# AI That Makes You Think: Designing Systems for Guided Reasoning and Reflection

NAILE B. HACIOGLU, Dublin City University, Ireland
MARIA CHIARA LEVA, Technological University Dublin, Ireland
HYOWON LEE, Dublin City University, Ireland

Current AI-augmented reasoning systems often optimize decision-making by providing rapid, automated insights. However, this can lead to cognitive overload and over-reliance, undermining human critical thinking. This paper explores how AI can take on the role of a structured reasoning guide rather than a passive assistant by actively shaping cognitive engagement. We propose that AI should strategically introduce guided reflection pauses, scaffold reasoning skills, and track cognitive progress, ensuring users actively engage with reasoning tasks rather than passively consuming AI-generated insights. Our framework adapts principles from intelligent tutoring systems (ITS), ensuring that AI fosters structured problem-solving and metacognitive growth rather than replacing human thought.

Additional Key Words and Phrases: AI-augmented reasoning, Cognition-aware design, Guided reflection, Structured reasoning, Metacognitive scaffolding, Sustained attention, Human-AI interaction

**ACM Reference Format:**
Naile B. Hacioglu, Maria Chiara Leva, and Hyowon Lee. 2025. AI That Makes You Think: Designing Systems for Guided Reasoning and Reflection. 1, 1 (August 2025), 6 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

## 1 Introduction

AI-augmented reasoning is increasingly prevalent in fact-checking, argument mining, and decision-support systems. These cognitive assistants support human reasoning by providing AI-based feedback that helps users manage knowledge while refining their analysis and judgment on specific topics. Enabled by techniques such as argumentation mining, fact-checking, crowdsourcing, attention nudging, and large language models, AI-augmented reasoning systems offer real-time feedback on logical reasoning, help users identify flawed arguments and misinformation, suggest counter-arguments, provide evidence-based explanations, and foster deeper reflection [5, 8].

However, research suggests that higher confidence in AI capabilities can lead to diminished critical engagement, particularly in routine tasks, potentially eroding independent problem-solving skills [15]. Studies by Danry et al. [7] highlight that AI-generated explanations framed as questions can improve logical discernment, while other research warns that deceptive AI systems can manipulate decision-making if users trust them uncritically [9]. This paper argues

This paper was presented at the 2025 ACM Workshop on Human-AI Interaction for Augmented Reasoning (AIREASONING-2025-01). This is the authors' version for arXiv.

Authors' Contact Information: Naile B. Hacioglu, Dublin City University, Dublin, Ireland; Maria Chiara Leva, Technological University Dublin, Dublin, Ireland; Hyowon Lee, Dublin City University, Dublin, Ireland.

that AI should function as an active guide, actively structuring the reasoning process rather than merely supporting it. Instead of passively providing insights, AI should act as an interactive mentor, prompting users at key cognitive moments, controlling the flow of information, and ensuring that users engage in deliberate reflection before conclusions are drawn.

## 2 Background and Related Work

AI-augmented reasoning systems are widely used in information validation, automated argument analysis, and decision-making applications. While these systems enhance efficiency, research indicates they can also contribute to cognitive overload by presenting users with excessive information at once [14]. Cognitive science studies show that high cognitive load can inhibit deep reasoning and make users more susceptible to AI-driven suggestions without sufficient evaluation [22].

Historically, cognitive science and human-computer interaction (HCI) research have emphasized the importance of structured reasoning and deliberate reflection in learning and decision-making processes. Studies show that human cognition is context-dependent, meaning users rely on external tools and interfaces to offload cognitive effort [19]. Traditional models of learning, such as Bloom's taxonomy and constructivist learning theories, emphasize the importance of active learning and self-reflection in problem-solving, both of which AI should facilitate rather than replace [2].

Additionally, research in intelligent tutoring systems (ITS) and adaptive learning technologies provides insights into how AI can support human reasoning. Systems that personalize feedback, provide structured prompts, and encourage reflection have been shown to improve knowledge retention and decision-making accuracy [16]. These findings suggest that AI reasoning tools should be designed with adaptive guidance mechanisms that allow users to engage critically with information rather than passively accepting AI-generated conclusions.

