

Understanding AI Hallucination Effects on Human-AI Symbiosis Quality based on Extended Cognition Theory

Yayao Gao ¹
Qiong Jia ^{2,*}

*Corresponding author

¹ Hohai University, Nanjing, China, 2208050205@hhu.edu.cn

² Hohai University, Nanjing, China, jiaqionghit@163.com

ABSTRACT

AI hallucination has been a challenging phenomenon in human-AI collaboration. This study examines the relationship between AI hallucination and human-AI symbiosis based on extended cognition theory. We conducted a qualitative study and interviewed 15 college students with varying experiences with AI, analyzing the data using NVivo 12.0. Our research suggests that AI hallucination affects human-AI symbiosis quality by influencing users' AI self-efficacy. AI explainability and critical thinking act as moderating factors. Although AI hallucination may have a negative influence on human-AI symbiosis under some conditions, it can prompt critical thinking in humans. Our findings contribute to a better understanding of AI hallucination based on extended cognition theory and also guide for improving human-AI symbiosis.

Keywords: AI hallucination, human-AI symbiosis quality, extended cognition theory.

INTRODUCTION

Artificial intelligence (AI) technology is developing rapidly, and human-AI interaction is becoming a future working and studying pattern that influences managerial efficiency in today's society. Human-AI symbiosis quality refers to the quality of collaboration between humans and AI systems. High-quality collaboration means both sides support each other in decision-making. The value they create together is higher than the value either could create alone (Jarrahi, 2018). However, AI hallucination creates new challenges for this partnership. Large Language Models (LLMs) sometimes produce fake or misleading content. This occurs due to issues with training data or model limitations (Rawte *et al.*, 2023). AI hallucination poses a threat to the core value of human-AI symbiosis. It challenges the overall quality of collaboration (Almeida & Senapati, 2024).

Many researchers have studied AI hallucination deeply. They explore classification systems, measurement methods, ways to reduce hallucination, and impact analysis. In terms of classification research, several scholars mainly classify AI hallucination based on dimensions such as the degree of knowledge bias and the degree of accuracy matching between the answer and the explanation (Li *et al.*, 2024; Zhou *et al.*, 2025). Based on these theoretical foundations, researchers have developed various measurement methods. These include the reference hallucination score and evaluation of hallucination rates in legal research tools (Aljamaan *et al.*, 2024; Magesh *et al.*, n.d.). They also explore ways to reduce hallucinations, such as input checking (Jancarik & Dušek, 2024). Studies show mixed results about the impact of hallucinations. In medicine, AI hallucinations pose significant risks to health information access (Jin *et al.*, 2025). Some research suggests that hallucinations negatively impact how people perceive AI services (Song *et al.*, n.d.). It can also mislead human reasoning (Spitzer *et al.*, 2025). However, some interesting findings have emerged from the research. One study showed that while hallucination risk hurts AI attitudes, it slightly increases continued use intentions. This suggests that users recognize the risk but still want to use AI (Lim *et al.*, 2025). Further research suggests that AI hallucination may promote human-AI symbiosis. AI-generated hallucination content may contain unique perspectives. These novel combinations can inspire human creators to explore new directions (Haase & Pokutta, 2024).

Current research focuses more on trust (Bach *et al.*, 2024), interaction design (Heyder *et al.*, 2023), and user acceptance (C. Wang *et al.*, 2023). There is a research gap on how AI hallucination specifically affects the effectiveness of human-AI collaboration. Extended cognition theory offers a valuable perspective on understanding this issue. According to extended cognition theory, cognitive processes are not limited to the individual brain. They can extend to external tools and environments. This forms a unified cognitive system (Clark & Chalmers, 1998). Some scholars apply this theory to AI. They believe humans can augment their cognitive capabilities with those of machines under specific conditions. This creates a new form of human-machine cognitive coupling (Smart, 2018). In AI applications, AI systems are integrated with external resources and become a "cognitive extender" of user abilities. The user's thought process is coupled with AI's computational process. They complete complex cognitive tasks together (Hernández-Orallo & Vold, 2019). This cognitive extension relationship means AI hallucination is not just a technical output error. It is cognitive interference in the human-AI cognitive system. This will directly affect the performance of the entire extended cognitive system. Therefore, this study aims to

understand how AI hallucination affects the quality of human-AI symbiosis. Grounded in extended cognition theory, this study employs qualitative research methods. We want to identify key influencing factors and provide a theoretical basis and practical guidance for optimizing human-AI interaction design.

