

Human-AI collaboration fostering innovation performance: A theoretical framework

Thu Thao Nguyen, Trung Quang Nguyen, Giang Truong Hoang, Alvedi Sabani

RMIT University, The Business School

Email address: S3987152@student.rmit.edu.au

Human-AI collaboration fostering innovation performance: A theoretical framework

Summary

This paper investigates how human-AI collaboration contributes to innovation performance in organisational contexts. In light of task-technology fit theory (TTF), we develop a conceptual model that highlights the relationship between human-AI collaboration and innovation outcomes (i.e., employee's innovative behavior and firm innovation performance) and the mediating mechanism underlying these relationships. The paper also presents the methodology to be used in an empirical study to test the proposed model. This paper provides insights into how human-AI collaboration is linked to innovation outcomes and discusses the expected theoretical and practical contributions of the empirical study.

Keywords: Human-AI collaboration, innovative behaviour, innovation performance, psychological safety, thriving at work, task-technology fit

Track: Developmental Papers

Word count: 2039

Introduction

The rapid development and application of artificial intelligence (AI) technologies have urged scholars to explore its impacts on businesses and people. According to Constantinides et al. (2024), AI refers to a machine that runs large dense neural network training algorithms to mimic cognitive functions associated with human minds. These algorithms, often trained on unstructured data like photos, audio, or text, significantly enhance machines' ability to learn, reason, categorise, and understand. With potential applications in nearly every field, AI has become an essential component of many business models and a critical strategic factor for numerous industries and governments worldwide (Brougham and Haar, 2018; Pereira *et al.*, 2023). For example, AI expert systems can efficiently analyse large amounts of data and provide accurate insights, thereby enhancing the precision of human resources decisions in the workplace. These systems can assist in the automated pre-selection of candidates, ensuring compatibility with current teams and potential future colleagues (Pereira *et al.*, 2023). In marketing and retail, AI technology is leveraged through big data analytics to create personalised consumer profiles and forecast purchasing behaviours. In healthcare, AI has the potential to enhance patient care and diagnosis, especially in radiology, by improving the interpretation of medical imaging (Issa, Jaber and Lakkis, 2024).

According to Seeber et al. (2020), the incorporation of AI as an autonomous and interdependent team member has necessitated changes in team roles and responsibilities, prompting the development of new norms and processes to accommodate this new member. Given the inherent challenges to the future of work with AI, it is vital and timely to comprehend how humans and AI collaborate and excel with each other. According to Loske and Klump (2021), human-AI collaboration involves leveraging the unique strengths of both parties and evolving together through complimentary enhancements. Higher levels of human-AI collaboration indicate a willingness among users to enhance AI performance by integrating their intelligence while the AI system continues to learn from users to augment human intelligence simultaneously (Jiang *et al.*, 2022). One way to realise this synergistic relationship is by combining AI's speed in collecting and analysing information with humans' superior intuitive judgment and insight. Through collaborative efforts and building on each other's strengths rather than competing, people and AI can achieve mutually beneficial results, fostering synergy among all involved parties.

Although several studies have been conducted to explore human-AI collaboration, there is a gap in understanding how human-AI collaboration facilitates innovative outcomes at both individual and firm levels (Bankins *et al.*, 2023; Gama and Magistretti, 2023; Cai, Xiao and Ji, 2023). To address this gap, our study aims to develop a theoretical framework based on the premises of task-technology fit (TTF) theory. Specifically, our framework proposes that the collaboration between humans and AI will promote employee's innovative behaviour, which in turn promotes firm innovation performance. Furthermore, our study posits that employees' psychological safety and thriving at work are potential mediators that translate human-AI collaboration into innovation outcomes (i.e., employees' innovative behaviour and firm innovation performance).

The study will primarily focus on the Vietnamese context for three key reasons. Firstly, there is a timely need to gain a deeper understanding of the impacts of human-AI collaboration in emerging countries like Vietnam. According to Dwivedi et al. (2021), employees in these nations are more vulnerable to being replaced by AI in repetitive jobs, necessitating further preparation for this significant shift. Secondly, conducting research in less-explored contexts such as Vietnam will help

validate well-developed theories and associated concepts established in more developed regions. Thirdly, Vietnam presents an ideal setting for this study due to the government's proactive efforts to advance the digital economy and society (Dey *et al.*, 2023), wherein human-AI collaboration plays a pivotal role.

