Augmenting Intuition: Expanding AI-Mediated Reasoning Beyond Reflective Interventions

Valdemar Danry
vdanry@mit.edu
MIT Media Lab
Cambridge, Massachusetts, USA

Pattie Maes
pattie@media.mit.edu
MIT Media Lab
Cambridge, Massachusetts, USA

Abstract

Interactive AI systems increasingly mediate human reasoning by generating explanations, arguments, and recommendations on demand. Current tools for thought primarily support critical thinking by prompting reflection—encouraging users to slow down, justify conclusions, or consider alternatives. While effective, these approaches are effortful, interrupt workflows, and do not directly improve the intuitive processes that generate most judgments. We argue that AI-mediated reasoning should move beyond reflection-centric interaction toward supporting the development of stronger intuitive reasoning. Drawing on cognitive science and human–AI interaction, we introduce a framework that organizes reasoning augmentation techniques into five categories spanning reflection- and intuition-oriented interventions. We further outline how adaptive systems that sense cognitive signals and model reasoning tendencies can selectively trigger reflection and shape intuitive responses over time. This perspective expands tools for thought from momentary decision support toward systems that cultivate lasting improvements in human reasoning.

Keywords

Intuitions, Reasoning Augmentation, Tools for Thought, Artificial Intelligence, AI, Reflection, Cognitive Science, Dual Process Theory

ACM Reference Format:
Valdemar Danry and Pattie Maes. 2018. Augmenting Intuition: Expanding AI-Mediated Reasoning Beyond Reflective Interventions. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Conference acronym 'XX).* ACM, New York, NY, USA, 6 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Reasoning is increasingly mediated by interactive AI systems [8, 41]. Large language models can generate arguments, explanations, and recommendations on demand, transforming how people evaluate information and make decisions. This has led to a growing interest in reasoning augmentation: designing systems that do not merely produce answers, but help users become better reasoners through interaction [8, 41]. A dominant strategy in reasoning augmentation has been to encourage users to reflect — to slow down, reconsider initial responses, justify conclusions, or examine alternative interpretations. Interactive systems prompt users to explain their reasoning, present counterarguments, or answer reflective questions with the goal of promoting deeper cognitive engagement [5, 9, 13]. Educational AI systems similarly adopt Socratic dialogue and metacognitive scaffolding to guide users toward more deliberate reasoning processes [9, 20, 21]. Across domains, reasoning augmentation is largely treated as a problem of inducing reflection.

This emphasis on reflection is well motivated. Reflection can improve reasoning accuracy by enabling users to reconsider misleading initial impressions, attend to relevant evidence, search for additional information and revise incorrect conclusions. Prompts that encourage users to evaluate accuracy, explain their reasoning, or consider alternatives can improve information discernment and analytical performance [27, 33, 34, 40]. In human–AI interaction, such reflective interventions are widely used to improve users' evaluation of AI-generated outputs and support more thoughtful decision-making [31]. These approaches assume that better reasoning emerges from engaging users in more deliberate, reflective cognitive processes.

However, reflection is not the default mode of human cognition[22]. Most judgments arise rapidly and automatically, accompanied by intuitive feelings of rightness, coherence, or conflict that guide acceptance or rejection before deliberate reasoning begins [10, 32]. Reflection typically occurs only intermittently — when intuitive responses feel uncertain, when stakes are high, or when cognitive resources permit extended deliberation. And it often requires us to shift our attention and give up whatever else we were currently doing. As a result, reflection is inherently sparse and costly [38]. Interactive systems that rely on prompting reflection must compete with users' ongoing goals, limited attention, and cognitive constraints. Even well-designed reflective prompts introduce friction, interrupt workflows, and depend on users' willingness to engage in effortful reasoning [5, 26]. This creates a fundamental limitation: reasoning augmentation strategies that rely exclusively on reflection can improve reasoning locally, but do not directly improve the intuitive processes that generate most judgments.