RESEARCH METHOD AND MATERIALS

Research Method

This study is based on the extended cognition theory and adopts grounded theory methodology, which helps us explore the theoretical model of the relationship between AI hallucination and the quality of human-AI symbiosis. We want to understand the multi-dimensional influencing mechanism of AI hallucination. On this basis, we will develop a systematic theoretical framework for explanation. Grounded theory is a systematic qualitative research method. An inductive analytical path characterizes it. It generates theories from empirical data rather than verifying established theoretical assumptions (Wolfswinkel *et al.*, 2013). Unlike traditional deductive research methods, grounded theory emphasizes that researchers should maintain open theoretical sensitivity. This allows core concepts and theoretical insights to emerge naturally during data analysis. Hence, it ensures the solidity and relevance of theory construction (Strauss & Corbin, 1998). However, AI hallucination is an emerging phenomenon that warrants exploration using grounded theory. Grounded theory can systematically construct theoretical models from empirical data through its inductive and iterative analytical features. It can reveal the mechanism of how AI hallucination influences the quality of human-AI symbiosis. In this study, all interview data have been imported into the software for data processing. We used open coding, axial coding, and selective coding methods.

Research Data Collection

This study examines how AI hallucination impacts the quality of human-AI symbiosis. We selected interview subjects based on several considerations: (1) Age matches AI usage experience. Interview subjects are aged 20-22 years old. This group is currently in a critical period of rapid development in AI technology. They have in-depth experience using AI tools. They can also clearly express the AI hallucination phenomena they encountered during human-AI interaction. We selected interviewees with varying durations of AI usage. These range from beginners with less than 1 year of experience to senior users with more than 3 years of experience. (2) Professional backgrounds are representative. Interviewees mainly came from human resource management majors. Nevertheless, they also covered various professional fields. These include financial management, marketing, international economics and trade, information management, control science, French, and forestry. This ensures that our study results accurately reflect the diverse user experiences across various professional backgrounds. (3) Gender coverage is comprehensive. We included interviewees of different genders in the sample. This helps capture possible gender differences in how people perceive and process AI hallucination. (4) Use of AI tools is diverse. Interviewees used various AI tools. These include DeepSeek, Doubao, ChatGPT, ERNIE Bot, Claude, Gemini, and Kimi. This helps us understand how hallucination affects the quality of human-AI symbiosis across different AI systems. The details of the interviewees are shown in Table 1.

The interviews were semi-structured. We conducted half-hour in-depth interviews with each participant. The formal interview outline had three sets of questions. The first set of questions was used to determine whether the interviewees' experiences fit our research theme. We guided suitable AI users to describe their experiences with AI hallucinations. The second set of questions is used to determine the decision-making behavior of the interviewees. We guided AI users to recall how they dealt with tasks when AI hallucination occurred. The third set of questions was used to assess the effectiveness of the interviewees' tasks. We investigated whether the effectiveness of their task completion changed after the emergence of AI hallucination. To ensure the reliability and validity of the research, this study adopted the principles of typicality and representativeness. We recruited college students online through WeChat's social platform. We interviewed 15 users with diverse genders, majors, and experiences with AI. During interviews, we promptly followed up on unclear expressions. We made appropriate adjustments to the outline after interviews. Eventually, we obtained 15 valid interview data sets, including two pre-interviews. Saturation was confirmed after 15 interviews, as further data collection only replicated existing patterns without adding conceptual depth. This resulted in 83,860 words of interview transcripts.

Table 1: Information of interviewees

	Basic Information	Frequency	Percentage
Gender	Male	4	27
	Female	11	73
Age	20 years old	3	20
	21 years	8	53
	22 years	4	27
Major	Financial management	1	7
	French	1	7
	International economics and trade	1	7
	Control science	1	7

	Forestry	1	7
	Human resources management	8	53
	Marketing	1	7
	Information management and information systems	1	7
Length of AI use	Less than 1 year	3	20
	1~2 years	7	47
	2~3 years	2	13
	3 years and above	3	20
Frequently used AI tools	Doubao	9	20
	ERNIE Bot	5	11
	ChatGPT	7	16
	Claude	2	4
	DeepSeek	15	33
	Gemini	1	2
	Kimi	6	13

Source: This study.