## 3 Cognitive Challenges in AI-Augmented Reasoning

### 3.1 Over-Reliance and Diminished Critical Thinking

Research has shown that users become dependent on AI-generated outputs, often reducing engagement in active reasoning [15]. When AI prioritizes efficiency, users may accept conclusions without questioning validity [7]. Over time, this dependence can erode domain expertise and problem-solving abilities, as users gradually defer reasoning tasks to AI systems rather than applying their own analytical skills [19].

Additionally, studies suggest that excessive AI reliance can lead to automation bias, where users develop an uncritical trust in AI-generated recommendations despite evidence of errors [1, 18]. This phenomenon has been observed in fields such as healthcare and finance, where professionals may override their own expertise in favor of AI-suggested diagnoses or investment strategies [10, 21].

### 3.2 Cognitive Overload in AI-Supported Decision-Making

Many AI tools present excessive information at once, leading to decision fatigue and reducing the ability to evaluate competing perspectives [4]. AI interfaces should pace information flow to optimize cognitive engagement rather than overwhelm users.

Cognitive load theory suggests that information processing is constrained by the limits of working memory, and poorly structured interfaces exacerbate cognitive strain. When AI generates large volumes of text, multiple decision pathways, or conflicting evidence, users might struggle to process the information effectively, leading to shallow processing and heuristic-based decision-making rather than deep analysis [11].

## 4 Beyond Answers: AI As A Mentor For Thoughtful Reasoning

AI-augmented reasoning systems should go beyond simply providing answers—they should actively guide users through structured cognitive processes to foster deeper engagement, critical reflection, and independent problem-solving. Rather than acting as passive assistants that supply information on demand, AI should function as an interactive mentor that encourages users to evaluate evidence, reflect on their thought process, and refine their reasoning over time.

To achieve this, AI systems must integrate adaptive reasoning scaffolding, allowing users to pause, reflect, and adjust their conclusions before receiving AI-generated insights. Drawing from cognitive psychology and intelligent tutoring systems (ITS), AI can facilitate structured learning by providing guided reflection pauses, interactive feedback loops, and metacognitive tracking mechanisms that reinforce active reasoning.

Controlled information flow and strategic pauses, as inspired by Attention Mode from our previous work [12, 13], should not merely be available to users but should be intelligently inserted into reasoning workflows based on user engagement levels and cognitive patterns and activated [17].

We propose three key strategies AI can adopt to support structured reasoning:

(1) Encouraging metacognition through reflection pauses, where users actively engage with their reasoning process before receiving AI feedback.
(2) Visualizing learning trajectories, enabling users to track their cognitive progress and refine their decision-making strategies over time.
(3) Guiding self-explanation techniques, prompting users to articulate their reasoning and evaluate counterarguments to enhance critical thinking.

By implementing these techniques, AI systems can help users develop long-term cognitive resilience, ensuring they do not merely consume AI-generated insights but actively engage with and refine their reasoning abilities.

### 4.1 Encouraging Metacognition Through Reflection Pauses

AI can introduce deliberate reflection pauses by detecting moments when users are rushing through reasoning steps. Instead of providing constant feedback, AI can detect cognitive fatigue and strategically introduce prompts that ask users to summarize, re-evaluate, or articulate their thought process.

#### *4.1.1 Example: Cognitive Checkpoints in AI-Augmented Reasoning.*

- Prompt users to explain why they arrived at a certain conclusion before revealing AI feedback: Instead of immediately presenting AI-generated insights, the system would require users to articulate their reasoning process. By encouraging users to verbalize or write down their thought process, the AI ensures they engage in deliberate reflection before receiving external validation.
- Encourage users to reconsider counterarguments before finalizing a decision: Before users settle on a conclusion, the AI could present them with alternative perspectives or counterarguments. This would challenge them to critically evaluate different viewpoints and assess the strength of their original stance, reducing confirmation bias and promoting deeper reasoning.
- Introduce structured note-taking spaces where users can externalize thoughts before moving to the next step [3]. The system would provide an interactive note-taking interface, allowing users to summarize key insights,

document their reasoning, or outline supporting evidence. By externalizing their thoughts, users would develop a clearer understanding of their own cognitive process, reinforcing long-term retention and structured thinking.