Cognitive science suggests that reasoning competence depends not only on reflective abilities, but also on the quality of intuitive responses, or so-called reflexive decisions, themselves. With training and experience, individuals can develop more reliable intuitive judgments, including rapid sensitivity to logical inconsistencies and misleading information [4, 11, 14]. These findings challenge the assumption that reasoning augmentation must operate primarily by triggering reflection. Instead, they suggest an alternative possibility: interactive systems can augment reasoning by shaping the

Conference acronym 'XX, Woodstock, NY

ACM ISBN 978-1-4503-XXXX-X/2018/06
https://doi.org/XXXXXXX.XXXXXXX

intuitive processes that guide users' initial judgments. Rather than repeatedly interrupting users to induce reflection, systems can help users develop better intuitions—improving reasoning at its source.

Despite rapid advances in AI-assisted reasoning, much existing work still operationalizes reasoning support primarily through explicit reflective scaffolds during interaction, such as explanations, metacognitive prompts, Socratic questioning, and critique [3, 13, 21, 31, 39, 41, 42]. Some of this work also aims to cultivate longer-term habits of reflection, metacognitive oversight, or learning through repeated interaction [21, 42]. However, the dominant mechanism of support remains explicit deliberative engagement at interaction time: users are asked to slow down, inspect, justify, compare, or reconsider. As a result, reasoning augmentation has focused far more on eliciting reflection than on directly shaping the intuitive appraisals, cue sensitivities, and learned default responses that guide most everyday judgments. This reveals an important imbalance in current approaches: substantially more attention has been given to triggering reflection than to designing systems that shape intuition. As a result, reasoning augmentation has focused far more on eliciting reflection than on directly shaping the intuitive appraisals, cue sensitivities, and learned default responses that guide most everyday judgments.

In this paper, we argue that reasoning augmentation should move beyond reflection-centric interaction toward supporting the development of better intuitions. We introduce a framework that conceptualizes reasoning augmentation as an adaptive process that senses users' cognitive and behavioral signals, models their reasoning tendencies, and shapes intuitive judgments over time. We synthesize evidence from human–AI interaction, cognitive science, and decision-making to map the design space of reasoning augmentation strategies and identify opportunities for systems that support both reflection and intuition. By shifting the focus from prompting reflection to shaping intuition, we outline a path toward reasoning augmentation systems that not only help users reason in the moment, but help them become better reasoners over time.

2 Defining Intuition and Reflection

Reasoning augmentation work often treats "better reasoning" as "more reflection": slowing down, examining evidence, and explicitly justifying conclusions [5, 8, 41]. In contrast, *intuitions* are fast appraisals—often accompanied by affective signals like confidence, coherence, or conflict—that guide acceptance or rejection before deliberate reasoning begins [10, 44]. *Reflection* refers to slower, effortful operations such as rule-based evaluation, hypothesis testing, and revising conclusions [22] — or experientially as "mulling over a problem until it feels right"[19]. These modes are not strict stages: intuitive appraisals can trigger reflection when uncertainty or conflict is detected, and reflection can shape later intuitions through learning. However, reflection is sparse and costly [38]. It competes with ongoing goals, consumes attention, and introduces friction [22]. As a result, reflection-centric systems can improve reasoning locally, but they do not directly improve the intuitive processes that generate most everyday judgments.

3 Intuition as a "Superpower"

Although intuition is often discussed as a source of bias, it can become highly reliable through training and experience. In many domains, expertise appears as rapid, non-deliberative discrimination. Radiologists, for example, learn to detect abnormalities in medical images within milliseconds, and perceptual learning work shows that this speeded sensitivity can be accelerated through structured exposure and feedback [24, 49, 50]. In high-stakes practice, experienced firefighters rely on recognition-primed decision making: they match situational patterns to action schemas rather than analytically comparing options [25]. Related effects occur in reasoning: training on logical structures or informal fallacies improves later detection of flawed arguments, suggesting that "something is wrong" can become a default appraisal rather than a reflective achievement [4, 14]. These "superpowers" reflect learned cue–response mappings strengthened through feedback and deliberate practice [12]. With sufficient training, conflict detection becomes faster and error signals more reliable, allowing correct judgments to emerge earlier in processing.