Data Analysis and Model Construction

Open coding

Open coding is the first step in grounded theory. It aims to break down large amounts of raw data into smaller, manageable units. From these, we can extract meaningful concepts and categories. This stage of coding is open. It allows researchers to extract and explore possibilities from the data freely. According to the interview text, interviewees first described specific experiences with AI hallucination, such as fabricating references. These descriptions helped us identify basic concepts like AI hallucination types. As the interview progressed, we found that the respondent developed complex cognitive responses after encountering AI hallucination, including changes in AI trust and adjustments in AI dependency. These responses pointed to changes in the users' internal cognitive state. Subsequently, the respondent adopted various coping strategies, such as adjusting question phrasing and verifying information sources. These behaviors reflected the user's active efforts to adapt to AI hallucination. We used NVivo 12.0 to analyze the interview data word by word. This resulted in 158 initial concepts and 26 categories, as shown in Table 2.

Table 2: Analysis of interview data of parents (open coding)

Sentence excerpts from raw data	Initial concepts	Category
I was writing my thesis. I asked the AI to generate information about current research. The AI created fake names and fake references. The research section became entirely fictional. These people and references do not exist.	A16 citation fabrication	B1 type of AI hallucination
AI systems should not create completely wrong answers. They should inform users when content may be fabricated. This helps users identify fictional content. Users can then avoid being deceived.	A39 demand for suggestive explanations	B8 differences in user needs for AI explanations
Resume information matters to me. Therefore, I would verify every AI answer to ensure that important information is accurate.	A50 important task to scrutinize AI Answers	B10 checking AI answer conditions
I use different AI models. I would like to give each model a clear framework for creating content. More detailed instructions help the AI give better answers. When the AI makes mistakes, this also pushes me to improve how I use it.	A81 Adjusting the way of asking questions to AI	B14 the promotion of critical thinking by hallucination
I focus on where the AI gets its data. I think the AI's reasoning process works well.	A98 refined AI logical reasoning	B16 AI trustworthiness
I still trust AI a lot. It has helped me solve complex problems in my field. The AI makes mistakes sometimes. I try to help it fix these mistakes with my feedback.	A123 positive feedback on AI use experiences	B20 confidence in using AI
I experienced AI creating fictional content. My work with AI might be less effective right now. However, this will get better as I learn more about AI. My past failures taught me what AI cannot do. Now I can let AI focus on what it does	A158 overall effectiveness decreases before increasing	B26 change in quality of human-AI symbiosis

best.		
-------	--	--

Source: This study.

Axial coding

Axial coding is a further development of open coding. Its main aim is to identify and establish relationships between categories. This forms a more systematic and structured framework for analysis. We have noticed that AI hallucination affects the AI-human collaboration by influencing users' cognitive states. When users encounter AI hallucination, their trust evaluation of AI systems and their perception of their own usage capabilities are primarily affected. This cognitive change is then reflected in actual AI-human collaboration performance. Regarding the relationship between categories, the same AI hallucination can produce significantly different effects under varying levels of explanation transparency and user cognitive abilities. This prompts us to identify AI explainability and users' AI critical thinking as important moderating variables. Additionally, multiple interviewees mentioned that they accumulate experience in handling AI hallucination during the process. This will in turn affect the quality of subsequent collaboration. Therefore, the AI-human collaboration process can influence users' cognitive abilities and coping strategies in reverse, forming a dynamic feedback loop. The analysis of interview data produced five main categories and ten subcategories. The details are as shown in Table 3.