Literature review

Task- technology fit theory (TTF)

The TTF theory suggests that technologies enhance performance when they are compatible with the tasks to be performed (Liang *et al.*, 2021). Goodhue and Thompson (1995, p. 216) defined TTF as "*the degree to which a technology assists an individual in performing his or her portfolio of tasks*". Due to the intricate nature of AI adoption in the workplace, which encompasses a variety of techniques, manipulation, and modes of interaction, it is imperative to comprehend the specific roles that AI plays in particular tasks (Jiang, Yang and Zheng, 2023). For example, the ethical considerations for AI implementation may differ across various contexts, and AI capabilities in healthcare may differ considerably from AI applications in entertainment or education (Aker *et al.*, 2023)

Building upon the TTF framework, the research posits that human-AI collaboration nurtures positive psychological states among employees, such as psychological safety and thriving at work. These conducive psychological conditions, in turn, catalyse innovative outcomes manifested through employees' innovative behaviours and the overall innovative performance of the organisation (Cai, Xiao and Ji, 2023; Edmondson and Bransby, 2023).

Human- AI collaboration and firm innovation performance

In today's rapidly evolving markets, innovation is essential for organisations to survive, develop, and achieve exceptional performance (Newman *et al.*, 2020; Cai, Xiao and Ji, 2023).

To thrive for a transformation in human-machine collaboration, both people and organisations must be able to innovate (Yin, Jiang and Niu, 2024; Bankins *et al.*, 2023). Employees, as essential sources of innovation within most organisations, contribute roughly 80% of new ideas that are subsequently implemented (Getz and Robinson, 2003). More specifically, continuous innovation is required and primarily driven by employee innovative behaviour (Verma and Singh, 2022; Gama and Magistretti, 2023). This behaviour enables organisations to explore and apply novel ideas across their operations, which, in turn, increases firms' chance of uncovering non-obvious insights and augmenting their performance, including innovation performance at higher levels continuously (Nguyen and McGuirk, 2022; Zhang and Jin, 2023). Therefore, we posit that human-AI collaboration will foster employees' innovative behaviours, which in turn contribute to firm innovation performance.

H1: Human-AI collaboration is positively related to employees' innovative behaviour.

H2: Employees' innovative behaviour is positively related to firm innovation performance.

The mediating role of psychological safety

Psychological safety refers to employee's psychological state in the workplace where they feel comfortable and empowered to express their thoughts, ideas, and views and engage in risk-taking behaviours without fear of facing negative consequences or social exclusion (Edmondson and

Bransby, 2023). This cognitive tolerance of failure will help decrease employees' anxiety, making them feel more confident in sharing their ideas and engaging in innovative activities despite the risk of failure (Cai, Xiao and Ji, 2023).

Drawing upon the underlying premise of the task-technology fit model, the research posits that the partnership between humans and AI holds significant potential in encouraging employees to complete their tasks. By providing high-value information resources contributing to employees' learning, human-AI collaboration enables employees to comprehend forthcoming job assignments better, mitigating their anxiety in the face of uncertainty and enhancing their perception of psychological safety. In addition, this collaboration can facilitate a set of artificial scenarios that support employees in taking risks and help enhance their self-confidence. Those who perceive this collaboration as a means for the company to enhance overall performance cultivate a cognitive resilience towards failure, leading to reduced anxiety and increased courage in expressing ideas and pursuing innovative endeavours, even in the face of potential failure. Hence, the following hypothesis is suggested:

H3: Psychological safety mediates the relationship between human-AI collaboration and employees' innovative behaviour.