4 Intuition is Cue-Driven and Designable

Intuitive appraisals depend not only on what people have learned, but on which cues the environment makes salient [16]. Interfaces and representations can therefore shape intuition by suppressing misleading cues, amplifying diagnostic cues, or altering the intensity of intuitive reactions[17]. For example, systems might reduce reliance on irrelevant identity cues (e.g., removing names or photos), highlight indicators of weak evidence, or introduce triggers that make reasoning failures immediately feel wrong (e.g., making ad hominem attacks perceptually salient as a fallacy). This suggests a path for reasoning augmentation that does not require constant deliberate engagement: rather than repeatedly interrupting users to reflect, systems can reshape the conditions under which intuitions form, cultivating fast "error signals" that guide when reflection is needed.

5 Implications for Reasoning Augmentation

Taken together, intuition-centered augmentation complements reflection-centric approaches. Reflection remains essential for novel, ambiguous, or high-stakes problems, but intuition can be improved at its source: by changing cues that trigger appraisals, reshaping how cues are interpreted, and training new default sensitivities over time. In the next section, we synthesize prior work to outline a design space with five categories—Always-On Reflection, Selective Reflection, Cued Intuitions, Shaping Intuitions, and Creating New Intuitions—spanning momentary versus developmental interventions and reflective versus intuitive targets.

6 Design Space

Reasoning augmentation can be implemented through a range of interaction techniques that intervene at different points in the cognitive process. Some techniques explicitly prompt reflective reasoning, while others modify intuitive appraisals or support the development of new intuitive competencies over time. These techniques differ in cognitive cost, temporal scope, and durability of impact. We organize prior work into five categories: Always-On

Reflection, Selective Reflection, Cued Intuitions, Shaping Intuitions, and Creating New Intuitions.

6.1 Always-On Reflection

Always-on reflection techniques systematically require users to stop and think and engage in deliberate reasoning. These techniques introduce structured friction that prevents passive acceptance of information and encourages users to articulate, justify, and evaluate their conclusions. They are widely used in educational systems, decision aids, and AI-assisted reasoning tools. **Interaction techniques.** Cognitive forcing functions require users to generate conclusions, rationales, or confidence estimates before receiving AI output [5, 48]. Socratic prompting guides users through structured questions rather than providing answers [9, 21]. Devil's advocate techniques introduce counterarguments or alternative explanations [6]. Structured comparison techniques require users to evaluate multiple alternatives explicitly [2]. Metacognitive calibration prompts ask users to assess confidence and uncertainty [33, 34]. Fallacy identification tools highlight reasoning weaknesses and prompt revision [54]. **Why it works.** These techniques improve reasoning by forcing articulation, expanding the hypothesis space, and improving metacognitive awareness, reducing uncritical acceptance of flawed conclusions. **Failure modes.** Continuous reflection introduces cognitive cost, workflow disruption, and fatigue. Overuse may reduce engagement or create dependency on external scaffolding.

6.2 Selective Reflection

Selective reflection techniques aim to trigger reflection only when necessary. Rather than continuously imposing cognitive effort, these techniques identify situations in which intuitive reasoning is likely to fail and selectively introduce reflective interventions. This approach reduces cognitive burden while preserving the benefits of reflection. **Interaction techniques.** Systems trigger reflection when detecting uncertainty, disagreement, or high error likelihood [48]. Progressive disclosure provides partial hints before full solutions [1]. Confidence-triggered prompts intervene when users express strong confidence in incorrect responses [42]. **Why it works.** Reflection is deployed when its expected benefit is highest, aligning with evidence that reflective reasoning is selectively activated in response to conflict [10]. **Failure modes.** Effectiveness depends on accurately predicting when intervention is necessary. Incorrect timing may frustrate users or miss critical reasoning failures.