Table 3: Analysis of parents' interview data

Open code extraction category	Subcategory	Main category
B1 reasons for AI hallucination	C1 AI hallucination generation	D1 AI hallucination
B2 AI hallucination type		
B3 hallucination recognition ability	C2 hallucination recognition	
B4 reason for being able to recognize hallucination		
B5 hallucination inspiration		
B6 insufficient AI explanation	C3 AI explanation effectiveness	D2 AI explainability
B7 AI explanation usefulness		
B8 AI explains differences in user needs	C4 AI explanatory differences	
B9 AI model differences		
B10 checking AI answer conditions	C5 AI answer quality assessment	D3 AI evaluation and reflection
B11 dealing with AI hallucination behaviors		
B12 hallucination discrimination concerns		
B13 hallucination identifies limitations		
B14 hallucination's promotion of critical thinking	C6 experience formation and accumulation	
B15 historical experience formation		
B16 AI trustworthiness	C7 AI trust evaluation	D4 AI and self-perception
B17 AI dependence		
B18 confidence in the AI itself		
B19 sense of control over AI	C8 perception of the user's ability	
B20 confidence in using AI		
B21 role relationship with AI		
B22 dissatisfaction with AI	C9 current status of AI use	D5 current status of human-AI symbiosis
B23 AI satisfaction		
B24 purpose of AI use		
B25 frequency of AI use		
B26 change in collaboration effectiveness	C10 human-AI collaboration efficacy	

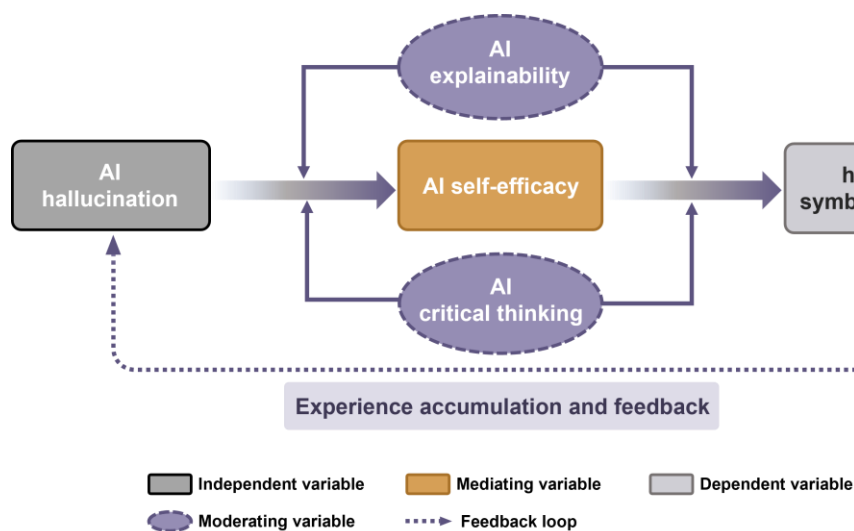
Source: This study.

Selective coding

Selective coding allows for further systematic processing of associations between categories. It provides a more logical description of human-AI symbiosis. It also helps construct models to represent the relationship between AI hallucination and the quality of human-AI symbiosis.

For the initial analysis, we had considered only AI hallucination as a core category. However, in-depth analysis revealed significant variability in users' responses to AI hallucination. A simple stimulus and response model cannot explain this variability. This prompted us to revisit the data. Ultimately, we determined that AI self-efficacy serves as a mediating variable connecting AI hallucination to the quality of human-AI symbiosis. After establishing the core category, we found it necessary to further clarify the mechanisms among the variables. Through the analysis of interview data, we identified that AI self-efficacy encompasses two key dimensions: trust in AI systems and self-perceived ability to use them. These two dimensions work together to form the cognitive foundation users employ to address AI hallucination. Finally, we integrated static relationships and dynamic processes. The key is to recognize that the quality of human-AI symbiosis is not a static outcome variable, but a dynamic process that is constantly evolving through continuous interaction. Users accumulate experience and enhance their cognitive abilities by handling AI hallucination, which in turn improves subsequent collaboration, creating a virtuous cycle.

