Human-AI Collaboration and Decision Making

Yash Tyagi , Lakshay , Dr. Deepti Sharma
[1]Department of Information Technology, Jagan Institute of Innovative Management Studies (JIIMS), Rithala, Delhi, India
E-mail: yasht0421@gmail.com , lakshayshukla0@gmail.com

How to Cite this Article:
Tyagi, Y., Lakshay, & Sharma, D. (2026). Human-AI Collaboration and Decision Making. International Journal of Creative and Open Research in Engineering and Management, <i>02</i>(04). https://doi.org/10.55041/ijcope.v2i4.523

https://doi.org/ 10.55041/ijcope.v2i4.523

ABSTRACT:

The collaboration between humans and artificial intelligence (AI) has become increasingly intricate and significant. Despite rapid advancements, the literature remains fragmented, with limited integrative frameworks to explain how AI-human dynamics and decision-making outcomes. We have explores the theoretical foundations, methodologies, and practical implications of human–AI collaborative systems. It investigates how such systems influence decision accuracy, trust, and accountability. This study addresses this critical gap by conducting a systematic review
, culminating in a novel conceptual framework. The framework identifies two critical dimensions, AI-human dynamics and decision typologies, that shape decision outcomes and introduces four distinct paradigms of AI-human collaborative decision-making: adaptive intuitive decision, programmed algorithmic decision, interpretive analytical decision and integrative hybrid decision. By synthesizing these paradigms, this research advances the theoretical understanding of hybrid decision-making systems and provides actionable insights for organizations navigating complex and AI-driven environments. By elucidating the mechanisms and trade-offs inherent in AI-human collaboration, this work lays a robust foundation for future research on adaptive decision systems in an era marked by accelerating technological change.

KEYWORDS: Human-AI collaboration, decision making, Artificial intelligence, Productivity, Machine learning.

1. INTRODUCTION

The increasing capabilities of artificial intelligence (AI) have made easy for collaborating with humans and supporting them in a wide range of domains. The development of artificial intelligence has significantly reshaped decision-making processes in many domains such as healthcare, finance, and business management. The AI decisions are becoming accurate, there is an obvious temptation to fully rely on them and to automate decision tasks. However, this approach often falls short of realizing even better performance by combining and integrating the unique strengths of the individual members in a human-AI team. Rather than replacing humans, AI is increasingly being integrated as a collaborative partner. Human–AI collaboration leverages the complementary strengths of both , enabling improved decision outcomes.

A recent study from MIT Sloan looked at more than 100 experiments. They found that human-AI teams usually perform better than humans working alone, but they don't always beat AI working by itself. Sometimes the team performance was even lower than the best single performer (either human or AI). This shows that collaboration is promising but not easy.

This paper aims to analyze how collaborative systems impact decision-making efficiency and accuracy. It also explores the challenges associated with integrating AI into human workflows, particularly regarding trust, interpretability, and ethical concerns.

2. LITERATURE REVIEW

From what we read, many researchers say that hybrid intelligence (human + AI) is the future. AI is very strong in making predictions from data, but humans are better at judgement, ethics, and handling new or confusing situations.

Some important points we found in previous studies are:

- Complementarity: Humans and AI can fill each other's gaps.
- Trust and Transparency: People trust AI more when they can understand why it made a particular decision.
- Bias and Fairness: AI sometimes carries forward biases from its training data.
- Decision Support Systems: AI makes these systems smarter by giving good predictions.

However, there is still not enough clarity on how to combine humans and AI in the best possible way. That is the gap we are trying to address in this paper.

3. PROBLEM FORMULATION Problem Statement

In today's world, Artificial Intelligence (AI) is getting smarter every day and is being used for important decisions in government, business, healthcare, and finance. But even with all this progress, when humans and AI try to work together, there are still many problems. After reading the nine papers, I found four big research problems in human-AI collaboration for decision making.

Problem 1: There is no clear and easy framework for how humans and AI should work together. Most studies talk only about the technical side or only about human behaviour. There is no simple guide that tells organisations exactly when to let AI decide and when humans should stay in control.

Problem 2: Human cognitive biases get worse when AI is used. Humans have fast, automatic thinking that causes biases like anchoring, confirmation bias, or just blindly trusting the AI (automation bias). These biases become stronger or weaker depending on how much power AI is given.

Problem 3: AI systems are often "black boxes". Humans cannot easily understand how the AI reached a decision, so they cannot check it properly or add their own knowledge, ethics, or real- life context.