6.3 Cued Intuition

Cued intuition techniques modify the cues that shape intuitive judgments without requiring explicit reflection. Rather than prompting deliberate reasoning, these techniques alter representations, interfaces, or environments to influence fast, automatic appraisals. These interventions operate directly on intuitive processes by modifying the signals that guide immediate evaluation. **Interaction techniques.** Choice architecture modifies defaults, ordering, and salience to influence intuitive judgments [43, 52]. Framing effects demonstrate that linguistic and visual presentation alters intuitive evaluation [46]. Cue removal, such as anonymization, reduces bias-triggering signals [18]. Language priming alters intuitive reasoning through linguistic context [15]. Cue amplification highlights diagnostic features such as logical inconsistencies [54]. **Why it works.** Intuition relies on cue-based appraisal. Modifying cue salience alters intuitive responses directly, enabling faster and more reliable judgments without explicit analysis. **Failure modes.** Effects may be context-dependent and raise ethical concerns if users are unaware of how cues are modified.

6.4 Shaping Intuitions

Shaping intuition techniques modify how users interpret cues, altering intuitive appraisals without necessarily requiring explicit reasoning. These techniques focus on restructuring interpretation rather than merely modifying cues. **Interaction techniques.** Cognitive reframing modifies how information is construed [37, 46]. Fallacy labeling transforms persuasive rhetoric into diagnostic categories [30]. Argument mapping restructures reasoning into visual representations that support intuitive recognition of logical relationships [47]. **Why it works.** Intuitive judgments depend on interpretation. Changing interpretation alters intuitive appraisal and improves reasoning reliability. **Failure modes.** Effects may depend on user acceptance and may not generalize across contexts.

6.5 Creating New Intuitions

Creating new intuitions involves training users to develop reliable intuitive reasoning abilities through repeated interaction, feedback, and exposure. These techniques focus on long-term learning rather than immediate intervention. **Interaction techniques.** Perceptual learning accelerates rapid discrimination through structured exposure and feedback [24]. Logical intuition training improves automatic conflict detection [4]. Argument mapping builds structural reasoning competence [47]. Inoculation training builds resistance to manipulation [45]. Sensorimotor augmentation demonstrates that new intuitive perceptions can be learned through repeated association [23]. **Why it works.** Intuition reflects learned cue–response mappings. Training strengthens these mappings, enabling faster and more reliable reasoning. **Failure modes.** Training requires time, and learned intuitions may not transfer across domains.

7 Realizing Selective Reflection and Intuition Augmentation

Realizing the full design space of reasoning augmentation requires systems that can determine when reflection is necessary, which intuitive cues to modify, and how reasoning tendencies evolve over time. This capability depends on integrating three components: sensing cognitive and affective signals, creating cognitive maps of reasoning behavior, and adaptively deploying interaction techniques based on these models.

7.1 Sensing Reflection & Intuition Signals

Recent work demonstrates that cognitive states relevant to reasoning—such as uncertainty, engagement, cognitive conflict, and intuitive reactions—can be inferred from multimodal behavioral and physiological signals. Eye gaze provides a continuous measure of attention and comprehension and can be used to personalize LLM assistance and trigger targeted scaffolding [35]. Facial expressions

Category	Cognitive cost	Timescale	Interaction techniques (examples)
Always-On Reflection: Force deliberate reasoning	High	Immediate	Cognitive forcing functions (user generates answer/rationale before AI) [5]; Socratic prompting [9, 21]; devil's advocate counterarguments [6]; structured comparison of alternatives [2]; metacognitive calibration prompts [33, 34, 42]; fallacy identification and critique [54].
Selective Reflection: Trigger reflection adaptively	Medium	Immediate	Reflection prompts triggered by uncertainty/disagreement/predicted error [48]; progressive disclosure of AI assistance [1]; confidence-triggered reflective interventions [42].
Cued Intuitions: Modify intuitive cues	Low	Immediate	Choice architecture / digital nudging (defaults, salience, ordering) [43, 52]; framing/representation changes [46]; cue removal/anonymization [18]; language priming [15]; cue amplification (highlight fallacies/weak evidence) [54].
Shaping Intuitions: Alter felt meaning	Medium	Intermediate	Cognitive reframing / reinterpretation scaffolds [37, 46]; fallacy labeling and category transformation [30]; argument mapping and reasoning visualization [47].
Creating New Intuitions: Train intuitive competence	Low (per instance)	Long-term	Perceptual learning with structured feedback [24]; logical intuition training [4, 14]; argument mapping practice [47]; inoculation / prebunking [45]; sensorimotor contingency learning / sensory augmentation [23].