After iterative theory building, we concluded that AI hallucination affects the quality of human-AI symbiosis in three main ways: AI explainability, AI self-efficacy, and critical thinking. AI explainability refers to the process of demonstrating how AI systems operate. It gives people details and reasons to help them understand what AI is doing (Barredo Arrieta *et al.*, 2020). AI self-efficacy is about confidence in AI. It shows how much people believe they can use AI tools and how well they think they can work with AI (Y.-Y. Wang & Chuang, 2024). AI critical thinking refers to the evaluation of AI-generated results. People look at how well AI can handle information and think through problems (Oliveira *et al.*, 2025). AI hallucination mainly manifests as content fabrication, data fabrication, and information errors. These hallucination phenomena impact the quality of human-AI symbiosis by influencing users' AI self-efficacy. AI self-efficacy plays a key role as a mediating variable. It reflects both users' evaluation of AI's trustworthiness and their perception of their ability to use AI. When users can recognize and rationally view AI hallucination, they do not simply reduce their reliance on AI. Instead, they optimize human-AI collaboration by adjusting usage strategies and enhancing critical thinking. AI explainability and critical thinking act as important moderating variables. They affect the quality of human-AI collaboration in terms of technical transparency and user cognitive ability. Good AI explainability helps users understand the reasoning process and limitations of AI. Critical thinking ability enables users to evaluate the quality of AI-generated content reasonably. The model also shows a dynamic feedback loop. Users accumulate experience in human-AI symbiosis practice. This experience enhances their ability to recognize and cope with AI hallucinations. This forms a positive development mechanism. This cyclical process ultimately promotes continuous improvement of human-AI symbiosis quality. This results in increased collaboration efficiency, improved task completion quality, and enhanced user satisfaction. The details are as shown in Figure 1.



Source: This study.

Figure 1: Theoretical model diagram of AI hallucination and human-AI symbiosis quality

Theoretical saturation test

We kept raw data from three in-depth interviews for the theoretical saturation test. According to the three-level coding results, we found no new conceptual categories and relational patterns in the core categories of AI hallucination, AI explainability, AI assessment and reflection, AI self-awareness, and the current state of human-AI symbiosis. Therefore, we can assume that the theoretical model has reached theoretical saturation.

RESULTS

Users Focus on Data Accuracy and Task Completion When Discussing AI Hallucination

We analyzed word frequency statistics from interview texts. The study found that the impact of AI hallucination on the quality of human-AI symbiosis showed a multi-level topic structure. The word "data" has the highest frequency in interviews. This is

making. This is especially true when AI uses specialized terminology to package misinformation. The user's negative feelings intensify even further. The moderating effects of AI explainability prompt a reevaluation of evaluation criteria for the quality of human-AI symbiosis. Interview data indicate that users are increasingly emphasizing the transparency and understandability of the interaction process. They do not focus only on efficiency or accuracy of task completion. Being able to understand AI's reasoning process is seen as more valuable than obtaining exactly the correct result. Understanding enables users to actively engage in error identification and correction rather than passively accepting AI output. This cognitive shift suggests that the human-AI symbiosis relationship is evolving from simple tool use to a more profound level of cognitive collaboration. AI explainability becomes an essential foundational element in building high-quality human-AI symbiosis relationships. By providing effective explanatory mechanisms, AI systems can transform potential hallucination risks into opportunities to promote user understanding and learning. It will enhance the overall quality of human-AI symbiosis.

Critical Thinking Skills Determine How Effectively Users Respond to AI Hallucination

Users' critical thinking level directly affects the effect AI hallucination has on the quality of human-AI symbiosis. The same AI hallucination phenomenon will trigger wildly different results in user groups with different critical thinking levels. When users possessed high AI critical thinking skills, AI hallucination catalyzed to stimulate their cognitive potential and learning motivation. Interview data revealed that these users initially had high levels of trust in AI. However, after discovering that AI could make mistakes, they began to think more independently. They realized that different AI models had their characteristics. This cognitive shift led to the development of various coping strategies. These include searching for validated information sources independently, refining the way they ask questions to make them more specific, and switching flexibly between different AI models. In contrast, users with relatively weak critical thinking skills showed obvious cognitive difficulties when encountering AI hallucination. They often struggled to recognize the authenticity of AI-generated content. They were easily confused by superficial logic. Several users admitted that they are unable to identify AI errors. They thought AI-generated content was logical to some extent. Nevertheless, at the same time, they also recognized that they might be unable to detect errors. This cognitive limitation makes AI hallucination have a greater negative impact on the effectiveness of human-AI collaboration. The moderating effect of critical thinking is also reflected in users' attitudes and subsequent behavioral choices regarding the limitations of AI technology. High AI critical thinking users tend to understand AI hallucination as an inherent feature of current technological development. They proactively explore how to optimize human-AI collaboration models based on recognizing these limitations. Low AI critical thinking users are more likely to be frustrated by AI hallucination. They may even develop overall distrust of AI technology. Therefore, cultivating and improving users' critical thinking skills can help them better understand AI technology. This realizes higher-level human-AI cognitive collaboration and value creation.