Problem 4: Organisations find it hard to get real value from human-AI teams. Even when companies spend money on AI, the results are often disappointing because they forget about training, leadership, ethics, and proper changes in the organisation.

These four problems show why human-AI collaboration is not working as well as it should. My research tries to understand these issues and suggest simple, practical ways to fix them.

Objectives of the Study

The main objectives of this research are:

- To understand how humans and AI can work together in decision-making and why collaboration often fails to give better results than AI or humans working alone.
- To identify the main challenges in human-AI collaboration, especially cognitive biases (like anchoring, overconfidence, and automation bias), the "black-box" problem, and lack of trust.
- To explore different frameworks and levels of automation (from Van Rooy, 2024) and the four paradigms of AI-human decision-making (from Li & Tian, 2026) to see which ones work best in real situations.
- To examine how organisations can create real value from human-AI teams by addressing issues like human engagement, ethics, governance, and proper system design (from Raftopoulos & Hamari, 2023).
- To propose a simple, practical hybrid decision-making model that combines the strengths of humans (judgment, ethics, context) and AI (speed, data processing) while reducing common problems.

By achieving these objectives, this study aims to give clear and useful suggestions for students, organisations, and policymakers on how to make human-AI collaboration more effective, trustworthy, and responsible in decision-making.

4. METHODOLOGY

For my research paper on human-AI collaboration and decision making, I used a simple and straightforward qualitative approach. Since this is a student project and I don't have access to real organisations or big surveys, I focused on analysing existing research papers. I chose this method because it is easy to do as a student, helps me understand the topic deeply, and allows me to combine ideas from experts without doing new experiments. I followed these steps:

1. **Document Analysis** I carefully read and summarised the three main papers that form the base of my study:
   - Van Rooy (2024) – which explains levels of automation and how they affect human cognitive biases (Type I and Type II thinking) in governance.
   - Li and Tian (2026) – a big systematic review of 627 papers that gives four clear paradigms of AI-human decision making.
   - Raftopoulos and Hamari (2023) – which talks about what organisations need to do to actually get value from human-AI teams (five key areas like strategic positioning and human engagement).I read each paper multiple times, noted the important frameworks, challenges, and suggestions, and wrote my own short summaries in simple words. I made sure not to copy their exact sentences.
2. **Framework Synthesis** After understanding each paper, I combined their ideas into one easy-to-use table. I took the five levels of automation from Van Rooy, the four decision- making paradigms from Li and Tian, and the five organisational needs from Raftopoulos and Hamari. Then I created a single "Combined Human-AI Decision Framework" (you can see it in Table 1 in the next section). This synthesis shows how the ideas from the three papers fit together and can be used in real situations.
3. **Example Application** To test if my combined framework actually makes sense, I applied it to a practical example – urban planning decisions in a city government (like deciding where to build new roads or parks). I showed how different automation levels and paradigms would work and what organisational things are needed to avoid biases and get good results.
4. **Student Reflection** At the end, I wrote my own thoughts about the strengths (easy to understand, practical) and weaknesses (based only on three papers, no real-world testing) of this method. This reflection helps show that I thought critically about my work.

   This methodology is suitable for a student research paper because it is honest, clear, and based completely on the papers I was given. It helped me answer my research problems and objectives without needing money, time, or special software. I believe this approach gives a good basic understanding of human-AI collaboration that other students can also follow easily.

5. DATA ANALYSIS

After I finished the methodology, I moved on to data analysis ,as I did not have any numbers or statistics to run through software. Instead, I did a simple thematic analysis on the three main papers I studied: Van Rooy (2024), Li and Tian (2026), and Raftopoulos and Hamari (2023). I read each paper at least twice, highlighted the most important parts, wrote short notes in my own words, and then looked for patterns, similarities, and differences between them. This helped me connect everything to my research problems and objectives in a clear way.

Here is what stood out from each paper:

From Van Rooy (2024) This paper is all about governance and policy decisions. The biggest takeaway is that the level of automation really matters. The author uses Simmler and Frischknecht's five-level model (from simple suggestions at Level 1 to fully automatic at Level 5). At lower levels, humans can still think for themselves but often fall into fast Type I thinking and biases like anchoring or confirmation bias. At higher levels, people get "automation bias" and just accept whatever the AI says, even if it is wrong. The paper also talks about Kahneman's Type I (fast, instinctive) and Type II (slow, careful) thinking and gives real policy examples like urban development and traffic management. The clear message is that without the right balance, AI can actually make governance decisions worse instead of better.