Table 1: Design space of reasoning augmentation interaction techniques, organized by cognitive target, cost, and temporal scope.

and behavioral engagement signals similarly provide real-time indicators of confusion, attention, and cognitive effort [29]. Physiological measures such as electrodermal activity (EDA) enable detection of arousal and cognitive load, supporting adaptive intervention policies that adjust assistance based on inferred cognitive state [7]. More direct sensing of interoceptive signals—including emerging approaches to detecting "gut feeling" responses—demonstrates that intuitive reactions themselves can be measured and modeled [28]. Together, these signals provide a substrate for selective reflection, enabling systems to intervene when users are likely to benefit from additional reasoning support.

7.2 Creating Cognitive Maps of Reasoning Behavior

Beyond momentary sensing, enabling intuition-aware augmentation requires longitudinal models of reasoning tendencies. Interaction traces from everyday activities—such as writing, search, and communication—contain stable patterns reflecting heuristics, biases, and reasoning strategies. Prior work shows that user models can be constructed from behavioral traces including email interaction patterns, document editing, and activity histories, enabling inference of user goals and behavioral tendencies [36, 53]. Interaction trace analysis has also been used to identify cognitive biases and deviations in decision-making workflows, demonstrating that reasoning tendencies can be inferred from behavioral patterns over time [51]. These models form cognitive maps: structured representations of how individuals reason, including characteristic strengths, vulnerabilities, and intuitive response patterns.

7.3 Adaptive Deployment of Interaction Techniques

Cognitive maps enable adaptive selection of reasoning augmentation strategies. Rather than applying fixed prompts or interface changes, systems can personalize interventions based on user-specific reasoning tendencies and momentary cognitive state. This enables dynamic coordination between reflection-oriented and intuition-oriented interaction techniques. For example, systems may trigger reflective prompting when cognitive maps indicate elevated risk of reasoning error, amplify diagnostic cues when intuitive responses are likely to be misleading, or reinforce reliable intuitive patterns through feedback.

Together, these capabilities enable a shift from static reasoning augmentation toward cognitive-aware systems that continuously model and support human reasoning. By integrating multimodal sensing, longitudinal cognitive maps, and adaptive intervention policies, such systems can selectively deploy reflection, shape intuitive responses, and support the development of stronger reasoning abilities over time.

8 Ethical Considerations

Reasoning augmentation systems that shape intuition and adapt to users' cognitive tendencies introduce significant ethical considerations. Because these systems operate by modifying cues, reinforcing associations, and influencing intuitive appraisals—often outside conscious awareness—they carry risks of manipulation, loss of autonomy, and covert persuasion. Interventions designed to improve reasoning could also be repurposed to steer beliefs or decisions in ways that benefit system designers rather than users. Longitudinal cognitive maps, built from behavioral and physiological data, raise additional concerns about privacy, surveillance, and the potential misuse of sensitive cognitive inferences. Ensuring ethical deployment therefore requires transparency about how systems influence reasoning, meaningful user control over interventions and data collection, and clear alignment with users' epistemic interests. Designers should prioritize interventions that enhance users' independent reasoning competence rather than fostering dependence on adaptive scaffolding. Finally, careful evaluation is needed to ensure that intuition-shaping interventions generalize appropriately and do not inadvertently reinforce biases or degrade reasoning in other domains. By centering user autonomy, transparency, and long-term cognitive benefit, reasoning augmentation systems can support human reasoning without undermining agency or trust.

