

# ***Impact of AI on Human Decision-Making: Analysis of Human, AI, and Environment of Interaction***

**Liwen, Zhang<sup>1,a,\*</sup>**

<sup>1</sup>*Zhou Enlai School of Government, Nankai University, Tongyan Street, Tianjin, 300350, China*  
*a. xuanjue0906@163.com*

*\*corresponding author*

**Abstract:** In recent years, the application of AI in the decision-making field is increasing, and AI is frequently and deeply involved in every aspect of decision-making. As AI-assisted decision-making is becoming an important intersection of computer science and psychology, there is an emerging need for understanding how would AI affect people's decision-making. Due to the rapid development of AI technology, a gap of literature reviews considering this field has arisen. In response, this article systematically reviews previous related researches from the classic perspectives of human-computer interaction: human, AI, and the environment of interaction, and each section contains several sub aspects. In detail, explainability and presentation of AI's decision, preference and demographic variables of human, and three specific circumstances of interaction are included. In addition, some principled issues such as research ethics have also been discussed. In sum, this article organizes structured empirical evidence for ongoing related researches, summarizes the shortcomings of previous researches, and proposes possible directions for more in-depth researches.

**Keywords:** AI, AI-assisted, human-computer interaction, decision-making

## **1. Introduction**

In recent years, the application of AI in the decision-making field is increasing, and AI is frequently and deeply involved in every aspect of decision-making, ranging from movie recommendations on Netflix and Amazon and tailored advertisements on Google search result pages [1] to human resource decisions and clinical decision making [2]. More recently, the rise of new applications of AI like AI painting and AI chatting (ChatGPT for example) is aggressively attempting even to replace people's thinking and decision-making process in brain [3].

As AI being a highly valued research direction in computer science, and decision-making also being a long-standing research object in psychology, it comes to the intersection of the two important subjects now and it is forming a brand new branch of science. According to Arrieta, the sophistication of AI-powered systems has lately increased to such an extent that almost no human intervention is required for their design and deployment [4]. Starting with simple mechanical movements like controlling smart furniture, AI is more and more often seen in people's daily lives. When AI ultimately affects people's decision-making which should be made extremely careful (e.g. medical issue, moral dilemma, financial risk), it is crucial to figure out the specific impact of AI on human decision-making.

However, as a newly emerged research field, although rapidly developing, the number of existing researches in it is currently not sufficient enough, suggesting a large number of blank areas have not been studied yet. At the same time, there are also various problems of existing researches. For example, the use of professional terminology is not uniform (e.g., researchers searching terminologies of XAI may be confused with understandability, comprehensibility, interpretability, explainability and transparency) [4]. Besides, the description of certain independent and dependent variables is unclear and inaccurate (e.g., the dependent variables used to describe the performance of people's decision-making could vary from accurate rate of a specific task to perceptual trust towards AI). In addition, there are no standard norms for research paradigms and experimental scenarios (e.g., paradigms like behavioral experiments and self-statement reports are frequently used but without a standard procedure).

In order to propose solutions to the problems mentioned above and provide directions for future researches towards the blank areas, this article systematically reviews past researches from the classic perspectives of human-computer interaction: human, AI, and the environment of interaction. This article also organizes structured empirical evidence for ongoing related researches, summarizes the shortcomings of previous researches, and proposes possible directions for more in-depth researches. The main body of this article consists of 5 sections. Section 1 introduces the research background and writing purpose. Section 2 analyzes how explainability and presentation of AI affect human decision-making. Section 3 reviews how preference of people and demographic characteristics affect human decision-making. Section 4 discusses how specific situations of AI-assisted task affect human decision-making. Finally, section 5 and section 6 make a discussion and conclusion to this article.

## **2. AI Factors Affecting AI-assisted Decision-making**

### **2.1. XAI**

XAI, standing for eXplainable(X) Artificial(A) Intelligence(I), is “one that produces details or reasons to make its functioning clear or easy to understand” [4]. Leichtmann revealed that explanations of the AI's predictions would lead to a statistically significant better performance of decision-making through a mushroom-picking task [5]. Hudon also found that transparent and explainable AI systems could result in improved confidence between human and AI, which further have a positive impact on decision-making [6]. From an opposite perspective, Ebermann proved that under situations with cognitive misfit, users would experience negative mood significantly more often and thus have a negative evaluation of AI's support [7]. In conclusion, XAI does have benefits for building up the bridge of trust between human and AI. As AI has experienced a rapid developing period of time in recent years, people are attaching more and more importance on AI's explainability. However, there are opposite ideas that the key of successful decision making is more likely depending on accuracy not explainability of AI.