The mediating role of thriving at work

Employees who thrive within the workplace will gauge whether they are navigating toward self-development and progress through vitality and learning (Niessen, Sonnentag and Sach, 2012; Yang *et al.*, 2021). While vitality represents workers' energy and liveliness, learning pertains to their acquisition and application of information and skills, ultimately fostering creativity and guaranteeing success for employee innovative efforts (Jiang *et al.*, 2022). Thriving is influenced by social factors, indicating that the social work environment enables thriving experiences such as leadership, workplace climates, and task resources (Nguyen and McGuirk, 2022; Yang *et al.*, 2021; Zhang *et al.*, 2023). In this case, in light of the TTF theory, the level of match between human and AI collaboration will define the task resources and further employees' thriving at work.

In contrast, human-AI collaboration might enhance employees' self-determination, meet employees' psychological needs, and promote learning and vitality, all of which are integral to thriving at work (Yang *et al.*, 2021). Individuals who are thriving have a perpetual inclination towards acquiring new information and achieving mastery, enabling them to explore, create, and implement novel concepts (Amabile *et al.*, 1996; Kleine, Rudolph and Zacher, 2019). For instance, when AI encounters situations outside its previous training, it may fail to perform correctly, requiring humans to detect and address these deficiencies. In such cases, thriving employees are more likely to be encouraged and motivated to challenge established practices and norms, acquire new knowledge, and generate innovative solutions to retrain AI.

Hence, the following hypothesis is formulated:

H4: Thriving at work mediates the relationship between human-AI collaboration and employees' innovative behaviour.

The research model with the hypothesised relationships is depicted in the following figure.

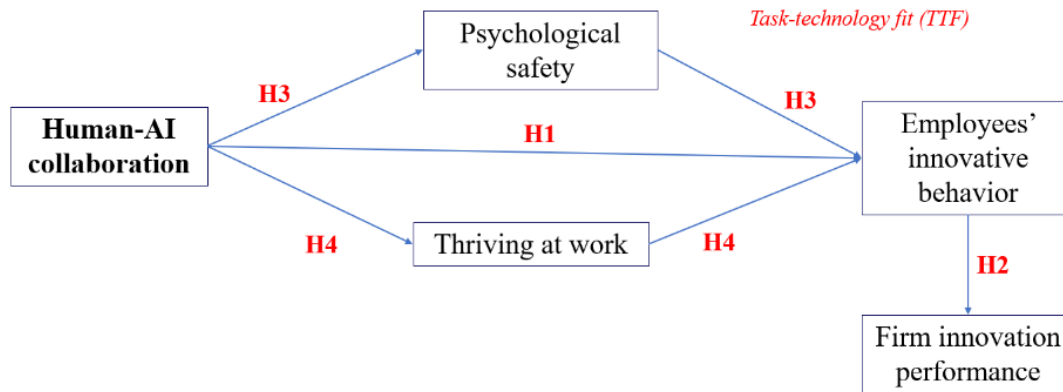


Figure 1. Theoretical framework

Proposed methodology

This research will apply mixed methods using quantitative and qualitative approaches to examine the proposed research model. An online survey and semi-structured interviews will be conducted to collect empirical data for the model. According to Neuman (2011), the synthesis of the two approaches supports and strengthens each other's findings. Applying this mixed approach will help to enrich the research, gather more reliable and valid findings, exhibit imaginative and emotional behaviours, and explain these behaviours through data interpretation (Creswell and Creswell, 2018).

This study will apply the explanatory mixed method design, which involves collecting quantitative data in the initial phase and qualitative data in the second phase. The quantitative research will examine the effects of human-AI collaboration on innovation outcomes and the mediating effects of psychological safety and thriving at work on these relationships. Furthermore, qualitative research can assist in resolving "how" and "why" questions by exploring further the topics under investigation. It helps bring many perspectives (multiple realities) to the situation, allowing for a complete understanding of the challenges.

Quantitative study

Data collection: The authors will apply the multi-source data collection technique in collecting data. Questionnaires will be sent to managers and employees in e-commerce companies in Vietnam. The participants will be asked to complete the survey anonymously to ensure their confidentiality. However, information about positions (employees or managers) and their working departments will be asked to better navigate the question flow and support for analysis later. For managerial roles, they will be asked to evaluate their employees' innovative behaviour (anonymously in the same department).