Three Different Patterns for Users to Identify AI Hallucination

The study found that users' ability to recognize and respond to AI hallucination shows significant individual differences. This variability not only affects users' trust in AI systems but also profoundly influences the quality and nature of human-AI symbiosis. Through systematic analysis of interview data, we identified three typical cognitive patterns. The first type is characterized by low recognition and high reliance. Users with this cognitive pattern often lack the professional knowledge to evaluate the authenticity of AI output. One interviewee confessed that he had never encountered AI fabrication. AI-generated content had some logic, but it is also possible that the AI could not determine whether the content it generated was accurate. This cognitive blindness makes it easy for users to fall into ignorant trust. The potential risk is high. The second type is the high recognition rational response pattern. These users typically have strong professional backgrounds and critical thinking skills in AI. They can keenly detect AI errors and take effective response strategies. They will not only search for validation on their own but also replace AI models or refine questions. It shows a mature human-AI collaboration ability. The third type is the moderate recognition context-dependent pattern. The recognition ability of these users varies depending on the task type and area of specialization. They show noticeable context sensitivity. Such cognitive differences are not fixed individual characteristics. They dynamically evolve during continuous human-AI interaction. Users who can transform AI hallucination recognition into learning opportunities tend to develop higher-quality human-AI symbiosis relationships. They no longer regard AI as a simple tool but as a cognitive partner. By calibrating each other, they realize mutual growth.

CONCLUSION AND DISCUSSION

Research Conclusions

Based on grounded theory analysis, this study offers a systematic understanding of how AI hallucination affects the quality of human-AI symbiosis. We conducted in-depth interviews with 15 college students about their experiences with AI use. Through data analysis, we found that AI hallucination is not merely a technical defect. It is a complex phenomenon that affects the quality of human-AI symbiosis. Multiple dimensions and levels characterize its action mechanism. There are significant individual differences in users' cognition and coping ability with AI hallucination. Such differences directly impact the effectiveness of human-machine collaboration. In this process, AI hallucination affects human-AI symbiosis quality by influencing users' AI self-efficacy. AI critical thinking ability and AI explainability play key roles as important moderating variables. Notably, users have gradually developed experience and capabilities in identifying and processing AI hallucinations. It forms a dynamic feedback mechanism that promotes the improvement of human-AI symbiosis quality. These findings offer new theoretical insights and practical guidance for understanding and optimizing human-AI symbiotic relationships. They reveal the positive role that AI hallucination may play in human-AI symbiosis.

Theoretical Findings

This study constructs a theoretical model of the impact of AI hallucination on the quality of human-AI symbiosis. This enriches the application of extended cognition theory in the field of artificial intelligence. The model reveals that AI hallucination, as cognitive interference, affects the entire human-AI cognitive system by impacting the user's AI self-efficacy. This provides a new explanatory dimension for extended cognition theory. Meanwhile, the moderating effect of AI explainability and AI critical thinking enriches extended cognition theory regarding the impact of instrumental transparency and cognitive ability. The three cognitive modes of users' responses to AI hallucination suggest that users' professional backgrounds, usage experiences, and AI critical thinking abilities jointly shape the operation of extended cognition systems.

The theoretical model constructed in this study breaks through the limitations of existing theoretical frameworks in several dimensions. In terms of self-efficacy theory, traditional theories emphasize the formation of self-efficacy through four fixed sources of information: achievement experience, alternative experience, verbal persuasion, and physiological state (Bandura, 1977). In contrast, this study found that AI self-efficacy presents dynamic change characteristics. Users actually become better at controlling AI systems when confronted with AI illusions. In terms of explainability theory, previous XAI theory focuses on technical transparency and error avoidance (Barredo Arrieta *et al.*, 2020). However, this study found that the main value of AI explainability is that it transforms AI errors into opportunities for users to learn. This demonstrates a new way for humans and AI to think and learn together. The above analysis suggests that under certain conditions, AI hallucinations can promote users' cognitive development, which challenges the traditional view of extended cognition theory. The traditional view regards external tool deficiencies as purely negative factors. Our study provides new theoretical insights into understanding dynamic adaptation and optimization mechanisms in extended cognitive systems. It pushes extended cognition theory in a more complex and dynamic direction.