Though there seems to be no contradiction between explainability and accuracy of AI, when it comes to the situation that explainability must trade off against accuracy, people will choose to prioritize accuracy over explainability in practice [8]. XAI often has a positive impact on people's sense of trust and fairness especially when considering profits to further influence decision-making. In other words, explainability sometimes equals to fairness because people are willing to see how a sufficient AI system operates. However, Langenberg emphasized that assuming the relationship between fairness and accuracy to be a trade-off, an increase in fairness would definitely lead to an unavoidable loss of accuracy [9]. It is of no doubt that explanations of an AI system are improving people's trust and dependence on AI, but only when AI provides correct answers do people make more intellectually correct decisions. According to Klichowski's research, most (more than 85%) participants would choose to agree with AI even when they were clearly aware of that the AI's choice

had no sense [10]. It is obvious that AI's assistance could lead to a blind answer of "yes" instead of rational decision-making. Above all, XAI is just like a coin, with people's trust on one side and AI's lie on the other side, strangely balanced. It is of urge to make sure that XAI won't provide fake information before it is put into use in human decision-making.

## **2.2. Presentation of AI's Decision**

AI's decision can not only be presented by text, but also by graphs, voices, robots, even Virtual Reality (VR). Hudon and Karran separately examined two presentation-order methods' and three AI decision visualization attribution models' impact on decision-making [6][11]. They found that a visual decision-making task could lower participants' cognitive load, increase participants' confidence and cognitive fit, and thus improve participants' performance on decision-making with AI's assistance. But this does not mean that more lively a situation is, more correct a decision is.

The presentation is not only influencing on cognitive mediators, but also provoking emotional arousal of decision makers. Patil compared moral dilemmas presented in text format and in VR, revealing that participants would act in a more utilitarian manner and emotionally arousing way in VR [12]. Niforatos further expanded the experiment and concluded that the virtual enactment of a moral dilemma can further foster utilitarianism, even when decisions are on average biased towards utilitarianism already [13]. The distance created by VR between human and AI will hinder people's empathy when facing each other in specific and universal moral dilemmas [14]. Despite moral dilemmas, an AI-enabled chatbot's anthropomorphization would activate great psychological risk attachment, which enacts people to manifest stronger risk aversion tendency [15]. It remains under explored that thoroughly how do different presentations of AI influence human decision-making, but an existing milestone is that methods like robots and VR can evoke more emotions and have a significant impact on decision-making when AI is assisting human.

## **3. Human Factors Affecting AI-assisted Decision-making**

### **3.1. Preference**

People prefer AI to human experts would probably choose to follow AI's decision while people with the opposite preference would possibly reject AI's recommendation. An interesting idea is that people's familiarity with AI shapes their preference. The word "familiarity" here includes two specific meanings: people's familiarity with decision-making tasks, and people's familiarity with AI. As for the former one, "familiarity" indicates whether one is an expert of the task's domain or not. Kramer discovered that the more familiar participants were with AI making decisions in specific scenarios, the more likely they were to prefer AI over human decision-makers [16]. However, Berger raised objection by claiming that humans do not generally prefer human decision-makers to AI but will reject AI's recommendation after becoming familiar with the task. A possible reason for this disagreement is that each experiment task under certain and unique circumstances could lead to significant individual differences [17]. In Kramer's experiment, participants need to decide whether the AI agents they are to consider have an option for human override or not. In Berger's experiment, participants are to predict the number of incoming calls of a client's hotline operation. A task that is more specific and close to real life may lead to more confidence for participants themselves, which may further result in rejection to AI. Fortunately, it is still credible that people's familiarity with specific domains is influencing on human decision-making with AI's assistance.

As for the latter one, "familiarity" indicates the experience of exposure to AI. Through a robot-assisted surgery investigation, Joan found that previous exposure to AI does not have a generally significant influence on people's trust on AI [18]. But Joan also noted that there are significant internal differences between demographic variables including gender, age and educational level.

Specifically, the effect of exposure on trust is greater among men, people between 40 and 54 years old, and those with higher educational levels. Another notable thing is that Joan's data were all collected on only Europeans, suggesting the regional difference remains a mystery.

Generally speaking, familiarity has a slight and vague impact on human decision-making, implying the possible existence of other potentially more explanatory influencing factors.