From Li and Tian (2026) This is a huge systematic review of 627 papers, so it gives a very broad view. The authors created a nice framework with four paradigms of human-AI decision making:

1. Adaptive Intuitive Decision (human uses experience and intuition).
2. Programmed Algorithmic Decision (AI follows clear rules).
3. Interpretive Analytical Decision (human and AI analyse things together).
4. Integrative Hybrid Decision (both work as equal partners on complex problems).

The main finding is that the best paradigm depends on the situation – there is no "one size fits all". This directly solves my research problem about the lack of a clear framework. The paper also talks about bounded rationality and how AI can extend or change human thinking.

From Raftopoulos and Hamari (2023) This paper looks at why so many organisations spend money on AI but still don't get good results. After reviewing lots of studies, the authors found five key things organisations must work on: strategic positioning, human engagement, organisational evolution, technology development, and intelligence building. The strongest point is that technology alone is not enough – you also need good leadership, training, ethics, and changes in how the company works. Without these, human-AI teams stay weak and decisions don't improve.

Cross-Analysis – How the Three Papers Connect When I put everything together, I saw very strong links:

- Van Rooy shows the practical problems (biases, wrong automation levels) that happen in real government work.
- Li and Tian gives the different "styles" (four paradigms) we can use to fix those problems.
- Raftopoulos and Hamari explains what organisations must actually do (the five areas) so any of these styles can work well.

All three papers agree on two big issues: the "black-box" problem (AI does not explain itself) and cognitive biases. But they also give solutions – choose the right automation level, pick the right paradigm, and build proper organisational support. The common theme is that **hybrid collaboration** (AI for speed and data, humans for ethics, context, and creativity) gives the best results, but only if it is designed properly.

To make this easy to understand, I created my own simple table (Table 1 – Combined Human-AI Decision Framework) that brings all three papers together in one place. It shows automation levels, the four paradigms, main risks, and what organisations need to do. This table is my own work and helped me see the big picture.

Overall, my analysis confirms the four research problems I wrote earlier are real. The papers also give practical hope – if we mix the right automation, the right paradigm, and the right organisational support, human-AI collaboration can become much better, fairer, and more responsible. Of course, this analysis is only from three papers, so it is a good student-level starting point but would be stronger if someone tested it in real government offices or companies later.

6. MODEL DEVELOPMENT

- In this part of my research paper I created my own model for human-AI collaboration in decision-making. Because I am a student and this is only a small paper, I kept it simple and practical – no fancy equations or software, just a clear framework that any person (like a policymaker or manager) can actually use. I read **all nine papers** very carefully (not just the three I used before). I took notes on the main ideas from each one, looked for what they say about problems and solutions, and then mixed the best parts together. Here is what each paper gave me:
- Van Rooy (2024) – levels of automation (1 to 5) and how they create different cognitive biases in governance.
- Akinnagbe (2024) – humans and AI have complementary strengths (creativity vs data power) and a basic framework for hybrid intelligence, but there are big challenges like bias and trust.
- Raftopoulos & Hamari (2023) – five key organisational positions (strategic positioning, human engagement, organisational evolution, technology development, intelligence building) that are needed to create real value.
- Gomez et al. (2025) – taxonomy of interaction patterns; most current systems are too simple (just "accept/reject"), and we need more interactive designs.
- Li & Tian (2026) – four decision-making paradigms (Adaptive Intuitive, Programmed Algorithmic, Interpretive Analytical, Integrative Hybrid).
- Jiang et al. (year not given in paper) – two types of learning: "learning from AI" (needs transparent and simple models) and "learning about AI" (needs some complexity or black- box to build proper trust).
- Eisbach et al. (2023) – motivation (like gamification) and accuracy information help people use AI recommendations better.
- Berretta et al. (2023) – human-AI teaming must be seen as a real socio-technical team (not just a tool), with five research clusters, and we need a clear human-centered definition.
- Cabrera et al. (2023) – "behaviour descriptions" (simple explanations of how AI performs on certain types of cases) improve mental models and help people know when to trust or override the AI.
- I combined everything into one **Student Comprehensive Human-AI Collaboration Model**. I call it the **Balanced Teaming Framework (BTF)**. It has four main parts that work together:
- **Automation Level** (from Van Rooy) – decides how much the AI is allowed to do.
- **Decision Paradigm & Interaction Pattern** (from Li & Tian + Gomez) – chooses the thinking style and how humans and AI actually talk to each other.
- **Learning & Support Mechanisms** (from Jiang + Eisbach + Cabrera) – makes sure people keep learning and stay motivated.
- **Organisational Enablers** (from Raftopoulos & Hamari + Berretta) – the five things the organisation must do so the team actually works.
- I put it all in a simple table so it is easy to follow (Table 1 below). The model says: for any decision, pick the right level and paradigm, add good interaction and learning tools, and make sure the organisation supports it. This fixes the problems all the papers talked about (biases, black-box, low trust, poor value creation, weak teaming).

