can distort hiring decisions. Olian et al. (1991) observed that even experienced interviewers are susceptible to non-verbal cues and implicit biases, which can undermine their objectivity.

Gigerenzer (2018) offered a counter-perspective, arguing that heuristics may function as adaptive tools in uncertain environments, challenging the notion that cognitive biases are inherently irrational.

2.2 Human-AI Collaboration: Ethical Considerations and Complementarity

AI's capacity for large-scale data analysis enables the identification of cognitive biases and enhances decision-making objectivity, thereby offering a potential remedy to bounded rationality (Selbst et al., 2018). Nonetheless, AI systems may internalize and perpetuate biases present in historical data

(Hall et al., 2022). Zowghi and Bano (2024) caution against the inadvertent marginalization of individuals or groups through algorithmic processes.

Consequently, ethical concerns such as bias amplification and algorithmic opacity necessitate vigilant human oversight. The final responsibility for decisions must remain with human agents. Hybrid decision-making models that integrate AI and human judgment are thus essential for effective HRM.

3. Case Studies in HR Analytics

HRM represents a domain in which human-AI collaboration has achieved notable progress. HR analytics, as defined by Fitz-enz (2010) and Marler & Boudreau (2017), involves the application of data-driven methodologies to support HR decision-making, with AI significantly enhancing its capabilities.

3.1 Mitigating Bias and Enhancing Efficiency in Recruitment

AI contributes to both the efficiency and fairness of recruitment processes. Rukadikar et al. (2025) highlight the utility of AI-driven tools such as resume screening and chatbot-based preliminary interviews in saving time and cost while enhancing objectivity.

- **SoftBank:** Implemented AI systems to reduce video interview screening time by approximately 70%. A hybrid model is employed in which AI determines qualification thresholds, while final decisions are made by HR personnel.
- **Unilever:** Utilizes game-based assessments from Pymetrics and AI video analysis from HireVue to eliminate biases associated with educational and professional backgrounds, thereby facilitating the identification of diverse talent.

These cases underscore the role of AI as a complement to human judgment, enabling more equitable and efficient recruitment practices.

3.2 Contributions to Talent Development and Employee Engagement

AI is increasingly employed in post-recruitment contexts, including talent development and employee engagement.

- **Coca-Cola:** Deploys the Eightfold AI platform to recommend career paths and learning opportunities based on employee skill profiles. AI chatbots automate routine inquiries, providing continuous support and enhancing psychological safety. This application exemplifies how AI can assist in overcoming bounded rationality by guiding employees through complex decision landscapes.
- **Organizational Sound Checks:** Maltseva (2023) notes that career consultation chatbots can effectively capture latent employee sentiments. Organizations such as Airbnb and GE utilize chatbots to analyze emotional states in real time, facilitating early problem detection.

The author hypothesizes that this mechanism parallels the allegory of “The King’s Ears Are Donkey’s Ears,” wherein AI provides a secure space for anonymous disclosure, yielding valuable insights into organizational health.

3.3 Ethical Challenges and the Role of Human Oversight

AI deployment entails significant ethical considerations (Zowghi & Bano, 2024), including the risk of opaque decision-making and bias reinforcement. To address these concerns, frameworks such as **Explainable AI** (Garcia et al., 2025) are essential for ensuring transparency and interpretability.

Ultimately, the concept of **dual dynamics**—integrating AI’s analytical rigor with human ethical sensibilities—is indispensable. While AI enhances objectivity and efficiency, human oversight remains critical for ethical governance.

4. Conclusion and Future Directions

This study has examined how human-AI collaboration may address the challenges posed by cognitive bias and bounded rationality in HRM. AI offers objective data and operational efficiency, complementing human judgment.

The transformative potential of AI in HRM hinges not merely on technological adoption, but on the co-evolution of human and artificial agents—leveraging mutual strengths and compensating for respective limitations. The **dual dynamics** approach proposed in this paper is posited as essential for the realization of sustainable organizations and societies.

Future research should investigate methods for enhancing AI’s capacity to interpret complex emotional and contextual cues, as well as mechanisms for fostering human understanding and trust in AI decision-making processes.