### **3.2. Demographic Variables**

Joan's work examined people's performance in an AI-assisted decision-making task, and further examined the result between different genders, ages, and educational levels [18]. Joan's work suggested that basically there were no significant difference between male and female participants with 51.8% of the sample being male and 45.1% being female. Participants in different ages may focus on different attributes of AI but their performance is not significantly different. When it comes to educational levels, it is interesting that the result only showed a significant positive effect for the group of 16 to 19 years of education. Unfortunately, similar studies on demographic variables have been conducted relatively little, thus comparative validation cannot be formed as a result.

## **4. Environment of Interaction Affecting AI-assisted Decision-making**

### **4.1. Moral Dilemma**

When faced with issues related to human ethics and morality, people often believe that AI can solve moral dilemmas more reasonably, thus relying more on AI in moral decision-making. When AI and people working together to solve a moral problem, although people may believe that human is more morally trustworthy, they would consider their problem-solving abilities to be far inferior to AI. That is to say, AI's decisions are more easily accepted than human decisions, and further more, people may believe that AI and its controllers have less responsibility to bear when the problem is not solved [2]. Similarly, when AI and humans make the same unethical decisions, people do not tend to attribute errors to AI, and even AI operators are less blamed than independent individuals [19]. But this tendency seems to be limited to people with a clear understanding and understanding of the moral issues they are currently facing. For example, when an AI system on an automatic car runs into a moral dilemma of tram problem, it is uncertain whether people are willing to trust AI's decision or not. Such moral dilemmas are abstract for human beings, and there are also some abstract and distant moral distances between human and AI as well.

This so-called moral distance is created by technology and will hinder people's empathy without the condition of facing each other in specific and universal moral dilemma [15]. When making moral decisions, due to AI's participation, people will not see the vulnerability of human nature and the characteristics, personality and situation of specific individuals, This actually makes people more likely to make choice to reject AI [20]. Overall, when making moral decisions, people not only consider the consequences of their actions, but also the situation in which they are executed. AI is difficult to receive and process this information, which may indicate that people are more willing to trust their own moral judgments in moral decisions. A solution to this may be to use game theory and machine learning methods to make AI learn situations that are difficult to describe in natural language. After all, the aggregation of multiple human moral viewpoints may produce a better moral system than any individual's morality. AI can also recognize the general principles of moral decision-making that humans have not been aware of before, so as to make more perfect decisions [21].

## 4.2. Medical Issue

Under medical circumstances, AI's diagnostic can hardly positively influence patients' decision making [22]. On issues related to life safety, explanations matter most as no one would like to hand over their own lives to something unpredictable and unstable. Triberti said that "Clinical decisions could even be delayed or paralyzed when AI's recommendations are difficult to understand or to explain to patients" [23]. Despite explainability of AI, who makes the recommendations and what the recommendations are like also matters. Formosa found evidence of a "human bias" (a preference for human over AI decision makers) and an "outcome bias" (a preference for positive over negative outcomes) [24]. It fits human intuitions that people instinctively pursue to live and would prefer more warmth than bias.

Another interesting phenomenon is that doctors and patients hold different attitudes towards AI-assisted medical choice. Doctors are willing to rely on AI even regardless of accuracy or explanations [25] while patients tend to be very rational and make the decision exclusively on their own wishes [18]. It is for sure that AI has a deeper knowledge reserve and more rational thinking ability than doctors, which may increase dependence on AI. Meanwhile, it can also be reasonable that doctors are not willing to take responsibility for the possibility of patients' exacerbation while patients often share no same value of life with AI and they tend to hold the hope of life on their own hands.

## 4.3. Financial Risk

People's preference towards AI bias with different levels of risk. Wang found that when the stakes of the decisions become larger, people tend to lower their belief in AI recommendation's correctness and rely more on their own judgement in AI-assisted decision making through a task of evaluating loan default risks [26]. This has been confirmed in an early experiment conducted by Tversky and Kahneman, in which people were more irrational and would tend to avoid economic risks more when facing larger financial risks especially larger stakes [27]. Dikmen found that participants who had access to domain knowledge relied less on AI assistant when AI assistant was incorrect and indicated less trust in AI assistant [28]. The words "domain knowledge" is similar to "explainability" but clearly not exactly the same. It is more like a source of confidence for decision-makers. The increase of confidence leads to lower cognition and estimation of level of risk, thus lower the possibility of taking AI's advice. As people's attitude towards AI in decision-making equals to their attitude towards financial risk, there will be corresponding results in specific situations. In addition, a related comprehensive research showed that AI's recommendation though improved task performance (as AI being absolutely rational), only limited impact on risk-taking behavior was observed (which is more dependent on the circumstances). In the meantime, AI was under-valued by participants (as people are sensitive to economic loss) [29].