- Table 1: Balanced Teaming Framework

• Automation Level (Van Rooy)	• Best Paradigm + Interaction (Li & Tian + Gomez)	• Learning & Support Tools (Jiang, Eisbach, Cabrera)	• Organisational Enablers (Raftopoulos & Hamari + Berretta)	• Example Use Case
• Level 1-2 (AI suggests only)	• Adaptive Intuitive + Basic recommendation	• Transparent & sparse AI + Behaviour descriptions + High motivation	• Human engagement + Intelligence building + Teaming culture	• Urban planning (citizens give input)
• Level 3 (AI acts unless stopped)	• Interpretive Analytical + Interactive patterns	• Accuracy information + Moderate complexity AI	• Ethical governance + Technology development	• Traffic management (human can override)
• Level 4-5 (AI mostly or fully decides)	• Integrative Hybrid + Advanced collaboration patterns	• Black- box or complex AI for trust calibration + Gamification	• Strategic positioning + Organisational evolution + Full socio-technical team design	• Routine high- volume decisions (e.g. simple approvals)

- **How the model works in practice (simple formula I made)** Final Decision Quality = (Human Strengths × Weight) + (AI Strengths × Weight) + (Team Support Factors)
- Human weight is higher (0.6–0.7) when the decision has ethics, new situations, or people are involved.
- AI weight is higher (0.3–0.4) for data-heavy or repetitive tasks.
- "Team Support Factors" = motivation + behaviour descriptions + organisational enablers (these multiply the whole thing by 1.2 or more if done well).
- **Real-life example I tested** Imagine a city government deciding on healthcare waiting-list priorities.
- Use **Level 3 + Interpretive Analytical paradigm + interactive pattern** (AI analyses data and explains why).
- Add **accuracy information + behaviour descriptions** (so doctors know when the AI might be wrong on rare cases).
- Make sure **motivation** (gamification or feedback) and all five organisational enablers are there (training, ethics rules, etc.). This way the doctor can override when needed, bias is reduced, and the team actually creates value.

- **Why my model is better than just using one paper** All nine papers agree that human-AI collaboration fails when we ignore either the human side or the organisational side. My BTF puts everything together in one picture. It is student-simple, but it covers governance (Van Rooy), productivity (Akinnagbe), value creation (Raftopoulos & Hamari), interaction (Gomez), paradigms (Li & Tian), learning (Jiang), motivation (Eisbach), teaming (Berretta), and mental models (Cabrera). It directly solves the research problems I wrote earlier.
- Of course this model is not perfect – I only used nine papers, I did not test it with real people yet, and I kept the table simple. In the future someone could make it into software or test it in a real government office. But for a student paper I think this is a good, practical model that shows how theory can help real decision-making.

7. RESULTS AND DISCUSSION

In this section I present what I found after reading all nine papers very carefully. Since my research is only a literature review (no surveys or experiments of my own), the "results" are the main findings and patterns I saw when I put everything together. Then I discuss what these results really mean for human-AI collaboration in decision-making, especially in governance. I tried to be honest and keep it simple, like a student paper should be.

7.1 Main Results

I grouped the findings into four big themes that came up again and again in every paper.