Data analysis: The research will assess the model and data compatibility by employing confirmatory factor analysis (CFA) with AMOS software. After that, the researcher will use

structural equation modelling to assess the variables' direct relationships. The mediation effects will be tested by conducting a bootstrapping test.

Qualitative study

Data collection: The recruiting procedure included purposive and snowball sampling methods to encourage Vietnam-based e-commerce companies' managers who are willing to engage in semi-structured interviews. These interviewees possess deep knowledge and experience in leading their organisations with the implementation of AI technologies. During interviews (around 30 minutes), these senior managers will be requested to provide their perspectives and first-hand encounters on the collaboration of humans and AI.

Data analysis: Once data is transcribed (automatically with MS Teams online interviews and manually with face-to-face interviews), the NVivo program will be used to assist with thematic coding. The data will be analysed using coding and pattern matching, with underlying themes determined by reference to the study's literature, interviews, and theoretical assertions. The coded data can also be generated into charts, maps, and diagrams in NVivo to help identify the conceptual patterns.

Expected contributions

The study is expected to provide important implications for both researchers and practitioners. Regarding theoretical contributions, this study will advance our understanding of the human-AI collaboration and bridge the identified gaps in its impacts on innovative outcomes and the mediating mechanisms underlying this relationship. In addition, the findings carry significant insights for corporations to cultivate an environment that develops more efficient and fruitful human-AI collaborations. Furthermore, this study will enhance knowledge about the impacts of AI in the specific context of Vietnam- an emerging country that has been investing resources to accelerate a digital economy to gain competitive edges.

References

- Akter, S., Hossain, M. A., Sajib, S., Sultana, S., Rahman, M., Vrontis, D. and McCarthy, G. (2023) 'A framework for AI-powered service innovation capability: Review and agenda for future research', *Technovation*, 125, pp. 102768.
- Amabile, T. M., Conti, R., Coon, H., Lazenby, J. and Herron, M. (1996) 'Assessing the Work Environment for Creativity', *Academy of Management journal*, 39(5), pp. 1154-1184.
- Banks, S., Ocampo, A. C., Marrone, M., Restubog, S. L. D. and Woo, S. E. (2023) 'A multilevel review of artificial intelligence in organisations: Implications for organisational behavior research and practice', *Journal of organisational behavior*.
- Brougham, D. and Haar, J. (2018) 'Smart Technology, Artificial Intelligence, Robotics, and Algorithms (STARA): Employees' perceptions of our future workplace', *Journal of management & organisation*, 24(2), pp. 239-257.
- Cai, L., Xiao, Z. and Ji, X. (2023) 'Impact of supervisor developmental feedback on employee innovative behavior: roles of psychological safety and face orientation', *Journal of managerial psychology*, 38(1), pp. 73-87.
- Constantinides, P., Monteiro, E. and Mathiassen, L. (2024) 'Human-AI joint task performance: Learning from uncertainty in autonomous driving systems', *Information and organisation*, 34(2), pp. 100502.
- Creswell, J. W. and Creswell, J. D. (2018) *Research design : qualitative, quantitative, and mixed methods approaches*. Fifth edition. edn. Thousand Oaks, California: SAGE Publications, Inc.
- Dey, P. K., Chowdhury, S., Abadie, A., Vann Yaroson, E. and Sarkar, S. (2023) 'Artificial intelligence-driven supply chain resilience in Vietnamese manufacturing small- and medium-sized enterprises', *International journal of production research*, ahead-of-print(ahead-of-print), pp. 1-40.
- Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., Medaglia, R., Le Meunier-FitzHugh, K., Le Meunier-FitzHugh, L. C., Misra, S., Mogaji, E., Sharma, S. K., Singh, J. B., Raghavan, V., Raman, R., Rana, N. P., Samothrakis, S., Spencer, J., Tamilmani, K., Tubadji, A., Walton, P. and Williams, M. D. (2021) 'Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy', *International journal of information management*, 57, pp. 101994.
- Edmondson, A. C. and Bransby, D. P. (2023) 'Psychological Safety Comes of Age: Observed Themes in an Established Literature', *Annual review of organisational psychology and organisational behavior*, 10(1), pp. 55-78.
- Gama, F. and Magistretti, S. (2023) 'Artificial intelligence in innovation management: A review of innovation capabilities and a taxonomy of AI applications', *The Journal of product innovation management*.
- Getz, I. and Robinson, A. G. (2003) 'Innovate or Die: Is that a Fact?', *Creativity and innovation management*, 12(3), pp. 130-136.
- Goodhue, D. L. and Thompson, R. L. (1995) 'Task-Technology Fit and Individual Performance', *MIS quarterly*, 19(2), pp. 213-236.
- Issa, H., Jaber, J. and Lakkis, H. (2024) 'Navigating AI unpredictability: Exploring technostress in AI-powered healthcare systems', *Technological forecasting & social change*, 202.