Practical Enlightenment

In terms of technical design, AI system developers should emphasize the importance of explainability functions. When users understand AI's reasoning process and decision-making basis, they can more easily recognize potential instances of AI hallucination. They can also build stronger trust relationships based on this understanding. Developers should actively explain to users possible cases of AI hallucination. They should provide corresponding identification methods and coping strategies. It will establish reasonable user expectations. In terms of user training, educational institutions and enterprises should incorporate AI literacy training into their regular educational systems. The key is to cultivate critical thinking skills and help users form rational usage attitudes. While utilizing AI to enhance efficiency, users should maintain their independent judgment. In terms of collaboration mechanisms, appropriate AI output validation processes should be established in work scenarios, particularly in tasks that involve significant decisions. This validation mechanism needs to find a balance between ensuring work efficiency and providing adequate validation. It should neither be too complicated to affect work progress nor too simplified to lose its gatekeeper role.

Research Limitations and Future Research Directions

The model construction is based on a relatively limited sample. Interview subjects are mainly concentrated in the 20-22 age group of college students. Most come from backgrounds in human resource management. This sample composition may affect the generalizability of research findings. We need to verify the model's applicability across user groups of different age groups, occupational fields, and educational backgrounds. While qualitative research methods can dig deeper into the phenomenon's nature, there are inherent limitations in establishing strict causal relationships. Experimental designs may provide more substantial evidence for verifying the causal influence mechanism of AI hallucination. Studies have mainly focused on the user perception level. The analysis of how different AI system technical features affect hallucination generation and user experience is relatively weak. The cross-sectional study design limits the ability to capture the dynamic, evolutionary process of human-AI symbiosis relationships. This is especially true for how users gradually adjust their coping strategies and cognitive modes for long-term interactions. Due to time and resource limitations in research, these issues require further exploration in subsequent studies.

REFERENCES

- Aljamaan, F., Temsah, M.-H., Altamimi, I., Al-Eyadhy, A., Jamal, A., Alhasan, K., Mesallam, T. A., Farahat, M., & Malki, K. H. (2024). Reference hallucination score for medical artificial intelligence chatbots: Development and usability study. *JMIR Medical Informatics*, *12*, e54345. <https://doi.org/10.2196/54345>
- Almeida, F., & Senapati, B. (2024). Striving for symbiosis: Human-machine relations in the AI era. *2024 IEEE Integrated STEM Education Conference (ISEC)*, 1–4. <https://doi.org/10.1109/ISEC61299.2024.10664823>
- Bach, T. A., Khan, A., Hallock, H., Beltrão, G., & Sousa, S. (2024). A systematic literature review of user trust in AI-enabled systems: An HCI perspective. *International Journal of Human-Computer Interaction*, *40*(5), 1251–1266. <https://doi.org/10.1080/10447318.2022.2138826>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible. *Information Fusion*, *58*, 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>