References

[1] Gagan Bansal, Tongshuang Wu, Joyce Zhou, Raymond Fok, Besmira Nushi, Ece Kamar, Marco Tulio Ribeiro, and Daniel Weld. 2021. Does the whole exceed its parts? the effect of ai explanations on complementary team performance. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–16.

[2] Frederic Becker, Julian Skirzyński, Bas van Opheusden, and Falk Lieder. 2022. Boosting human decision-making with AI-generated decision aids. *Computational Brain & Behavior* 5, 4 (2022), 467–490.

[3] Marit Bentvelzen, Paweł W Woźniak, Pia SF Herbes, Evropi Stefanidi, and Jasmin Niess. 2022. Revisiting reflection in hci: Four design resources for technologies that support reflection. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 1 (2022), 1–27.

[4] Esther Boissin, Serge Caparos, Aikaterini Voudouri, and Wim De Neys. 2022. Debiasing System 1: Training favours logical over stereotypical intuiting. *Judgment and Decision Making* 17, 4 (2022), 646–690.

[5] Zana Buçinca, Maja Barbara Malaya, and Krzysztof Z Gajos. 2021. To trust or to think: cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making. *Proceedings of the ACM on Human-computer Interaction* 5, CSCW1 (2021), 1–21.

[6] Chun-Wei Chiang, Zhuoran Lu, Zhuoyan Li, and Ming Yin. 2024. Enhancing ai-assisted group decision making through llm-powered devil's advocate. In *Proceedings of the 29th International Conference on Intelligent User Interfaces*. 103–119.

[7] Francesco Chiossi, Changkun Ou, and Sven Mayer. 2023. Exploring physiological correlates of visual complexity adaptation: Insights from EDA, ECG, and EEG data for adaptation evaluation in VR adaptive systems. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–7.

[8] Valdemar Danry, Pat Pataranutaporn, Christopher Cui, Jui-Tse Hung, Lancelot Blanchard, Zana Buçinca, Chenhao Tan, Thad Starner, and Pattie Maes. 2025. Human-AI Interaction for Augmented Reasoning: Improving Human Reflective and Critical Thinking with Artificial Intelligence. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–6.

[9] 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*. 1–13.

[10] Wim De Neys. 2014. Conflict detection, dual processes, and logical intuitions: Some clarifications. *Thinking & Reasoning* 20, 2 (2014), 169–187.

[11] Wim De Neys, Sofie Cromheeke, and Magda Osman. 2011. Biased but in doubt: Conflict and decision confidence. *PloS one* 6, 1 (2011), e15954.

[12] K Anders Ericsson, Michael J Prietula, Edward T Cokely, et al. 2007. The making of an expert. *Harvard business review* 85, 7/8 (2007), 114.

[13] Simon WS Fischer, Hanna Schraffenberger, Serge Thill, and Pim Haselager. 2025. A Taxonomy of Questions for Critical Reflection in Machine-Assisted Decision-Making. *arXiv preprint arXiv:2504.12830* (2025).

[14] Nina Franiatte, Esther Boissin, Alexandra Delmas, and Wim De Neys. 2024. Boosting debiasing: Impact of repeated training on reasoning. *Learning and Instruction* 89 (2024), 101845.

[15] Janet Geipel, Constantinos Hadjichristidis, and Luca Surian. 2015. The foreign language effect on moral judgment: The role of emotions and norms. *PloS one* 10, 7 (2015), e0131529.

[16] Gerd Gigerenzer and Wolfgang Gaissmaier. 2011. Heuristic decision making. *Annual review of psychology* 62, 2011 (2011), 451–482.

[17] Gerd Gigerenzer, Ulrich Hoffrage, and Daniel G Goldstein. 2008. Fast and frugal heuristics are plausible models of cognition: Reply to Dougherty, Franco-Watkins, and Thomas (2008). (2008).