### 4.2 Visualizing Learning Trajectories for Reasoning

Just as intelligent tutoring systems (ITS) track skill mastery, AI reasoning tools should provide progress tracking systems that visually represent users' critical thinking development.

*4.2.1 Example: Focused Reasoning Mode for Managing Information Flow.* Users could toggle "guided mode" to receive incremental AI interventions instead of full solutions. Rather than presenting all available reasoning insights at once, this mode would introduce information in stages, ensuring users engage with each step of the reasoning process before moving forward.

AI could visualize progress in metacognition using a reasoning dashboard that tracks key cognitive indicators:

- Logical consistency – The system would analyze how consistently users apply reasoning principles across different arguments. If contradictions are detected in their thought process, the AI could highlight these inconsistencies and prompt users to reconsider their conclusions.
- Bias detection – By analyzing users' interaction patterns and argument preferences, the AI could identify tendencies where users might favor certain viewpoints. When biases are detected, the system could present alternative perspectives, encouraging users to engage with a broader range of viewpoints.
- Depth of analysis – The AI would evaluate whether users engage in multi-layered reasoning or rely on surface-level conclusions. If shallow reasoning is detected, the system could prompt users with targeted reflection exercises to deepen their engagement with the topic.

### 4.3 AI-Guided Self-Explanation Techniques

Encouraging users to explain their reasoning in their own words has been shown to improve comprehension, retention, and logical accuracy [6, 20]. AI can integrate self-explanation exercises into reasoning workflows.

*4.3.1 Example: Adaptive AI Prompts Inspired by Cognitive Checkpoints.*

- "How did you arrive at this conclusion?" This prompt encourages users to break down their reasoning process step by step, making their thought process explicit. By requiring users to articulate how they reached their conclusion, the AI fosters metacognitive awareness, helping users identify potential gaps or biases in their thinking.
- "What assumptions does this argument rely on?" This question prompts users to examine the underlying premises of their reasoning. Many arguments rest on implicit assumptions, and by making these explicit, users can critically assess whether their reasoning holds in different contexts or if any unexamined biases are influencing their thought process.
- "What are two counterarguments that challenge this perspective?" This prompt encourages users to engage with opposing viewpoints, fostering a more balanced and comprehensive approach to reasoning. By considering counterarguments, users refine their analytical skills and strengthen their ability to evaluate multiple perspectives before reaching a final decision.

## 5 Future Research Directions

As AI reasoning tools evolve, further research is necessary to assess their effectiveness and cognitive impact. Several areas of study would help refine cognition-aware AI interfaces and optimize their role in decision-making and learning environments.

First, research should focus on understanding the impact of structured AI guidance. While structured AI feedback is designed to promote deliberate thinking, it remains unclear how different levels of intervention affect users' engagement and reasoning quality. Investigating whether structured guidance helps users retain reasoning skills beyond AI-assisted sessions will be critical in evaluating its long-term benefits.

Second, future studies should explore developing adaptive reflection mechanisms that dynamically adjust to individual users' cognitive states. AI should be able to recognize signs of cognitive fatigue, hasty decision-making, or overconfidence and provide appropriately timed reflection prompts. Experimenting with different reflection strategies—such as delaying feedback, prompting users with open-ended questions, or requiring justifications for conclusions—could help determine which methods best support deep reasoning.

Third, researchers should investigate AI-supported collaborative reasoning, analyzing how AI can facilitate group discussions and collective decision-making. Instead of acting as a single-user tool, AI could mediate structured debates, highlight conflicting viewpoints, and encourage users to construct counterarguments in real time. Studying how AI can foster productive disagreements and knowledge sharing will be key to designing reasoning tools that enhance collective intelligence rather than reinforcing individual biases.