- Clark, A., & Chalmers, D. J. (1998). The extended mind. *Analysis*, 58(1), 7–19.
- Haase, J., & Pokutta, S. (2024). *Human-AI Co-Creativity: Exploring Synergies Across Levels of Creative Collaboration*. *arXiv preprint arXiv:2411.12527*. <https://arxiv.org/abs/2411.12527v2>
- Hernández-Orallo, J., & Vold, K. (2019). AI extenders: The ethical and societal implications of humans cognitively extended by AI. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 507–513. <https://doi.org/10.1145/3306618.3314238>
- Heyder, T., Passlack, N., & Posegga, O. (2023). Ethical management of human-AI interaction: Theory development review. *The Journal of Strategic Information Systems*, 32(3), 101772. <https://doi.org/10.1016/j.jsis.2023.101772>
- Jancarik, A., & Dušek, O. (2024). The problem of ai hallucination and how to solve it. *European Conference on E-Learning*, 23(1), 122–128. <https://doi.org/10.34190/ecel.23.1.2584>
- Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, 61(4), 577–586. <https://doi.org/10.1016/j.bushor.2018.03.007>
- Jin, L., Shen, Z., Alhur, A. A., & Naeem, S. B. (2025). Exploring the determinants and effects of artificial intelligence (AI) hallucination exposure on generative AI adoption in healthcare. *Information Development*. <https://journals.sagepub.com/doi/abs/10.1177/02666669251340954>
- Li, J., Chen, J., Ren, R., Cheng, X., Zhao, W. X., Nie, J.-Y., & Wen, J.-R. (2024). *The Dawn After the Dark: An Empirical Study on Factuality Hallucination in Large Language Models* (arXiv:2401.03205). *arXiv*. <https://doi.org/10.48550/arXiv.2401.03205>
- Lim, J. S., Shin, D., Lee, C., Kim, J., & Zhang, J. (2025). The role of user empowerment, AI hallucination, and privacy concerns in continued use and premium subscription intentions: An extended technology acceptance model for generative AI. *Journal of Broadcasting & Electronic Media*. <https://www.tandfonline.com/doi/abs/10.1080/08838151.2025.2487679>
- Magesh, V., Surani, F., Dahl, M., Suzgun, M., Manning, C. D., & Ho, D. E. (2025). Hallucination-free? Assessing the reliability of leading AI legal research tools. *Journal of Empirical Legal Studies*, 22(2), 216-242. <https://doi.org/10.1111/jels.12413>
- Oliveira, L., Tavares, C., Strzelecki, A., & Silva, M. (2025). Prompting minds: Evaluating how students perceive generative AI's critical thinking dispositions. *Electronic Journal of E-Learning*, 23(2), Article 2. <https://doi.org/10.34190/ejel.23.2.3986>
- Rawte, V., Sheth, A., & Das, A. (2023). A survey of hallucination in large foundation models (arXiv:2309.05922). *arXiv*. <https://doi.org/10.48550/arXiv.2309.05922>
- Smart, P. R. (2018). Human-extended machine cognition. *Cognitive Systems Research*, 49, 9–23. <https://doi.org/10.1016/j.cogsys.2017.11.001>
- Song, Y., Cui, Mingjia, Wan, Fei, Yu, Zhengwang, & Jiang, J. (n.d.). AI hallucination in crisis self-rescue scenarios: The impact on AI service evaluation and the mitigating effect of human expert advice. *International Journal of Human-Computer Interaction*, 0(0), 1–21. <https://doi.org/10.1080/10447318.2025.2483858>
- Spitzer, P., Holstein, J., Morrison, K., Holstein, K., Satzger, G., & Kühn, N. (2025). Don't be fooled: The misinformation effect of explanations in Human-AI collaboration (arXiv:2409.12809). *arXiv*. <https://doi.org/10.48550/arXiv.2409.12809>
- Strauss, A., & Corbin, J. M. (1998). *Basics of Qualitative Research Techniques*. Sage Publications.
- Wang, C., Ahmad, S. F., Bani Ahmad Ayassrah, A. Y. A., Awwad, E. M., Irshad, M., Ali, Y. A., Al-Razgan, M., Khan, Y., & Han, H. (2023). An empirical evaluation of technology acceptance model for Artificial Intelligence in E-commerce. *Heliyon*, 9(8), e18349. <https://doi.org/10.1016/j.heliyon.2023.e18349>
- Wang, Y.-Y., & Chuang, Y.-W. (2024). Artificial intelligence self-efficacy: Scale development and validation. *Education and Information Technologies*, 29(4), 4785–4808. <https://doi.org/10.1007/s10639-023-12015-w>
- Wolfswinkel, J. F., Furtmueller, E., & Wilderom, C. P. M. (2013). Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems*, 22(1), 45–55. <https://doi.org/10.1057/ejis.2011.51>
- Zhou, J., Zhang, J., Wan, R., Cui, X., Liu, Q., Guo, H., Shi, X., Fu, B., Meng, J., Yue, B., Zhang, Y., & Zhang, Z. (2025). Integrating AI into clinical education: Evaluating general practice trainees' proficiency in distinguishing AI-generated hallucinations and impacting factors. *BMC Medical Education*, 25(1), 406. <https://doi.org/10.1186/s12909-025-06916-2>