[18] Claudia Goldin and Cecilia Rouse. 2000. Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American economic review* 90, 4 (2000), 715–741.

[19] Jonathan Haidt. 2001. The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological review* 108, 4 (2001), 814.

[20] Shirley S Ho, Stanley A Wijaya, and Mengxue Ou. 2025. Designing Chatbots for Misinformation Correction: Examining the Roles of Chatbot Expertise and Anthropomorphism. *International Journal of Human–Computer Interaction* (2025), 1–15.

[21] Jui-Tse Hung, Christopher Cui, Diana M Popescu, Saurabh Chatterjee, and Thad Starner. 2024. Socratic Mind: Scalable Oral Assessment Powered By AI. In *Proceedings of the Eleventh ACM Conference on Learning@ Scale*. 340–345.

[22] Daniel Kahneman. 2011. Thinking, fast and slow. *Farrar, Straus and Giroux* (2011).

[23] Kai Kaspar, Sabine König, Jessika Schwandt, and Peter König. 2014. The experience of new sensorimotor contingencies by sensory augmentation. *Consciousness and cognition* 28 (2014), 47–63.

[24] Philip J Kellman and Sally Krasne. 2018. Accelerating expertise: Perceptual and adaptive learning technology in medical learning. *Medical teacher* 40, 8 (2018), 797–802.

[25] Gary A Klein et al. 1993. A recognition-primed decision (RPD) model of rapid decision making. (1993).

[26] Hao-Ping 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*. 1–22.

[27] Hause Lin, Gordon Pennycook, and David G Rand. 2023. Thinking more or thinking differently? Using drift-diffusion modeling to illuminate why accuracy prompts decrease misinformation sharing. *Cognition* 230 (2023), 105312.

[28] Mia Huong Nguyen, Kian Peen Yeo, Yasith Samaradivakara, and Suranga Nanayakkara. 2024. Catch that butterfly: A Multimodal Approach for Detecting and Simulating Gut Feelings. In *Adjunct Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. 1–3.

[29] Akshay Palimar Pai, Jayasankar Santhosh, and Shoya Ishimaru. 2022. Real-time feedback on reader's engagement and emotion estimated by eye-tracking and physiological sensing. In *Adjunct Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2022 ACM International Symposium on Wearable Computers*. 97–98.

[30] Fengjun Pan, Xiaobao Wu, Zongrui Li, and Luu Anh Tuan. 2024. Are LLMs good zero-shot fallacy classifiers?. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*. 14338–14364.

[31] Samir Passi, Shipi Dhanorkar, and Mihaela Vorvoreanu. 2024. *Appropriate reliance on Generative AI: Research synthesis*. Technical Report. Technical Report MSR-TR-2024-7. Microsoft. https://www. microsoft. com/en-us ....

[32] Gordon Pennycook, Tyrone D Cannon, and David G Rand. 2018. Prior exposure increases perceived accuracy of fake news. *Journal of experimental psychology: general* 147, 12 (2018), 1865.

[33] Gordon Pennycook, Ziv Epstein, Mohsen Mosleh, Antonio A Arechar, Dean Eckles, and David G Rand. 2021. Shifting attention to accuracy can reduce misinformation online. *Nature* 592, 7855 (2021), 590–595.

[34] Gordon Pennycook and David G Rand. 2022. Accuracy prompts are a replicable and generalizable approach for reducing the spread of misinformation. *Nature communications* 13, 1 (2022), 2333.

[35] Jayasankar Santhosh, Andreas Dengel, and Shoya Ishimaru. 2024. Gaze-Driven Adaptive Learning System with ChatGPT-Generated Summaries. *IEEE Access* 12 (2024), 173714–173733.

[36] Omar Shaikh, Shardul Sapkota, Shan Rizvi, Eric Horvitz, Joon Sung Park, Diyi Yang, and Michael S Bernstein. 2025. Creating general user models from computer use. In *Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology*. 1–23.