References

- Adamopoulou, E., & Moussiades, L. (2020). An Overview of the Use of Chatbots in Human Resources. *Journal of Human Resources Management Research*, 10(1), 1-18.
- Delgado, S., Villamarin, A., & Insuasti, J. (2025). AI-Powered Chatbots in Organizations: A Systematic Literature Review. *Journal of Information Systems Engineering and Management*, 10 (22s), 1-12.

- Didar Zowghi & Muneera Bano (2024). AI for All: Diversity & Inclusion in AI. *AI & Ethics*, 4, 873–876.
- Fitz-enz, J. (2010). *The New HR Analytics: Predicting the Economic Value of Your Human Capital Investments*. AMACOM, 1-300.
- Garcia, J., et al. (2025). Explainable AI in Labor Market Applications. *Proceedings of the 17th International Conference on Agents and Artificial Intelligence (ICAART 2025)*, 3, 1450-1457.
- Gigerenzer, G. (2018). The Bias Bias: How the Cognitive Sciences Ignore the Success of Heuristics in Real-World Decisions. *Review of General Psychology*, 22(4), 304-315.
- Gok, S., & Gungor, O. (2022). A review of AI-powered chatbots in the workplace: Implications for employee engagement. *International Journal of Contemporary Management*, 11(1), 1-15.
- Hall, M., Maaten, L., Cutafdon, L., & Adcock, A. (2022). A Systematic Study of Bias Amplification. *arXiv:2201.11706v2*, 1-12.
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263-291.
- Madanchian, M. (2024). From Recruitment to Retention: AI Tools for Human Resource Decision-Making. *Applied Science*, 24(24), 11750.
- Maltseva, I. (2023). *Addressing bias and privacy challenges when using AI in HR*. Workable. <https://workable.com/blog/ai-bias-privacy>
- Marler, J. H., & Boudreau, J. W. (2017). An Evidence-Based Review of HR Analytics, People Analytics, and Workforce Analytics Research. *Journal of Human Resource Management*, 10(1), 1-18.

- Mayer, R. C., & Davis, J. H. (1999). "The Illusion of Anonymity: A Qualitative Study of Employee Perceptions of Organizational Surveys." *Academy of Management Journal*, 42(5), 458-472.
- Morwitz, V. G., & Fitzsimons, G. J. (2004). The mere-measurement effect: The effect of questionnaire answers on subsequent behavior. *Journal of Consumer Research*, 31(2), 263-272.
- Nonaka, Ikujiro (2024). "The concept of dual dynamics serves to strengthen the co-evolutionary relationship between human cognition and artificial intelligence." *Diamond Quarterly*, 2025, Summer, 28-29.
- Olian, J. D., & Rynes, S. L. (1991). "Determinants of Effective Interviewing: An Examination of Experienced Interviewers." *Journal of Applied Psychology*, 76(4), 578-587.
- Porter, S. R., Whitcomb, M. E., & Weitzer, H. (2004). Multiple surveys of students and survey fatigue: A longitudinal analysis. *New Directions for Institutional Research*, 2004(121), 63-73.
- Prentice, C., & Loureiro, S. M. C. (2024). The impact of AI chatbots on employee wellbeing and performance. *Journal of Service Theory and Practice*, 18(1), 1-15.
- Rukadikar, A., Khandelwal, K., & Warriar, U. (2025). Reimagining recruitment: traditional methods meet AI interventions- A 20-year assessment (2003–2023). *Cogent Business Management*, 12(1), 1-20.
- Schein, E. H. (1985). *Organizational Culture and Leadership: A Dynamic View*. Jossey-Bass, 1-325.
- Selbst, A. D., et al. (2018). Fairness in Machine Learning: Lessons from Political Philosophy. *Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency*, 1-14.

- Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, 69(1), 99-118.
- Tversky, A., & Kahneman, D. (1974). "Judgment Under Uncertainty: Heuristics and Biases." *Science*, 185(4157), 1124-1131.
- Yuval Noah Harari (2025). NEXUS: A Human History of Information, Volume I – The Networked Species. *Kawadeshboboshinsha*.