Finally, future research should evaluate the long-term cognitive impacts of AI-assisted reasoning. If users frequently rely on AI-generated insights, there is a risk of developing passive engagement habits that diminish independent problem-solving skills. Longitudinal studies should examine whether prolonged exposure to AI-guided reasoning strengthens or weakens users' ability to think critically without AI assistance. Understanding how these tools shape reasoning behaviors over time will help developers refine AI systems to maximize cognitive benefits while mitigating dependency risks.

## 6 Conclusion

Rather than serving as a passive support system, AI should be designed as an active mentor, structuring the reasoning process in a way that promotes deep engagement, critical reflection, and sustained learning. By adopting cognition-aware design principles, AI systems can empower users to engage in deliberate, thoughtful reasoning, balancing efficiency and cognitive engagement.

Moving forward, AI developers, cognitive scientists, and HCI researchers should work collaboratively to ensure that AI-augmented reasoning tools are built with human cognitive well-being in mind, rather than merely optimizing for speed and convenience.

## Acknowledgments

This work was conducted with the financial support of the Research Ireland Centre for Research Training in Digitally-Enhanced Reality (d-real) under Grant No. 18/CRT/6224 and the Research Ireland Centre Grant No. 12/RC/2289_P2 at Insight Centre for Data Analytics at Dublin City University. 

## References


[1] Reuben Binns, Max Van Kleek, Michael Veale, Ulrik Lyngs, Jun Zhao, and Nigel Shadbolt. 2018. 'It's Reducing a Human Being to a Percentage': Perceptions of Justice in Algorithmic Decisions *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3173574.3173951

[2] Benjamin S. Bloom, Max D. Engelhart, Edward J. Furst, Walker H. Hill, and David R. Krathwohl. 1956. *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain.* Longmans, Green, New York, NY, USA.

[3] Mark Bohay, Daniel Blakely, and Andrea Tamplin. 2011. Note Taking, Review, Memory, and Comprehension. *The American journal of psychology* 124 (01 2011), 63–73. doi:10.5406/amerjpsyc.124.1.0063

[4] John Buchanan and Ned Kock. 2001. *Information Overload: A Decision Making Perspective.* 49–58. doi:10.1007/978-3-642-56680-6_4

[5] Ed H. Chi. 2009. Augmented social cognition: using social web technology to enhance the ability of groups to remember, think, and reason. In *Proceedings of the 2009 ACM SIGMOD International Conference on Management of Data* (Providence, Rhode Island, USA) *(SIGMOD '09)*. Association for Computing Machinery, New York, NY, USA, 973–984. doi:10.1145/1559845.1559959

[6] Michelene T.H. Chi, Nicholas De Leeuw, Mei Hung Chiu, and Christian Lavancher. 1994. Eliciting self-explanations improves understanding. *Cognitive Science* 18, 3 (1994), 439–477. doi:10.1016/0364-0213(94)90016-7

[7] Valdemar Danry, Pat Pataranutaporn, Ziv Epstein, Matthew Groh, and Pattie Maes. 2022. Deceptive AI Systems That Give Explanations Are Just as Convincing as Honest AI Systems in Human-Machine Decision Making. arXiv:2210.08960 [cs.CY] https://arxiv.org/abs/2210.08960

[8] Valdemar Danry, Pat Pataranutaporn, Yaoli Mao, and Pattie Maes. 2023. Don't Just Tell Me, Ask Me: AI Systems that Intelligently Frame Explanations as Questions Improve Human Logical Discernment Accuracy over Causal AI explanations. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 352, 13 pages. doi:10.1145/3544548.3580672

[9] Douglas C. Engelbart. 1962. *Augmenting Human Intellect: A Conceptual Framework.* Technical Report. Stanford Research Institute. https://www.dougengelbart.org/pubs/papers/scanned/Doug_Engelbart-AugmentingHumanIntellect.pdf Retrieved from https://www.dougengelbart.org/pubs/papers/scanned/Doug_Engelbart-AugmentingHumanIntellect.pdf.