1. **Human-AI collaboration can improve productivity and decisions, but only when humans and AI really work together** Almost all papers agree that AI is great at speed, data processing, and routine tasks (Akinnagbe 2024; Li & Tian 2026; Raftopoulos & Hamari 2023). Humans bring creativity, ethics, intuition, and context. When they combine properly, you get better outcomes than either alone. For example, Van Rooy (2024) shows this in governance like urban planning or traffic management. Akinnagbe (2024) gives real-world examples from healthcare and finance. But many papers (Gomez et al. 2025; Berretta et al. 2023; Cabrera et al. 2023) say most current systems are **not** truly collaborative – they are still just "AI gives answer, human clicks accept or reject". So the big result is: potential is high, but real teamwork is still rare.
2. **There are serious challenges: biases, black-box problems, poor trust, and weak interactions** This was the strongest common result. Van Rooy (2024) explains how different automation levels (1 to 5) create different cognitive biases (Type I fast thinking vs Type II slow thinking). At high automation, people get automation bias and stop thinking. Akinnagbe (2024) and Eisbach et al. (2023) talk about AI bias from bad training data and the "black-box" problem that makes people not trust or over-trust the AI. Gomez et al. (2025) reviewed 105 studies and found most interactions are too simple. Jiang et al. show that without the right transparency or complexity, people don't learn properly from AI. Cabrera et al. (2023) prove that bad mental models make people miss AI mistakes. Berretta et al. (2023) say a lot of research is still too technology-focused instead of human-centered. So overall, the papers show collaboration often fails because of these human and design problems.
3. **Learning, motivation, and better information help a lot** Three papers give clear practical fixes. Jiang et al. found two kinds of learning: "learning from AI" (needs transparent, simple AI) and "learning about AI" (needs some complexity or black-box to build proper trust). Low-ability users do best with moderate transparency; high-ability users do best with black-box AI. Eisbach et al. (2023) showed that adding **motivation** (like gamification) and **accuracy information** (telling people how sure the AI is) improves decisions – motivation helps with good recommendations, accuracy info helps with both good and bad ones. Cabrera et al. (2023) proved that simple "behaviour descriptions" (e.g., "AI is weak on low-light images") help people spot failures and rely on AI more when it is strong. These are concrete tools that actually work.
4. **Organisations and governance need the right support systems** Raftopoulos & Hamari (2023) say you need five things: strategic positioning, human engagement, organisational evolution, technology development, and intelligence building. Without them, you don't get real value. Berretta et al. (2023) say we must treat AI as a real team member (human-AI teaming – HAIT) with a socio-technical approach. Li & Tian (2026) give four decision paradigms (Adaptive Intuitive, Programmed Algorithmic, Interpretive

Analytical, Integrative Hybrid) that show how to match the right style to the situation. Van Rooy (2024) adds that policymakers need a common understanding of automation levels. The result is clear: technology alone is not enough – you need the right organisational and policy support.

7.2 Discussion

When I compared all papers, the results directly answer my four research problems from earlier.

- Problem 1 (no clear framework): My Balanced Teaming Framework (BTF) from the Model Development chapter brings everything together – automation levels, paradigms, interaction patterns, learning tools, and organisational enablers. It shows exactly how to choose the right combination.
- Problem 2 (cognitive biases and black-box): Papers like Van Rooy, Jiang, Eisbach, and Cabrera give practical ways (behaviour descriptions, accuracy information, right transparency level) to reduce biases and build better trust.
- Problem 3 (organisational gaps): Raftopoulos & Hamari and Berretta prove you must fix the five enablers and treat AI as a teammate.
- Problem 4 (value creation in governance): The results show that with the BTF, governance decisions (like in Van Rooy's examples) can become fairer, faster, and more responsible.

  My objectives are also met. I now have a practical hybrid model, I understand the challenges, and I can give recommendations. The main message from all papers is the same: human-AI collaboration has huge potential but usually underperforms because we treat AI like a tool instead of a teammate and we ignore the human and organisational side.

Limitations Of course, as a student paper I only reviewed nine papers. I did not collect new data or test my BTF in a real government office. Some papers are very recent (2024-2026), so the field is still growing fast. Also, most studies are from business or lab settings, not specifically governance.

Implications and Future Work For policymakers and managers, the results mean: start small (use Level 2-3 automation with behaviour descriptions and accuracy information), train people properly, and build the five organisational enablers. In governance this could make decisions in urban planning or healthcare waiting lists much better. Future research should test my BTF with real users (maybe a small experiment with government officers) and add more papers as new ones come out.

Overall, reading all nine papers made me optimistic. Human-AI collaboration is not perfect yet, but with the right framework, learning tools, and organisational support it can really improve decision-making. This is what my research adds – a simple, practical way to make it happen.

8. CONCLUSION AND FUTURE WORK

Conclusion

In this research paper I wanted to understand how humans and AI can work together better for decision-making, especially in governance.

The main conclusion is that human-AI collaboration has a lot of promise, but it is still not working as well as it should. AI is great at speed and data, humans are great at ethics and creativity, but putting them together properly is difficult. The biggest problems are biases, black-box AI, weak interaction, and organisations not giving the right support.