- Jiang, N., Liu, X., Liu, H., Lim, E. T. K., Tan, C.-W. and Gu, J. (2022) 'Beyond AI-powered context-aware services: the role of human-AI collaboration', *Industrial management + data systems*.
- Jiang, Y., Yang, X. and Zheng, T. (2023) 'Make chatbots more adaptive: Dual pathways linking human-like cues and tailored response to trust in interactions with chatbots', *Computers in human behavior*, 138, pp. 107485.
- Kleine, A. K., Rudolph, C. W. and Zacher, H. (2019) 'Thriving at work: A meta-analysis', *Journal of organisational behavior*, 40(9-10), pp. 973-999.
- Liang, T.-P., Kohli, R., Huang, H.-C. and Li, Z.-L. (2021) 'What Drives the Adoption of the Blockchain Technology? A Fit-Viability Perspective', *Journal of management information systems*, 38(2), pp. 314-337.
- Loske, D. and Klumpp, M. (2021) 'Human-AI collaboration in route planning: An empirical efficiency-based analysis in retail logistics', *International journal of production economics*, 241, pp. 108236.
- Neuman, W. L. (2011) *Social research methods : qualitative and quantitative approaches*. 7th edn. Boston: Allyn & Bacon.
- Newman, A., Round, H., Wang, S. and Mount, M. (2020) 'Innovation climate: A systematic review of the literature and agenda for future research', *Journal of occupational and organisational psychology*, 93(1), pp. 73-109.
- Nguyen, N. P. and McGuirk, H. (2022) 'Evaluating the effect of multifactors on employee's innovative behavior in SMEs: mediating effects of thriving at work and organisational commitment', *International journal of contemporary hospitality management*, 34(12), pp. 4458-4479.
- Niessen, C., Sonnentag, S. and Sach, F. (2012) 'Thriving at work-A diary study', *Journal of organisational behavior*, 33(4), pp. 468-487.
- Pereira, V., Hadjielias, E., Christofi, M. and 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), pp. 100857.
- Verma, S. and Singh, V. (2022) 'Impact of artificial intelligence-enabled job characteristics and perceived substitution crisis on innovative work behavior of employees from high-tech firms', *Computers in human behavior*, 131, pp. 107215.
- Yang, Y., Li, Z., Liang, L. and Zhang, X. (2021) 'Why and when paradoxical leader behavior impact employee creativity: Thriving at work and psychological safety', *Current psychology (New Brunswick, N.J.)*, 40(4), pp. 1911-1922.
- Yin, M., Jiang, S. and Niu, X. (2024) 'Can AI really help? The double-edged sword effect of AI assistant on employees' innovation behavior', *Computers in human behavior*, 150, pp. 107987.
- Zhang, R., Kang, H., Jiang, Z. and Niu, X. (2023) 'How does workplace ostracism hurt employee creativity? Thriving at work as a mediator and organisation-based self-esteem as a moderator', *Applied psychology*, 72(1), pp. 211-230.
- Zhang, X. and Jin, H. (2023) 'How does smart technology, artificial intelligence, automation, robotics, and algorithms (STAARA) awareness affect hotel employees' career perceptions? A disruptive innovation theory perspective', *Journal of hospitality marketing & management*, 32(2), pp. 264-283.