## 5. Discussion

This article systematically reviews previous researches on field of AI and human decision-making and tries to reach a conclusion on how would AI affect human decision-making. As a newly emerged field, the connection between AI and human decision-making is changing rapidly due to the even higher speed of AI tech development. Such a relatively short period of time contains a large amount of information gap, bias in point of view, and some potential crisis. The main problems in this field are as follows:

There is simply too little researches on demographic variables. As a newly emerged field, such basic research is actually essential, as its effects may mask the main effects of many empirical studies. At the same time, there has been significant breakthroughs in this field in the past two to three years.



Before a large amount of academic researches were conducted, there would be countless people who had sufficiently exposed to AI. That is why it is necessary to conduct relevant researches.

The black box issue of AI also needs attention. Just like machine learning, if people cannot fully understand the operating principles and mechanisms of AI, it will remain impossible to achieve complete utilization of AI, and the scenes in science fiction where humans are in turn dominated by AI may eventually come true.

The ethical issues of AI need to be taken seriously too. After all, AI is not an exact human being, and AI under current technology cannot perfectly simulate human emotions. Therefore, when it comes to ethical issues, AI can hardly man to provide answers that human will fully accept. Notably, this may also be a good new, as AI may never be able to completely replace human decision-making.

## 6. Conclusion

With the integration of computer science and psychology, it is crucial to figure out the specific impact of AI on human decision-making.

This article reaches the conclusion that the impact of AI on human decision-making is reflected in multiple aspects. From the perspective of AI, the interpretability and presentation of AI have a significant impact on human decision-making. From a human perspective, demographic variables and familiarity with AI do not seem to have a significant impact on human decision-making. From the perspective of various problem situations, specific problem situations often correspond to people's different reactions and tendencies, as human are always complex and difficult to perfectly predict.

Based on the viewpoints of this article, future researches could focus more on AI's impact on human decision-making under specific circumstances. It is a huge and difficult project but general patterns only come from accumulation of a large amount of empirical evidences. The in-depth exploration of interpretability is also an important part of future researches. Only when people can thoroughly analyze and understand AI can we trust, control, and utilize AI for the benefit of humanity without any risk or burden.

## References

- [1] Adadi, A., & Berrada, M. (2018). *Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)*. *IEEEAccess*, 6, 52138–52160.
- [2] Tolmeijer, S., Christen, M., Kandul, S., Kneer, M., & Bernstein, A. (2022). *Capable but Amoral? Comparing AI and Human Expert Collaboration in Ethical Decision Making*. *CHI'22, New Orleans, LA, USA*.
- [3] Thorp, H. H. (2023). *ChatGPT is fun, but not an author*. *Science*, 379(6630), 313–313.
- [4] Arrieta, A. B., Díaz-Rodríguez, N., Ser, J. D., Benetot, 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 AI*. *Information Fusion*, 58, 82-115.
- [5] Leichtmann, B., Hinterreiter, A., Humer, C., Streit, M., & Mara, M. (2023). *Explainable Artificial Intelligence Improves Human Decision-Making: Results from a Mushroom Picking Experiment at a Public Art Festival*. *International Journal of Human-Computer Interaction*.
- [6] Hudon, A., Demazure, T., Karran, A., Léger, P., & Sénécal, S. (2021). *Explainable Artificial Intelligence (XAI): How the Visualization of AI Predictions Affects User Cognitive Load and Confidence*. *NeuroIS*, 52, 237-246.
- [7] Ebermann, C., Selisky, M., & Weibelzahl, S. (2022). *Explainable AI: The Effect of Contradictory Decisions and Explanations on Users's Acceptance of AI Systems*. *International Journal of Human-Computer Interaction*, 39(9), 1807-1826.
- [8] Nussberger, A., Luo, L., Celis, L. E., & Crockett, M. J. (2022). *Public attitudes value interpretability but prioritize accuracy in Artificial Intelligence*. *Nature Communications*, 13, 5821.
- [9] Langenberg, A., Ma, S., Ermakova, T., & Fabian, B. (2023). *Formal Group Fairness and Accuracy in Automated Decision Making*. *Mathematics*, 11, 1771.
- [10] Klichowski, M. (2020). *People Copy the Actions of Artificial Intelligence*. *Frontiers in Psychology*, 11, 1130.
- [11] Karran, A., Demazure, T., Hudon, A., Senecal, S., & Léger, P. (2022). *Designing for Confidence: The Impact of Visualizing Artificial Intelligence Decisions*. *Frontiers in Neuroscience*, 16, 883385.