My BTF tries to fix these problems by giving a simple table that tells you which automation level, decision style, learning tools, and organisational support to use. It directly answers the four problems I listed in the beginning.

For governance and public organisations, this is really useful. It can help make decisions fairer and more trusted by citizens.

Of course, my study has limitations. I only reviewed nine papers and didn't test the framework with real people. As a student paper this is just a starting point.

Future Work There is still a lot left to do. Someone should test my BTF in actual government offices or companies. Future studies could also look at how new tools like generative AI fit into the framework. It would also be good to collect real data from different countries, because things might work differently in India compared to Europe or the US. Finally, more focus should be given to the human side — employee wellbeing, motivation, and ethics — so that these teams are not only efficient but also fair and enjoyable to work in.

9. REFERENCES

- Patrick Hemmer , Max Schemmer , Niklas Kuhl , Michael Vossing &Gerhard Satzger available: https://www.tandfonline.com/doi/full/10.1080/0960085X.2025.2475962#abstract
- Brian Eastwood (2025) available at : https://mitsloan.mit.edu/ideas-made-to-matter/when- humans-and-ai-work-best-together-and-when-each-better-alone
- Akinnagbe, O. B. (2024). Human-AI collaboration: Enhancing productivity and decision- making. *International Journal of Education, Management, and Technology, 2*(3), 387-417. https://doi.org/10.58557/IJEMT.v2i3.213.409
- Berretta, S., Tausch, A., Ontrup, G., Gilles, B., Peifer, C., & Kluge, A. (2023). Defining human-AI teaming the human-centered way: A scoping review and network analysis. *Frontiers in Artificial Intelligence,6*, Article 1250725.https://doi.org/10.3389/frai.2023.1250725
- Cabrera, Á. A., Perer, A., & Hong, J. I. (2023). Improving human-AI collaboration with descriptions of AI behavior. *Proceedings of the ACM on Human-Computer Interaction, 7*(CSCW1), Article 136. https://doi.org/10.1145/3579612
- Eisbach, S., Langer, M., & Hertel, G. (2023). Optimizing human-AI collaboration: Effects of motivation and accuracy information in AI-supported decision-making. *Computers in Human Behavior: Artificial Humans, 1*(2), Article 100015.https://doi.org/10.1016/j.chbah.2023.100015
- Gomez, C., Cho, S. M., Ke, S., Huang, C.-M., & Unberath, M. (2025). Human-AI collaboration is not very collaborative yet: A taxonomy of interaction patterns in AI- assisted decision making from a systematic review. *Frontiers in Computer Science, 6*, Article 1521066. https://doi.org/10.3389/fcomp.2024.1521066
- Jiang, Z., Argote, L., & Singh, P. V. (2024). Learning in human–AI collaboration [SSRN preprint]. https://ssrn.com/abstract=5310152
- Li, H., & Tian, F. (2026). Advancing decision-making through AI-human collaboration: A systematic review and conceptual framework. *Group Decision and Negotiation, 35*, 25-36. https://doi.org/10.1007/s10726-025-09890-1
- Raftopoulos, M., & Hamari, J. (2023). Human-AI collaboration in organisations: A literature review on enabling value creation. In *Proceedings of the 31st European Conference on Information Systems (ECIS 2023)*, Kristiansand, Norway.
- Van Rooy, D. (2024). Human-machine collaboration for enhanced decision-making in governance. *Data & Policy, 6*, Article e60. https://doi.org/10.1017/dap.2024.72
- **Horowitz, M. C.** (2024). Bending the automation bias curve: A study of human and AI decision-making. *International Studies Quarterly, 68*(2), Article sqae020. https://doi.org/10.1093/isq/sqae020
- **Leitão, D., Saleiro, P., Figueiredo, M. A. T., & Bizarro, P.** (2022). Human-AI collaboration in decision-making: Beyond learning to defer. *arXiv preprint arXiv:2206.13202*. https://arxiv.org/abs/2206.13202
- **Romeo, G.** (2025). Exploring automation bias in human–AI collaboration. *AI & Society*. Advance online publication. https://doi.org/10.1007/s00146-025-02422-7
- **National Academies of Sciences, Engineering, and Medicine.** (2022). *Human-AI teaming: State-of-the-art and research needs*. The National Academies Press. https://doi.org/10.17226/26355