[10] Kate Goddard, Abdul Roudsari, and Jeremy C. Wyatt. 2012. Automation Bias: A Systematic Review of Frequency, Effect Mediators, and Mitigators. *Journal of the American Medical Informatics Association: JAMIA* 19, 1 (2012), 121–127. doi:10.1136/amiajnl-2011-000089

[11] Nikolos Gurney, John H. Miller, and David V. Pynadath. 2023. The role of heuristics and biases during complex choices with an AI teammate *(AAAI'23/IAAI'23/EAAI'23)*. AAAI Press, Article 672, 9 pages. doi:10.1609/aaai.v37i5.25741

[12] Naile Hacioglu, Maria Chiara Leva, and Hyowon Lee. 2023. Designing Interaction to Support Sustained Attention. In *Human-Computer Interaction – INTERACT 2023: 19th IFIP TC13 International Conference, York, UK, August 28 – September 1, 2023, Proceedings, Part IV* (York, United Kingdom). Springer-Verlag, Berlin, Heidelberg, 454–458. doi:10.1007/978-3-031-42293-5_53

[13] Naile Hacioglu, Maria Chiara Leva, and Hyowon Lee. 2023. Re-designing the Interaction of Day-to-Day Applications to Support Sustained Attention Level. In *HCI International 2023 – Late Breaking Papers: 25th International Conference on Human-Computer Interaction, HCII 2023, Copenhagen, Denmark, July 23–28, 2023, Proceedings, Part I* (Copenhagen, Denmark). Springer-Verlag, Berlin, Heidelberg, 363–377. doi:10.1007/978-3-031-48038-6_23

[14] Ralph Hertwig and Till Grüne-Yanoff. 2017. Nudging and Boosting: Steering or Empowering Good Decisions. *Perspectives on Psychological Science* 12, 6 (2017), 973–986. doi:10.1177/1745691617702496

[15] Hao-Ping (Hank) Lee, Advait Sarkar, Lev Tankelevitch, Ian Drosos, Sean Rintel, Richard Banks, and Nicholas Wilson. 2025. The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. doi:10.1145/3706598.3713778

[16] Wenting Ma, Olusola O. Adesope, John C. Nesbit, and Qing Liu. 2014. Intelligent Tutoring Systems and Learning Outcomes: A Meta-Analysis. *Journal of Educational Psychology* 106, 4 (2014), 901–918. doi:10.1037/a0037123

[17] Christopher Miller and Michael Dorneich. 2006. *From Associate Systems to Augmented Cognition: 25 Years of User Adaptation in High Criticality Systems*. 344–353.

[18] Raja Parasuraman, Thomas Sheridan, and Christopher Wickens. 2000. A model for types and levels of human interaction with automation. IEEE Trans. Syst. Man Cybern. Part A Syst. Hum. 30(3), 286-297. *IEEE transactions on systems, man, and cybernetics. Part A, Systems and humans : a publication of the IEEE Systems, Man, and Cybernetics Society* 30 (06 2000), 286–97. doi:10.1109/3468.844354

[19] Evan F. Risko and Sam J. Gilbert. 2016. Cognitive Offloading. *Trends in Cognitive Sciences* 20, 9 (2016), 676–688. doi:10.1016/j.tics.2016.07.002

[20] Bethany Rittle-Johnson. 2006. Promoting Transfer: Effects of Self-Explanation and Direct Instruction. *Child Development* 77, 1 (2006), 1–15. doi:10.1111/j.1467-8624.2006.00852.x arXiv:https://srcd.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-8624.2006.00852.x

[21] Shivam Shukla and Sudhir Kumar Shukla. 2023. Impact of Automation Bias and Status Quo Bias on Capital Market Investment Decisions of Indian Investors: An Explanatory Research. *Financial Markets, Institutions and Risks* 7, 3 (2023), 1–11. doi:10.61093/fmir.7(3).1-11.2023

[22] John Sweller, Paul Ayres, and Slava Kalyuga. 2011. Cognitive Load Theory. (01 2011). doi:10.1007/978-1-4419-8126-4