- [12] Patil, I., Cogoni, C., Zangrando, N., Chittaro, L. & Silani, G. (2014). Affective basis of judgment-behavior discrepancy in virtual experiences of moral dilemmas. *Soc Neurosci*, 9(1), 94-107.
- [13] Niforatos, E., Palma, A., Gluszny, R., Vourvopoulos, A., & Liarokapis, F. (2020). Would you do it?: Enacting Moral Dilemmas in Virtual Reality for Understanding Ethical Decision-Making. *CHI'20, Honolulu, HI, USA*.
- [14] Montemayor, C., Halpern, J., & Fairweather, A. (2021). In principle obstacles for empathic AI: why we can't replace human empathy in healthcare. *AI & SOCIETY*, 37, 1353-1359.
- [15] Cui, Y. (2022). Sophia Sophia tell me more, which is the most risk-free plan of all? AI anthropomorphism and risk aversion in financial decision-making. *International Journal of Bank Marketing*, 40(6), 1133-1158.
- [16] Kramer, M. F., Borg, J. S., Conitzer, V., & Sinnott-Armstrong, W. (2018). When Do People Want AI to Make Decisions? *AIES'18, New Orleans, LA, USA*.
- [17] Berger, B., Adam, M., Ruhr, A., & Benlian, A. (2021). Watch Me Improve—Algorithm Aversion and Demonstrating the Ability to Learn. *Bus Inf Syst Eng*, 63(1), 55-68.
- [18] Joan, T., Ana, I. J., & Francesc, S. (2021). Do People Trust in Robot-Assisted Surgery? Evidence from Europe. *International Journal of Environmental Research and Public Health*, 18, 12519.
- [19] Shank, D. B., DeSanti, A., & Maninger, T. (2019). When are artificial intelligence versus human agents faulted for wrongdoing? Moral attributions after individual and joint decisions. *Information, Communication & Society*, 22, 648-663.
- [20] Carolina, V., & Martin, K. (2023). Moral distance, AI, and the ethics of care. *AI & SOCIETY*, 1-12.
- [21] Conitzer, V., Sinnott-Armstrong, W., Borg, J. S., Deng, Y., & Kramer, M. (2017). Moral Decision Making Frameworks for Artificial Intelligence. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, 31(1).
- [22] McDougall, R. J. (2018). Computer knows best? The need for value-flexibility in medical AI. *Med Ethics*, 0, 1-5.
- [23] Triberti, S., Durosini, I., & Pravettoni, G. (2020). A "Third Wheel" Effect in Health Decision Making Involving Artificial Entities: A Psychological Perspective. *Frontiers in Public Health*, 8, 117.
- [24] Formosa, P., Rogers, W., Griep, Y., Banks, S., & Richards, D. (2022). Medical AI and human dignity: Contrasting perceptions of human and artificially intelligent (AI) decision making in diagnostic and medical resource allocation contexts. *Computers in Human Behavior*, 133(1).
- [25] Meyer, J., Khademi, A., Tetu, B., Han, W., Nippak, P., & Remisch, D. (2022). Impact of artificial intelligence on pathologists' decisions: an experiment. *J Am Med Inform Assoc*, 29(10), 1688-1695.
- [26] Wang, X., Lu, Z., & Yin, M. (2022). Will You Accept the AI Recommendation? Predicting Human Behavior in AI-Assisted Decision Making. *WWW '22, Virtual Event, Lyon, France*.
- [27] Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica: Journal of the Econometric Society*, 47, 263-291.
- [28] Dikmen, M., & Burns, C. (2022). The effects of domain knowledge on trust in explainable AI and task performance: A case of peer-to-peer lending. *Int. J. Human-Computer Studies*, 162(2).
- [29] Elder, H., Canfield, C., Shank, D. B., Rieger, T., & Hines, C. (2022). Knowing When to Pass: The Effect of AI Reliability in Risky Decision Contexts. *Human Factors*.