[37] Ashish Sharma, Kevin Rushton, Inna Lin, David Wadden, Khendra Lucas, Adam Miner, Theresa Nguyen, and Tim Althoff. 2023. Cognitive reframing of negative thoughts through human-language model interaction. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 9977–10000.

[38] Herbert A Simon. 1990. Bounded rationality. In *Utility and probability*. Springer, 15–18.

[39] Anjali Singh, Zhitong Guan, and Soo Young Rieh. 2025. Enhancing critical thinking in generative AI search with metacognitive prompts. *Proceedings of the Association for Information Science and Technology* 62, 1 (2025), 672–684.

[40] Mark Steyvers and Aakriti Kumar. 2024. Three challenges for AI-assisted decision-making. *Perspectives on Psychological Science* 19, 5 (2024), 722–734.

[41] Lev Tankelevitch, Elena L Glassman, Jessica He, Majeed Kazemitabaar, Aniket Kittur, Mina Lee, Srishti Palani, Advait Sarkar, Gonzalo Ramos, Yvonne Rogers, et al. 2025. Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–8.

[42] Lev Tankelevitch, Viktor Kewenig, Auste Simkute, Ava Elizabeth Scott, Advait Sarkar, Abigail Sellen, and Sean Rintel. 2024. The metacognitive demands and opportunities of generative AI. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–24.

[43] Richard H Thaler and Cass R Sunstein. 2009. *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

[44] Valerie A Thompson, Jamie A Prowse Turner, and Gordon Pennycook. 2011. Intuition, reason, and metacognition. *Cognitive psychology* 63, 3 (2011), 107–140.

[45] Cecilie S Traberg, Jon Roozenbeek, and Sander Van Der Linden. 2022. Psychological inoculation against misinformation: Current evidence and future directions. *The ANNALS of the American Academy of Political and Social Science* 700, 1 (2022), 136–151.

[46] Amos Tversky and Daniel Kahneman. 1981. The framing of decisions and the psychology of choice. *science* 211, 4481 (1981), 453–458.

[47] Tim Van Gelder. 2007. The rationale for Rationale™. *Law, probability and risk* 6, 1-4 (2007), 23–42.

[48] Helena Vasconcelos, Matthew Jörke, Madeleine Grunde-McLaughlin, Tobias Gerstenberg, Michael S Bernstein, and Ranjay Krishna. 2023. Explanations can

reduce overreliance on ai systems during decision-making. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–38.

[49] Stephen Waite, Arkadij Grigorian, Robert G Alexander, Stephen L Macknik, Marisa Carrasco, David J Heeger, and Susana Martinez-Conde. 2019. Analysis of perceptual expertise in radiology–current knowledge and a new perspective. *Frontiers in human neuroscience* 13 (2019), 213.

[50] Stephen Waite, Jinel Scott, Brian Gale, Travis Fuchs, Srinivas Kolla, and Deborah Reede. 2017. Interpretive error in radiology. *American Journal of Roentgenology* 208, 4 (2017), 739–749.

[51] Emily Wall, Arpit Narechania, Adam Coscia, Jamal Paden, and Alex Endert. 2021. Left, right, and gender: Exploring interaction traces to mitigate human biases. *IEEE transactions on visualization and computer graphics* 28, 1 (2021), 966–975.

[52] Markus Weinmann, Christoph Schneider, and Jan vom Brocke. 2016. Digital nudging. *Business & Information Systems Engineering* 58, 6 (2016), 433–436.

[53] Justine Zhang, James Pennebaker, Susan Dumais, and Eric Horvitz. 2020. Configuring audiences: A case study of email communication. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (2020), 1–26.

[54] Zheng Zhang, Jie Gao, Ranjodh Singh Dhaliwal, and Toby Jia-Jun Li. 2023. Visar: A human-ai argumentative writing assistant with visual programming and rapid draft prototyping. In *Proceedings of the 36th annual ACM symposium on user interface software and technology*. 1–30.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009