

From Consumption to Collaboration: Measuring Interaction Patterns to Augment Human Cognition in Open-Ended Tasks

Joshua Holstein
joshua.holstein@kit.edu
Karlsruhe Institute of Technology
Karlsruhe, Germany

Moritz Diener
moritz.diener@kit.edu
Karlsruhe Institute of Technology
Karlsruhe, Germany

Philipp Spitzer
philipp.spitzer@kit.edu
Karlsruhe Institute of Technology
Karlsruhe, Germany

Abstract

The rise of Generative AI, and Large Language Models (LLMs) in particular, is fundamentally changing cognitive processes in knowledge work, raising critical questions about their impact on human reasoning and problem-solving capabilities. As these AI systems become increasingly integrated into workflows, they offer unprecedented opportunities for augmenting human thinking while simultaneously risking cognitive erosion through passive consumption of generated answers. This tension is particularly pronounced in open-ended tasks, where effective solutions require deep contextualization and integration of domain knowledge. Unlike structured tasks with established metrics, measuring the quality of human-LLM interaction in such open-ended tasks poses significant challenges due to the absence of ground truth and the iterative nature of solution development. To address this, we present a framework that analyzes interaction patterns along two dimensions: cognitive activity mode (exploration vs. exploitation) and cognitive engagement mode (constructive vs. detrimental). This framework provides systematic measurements to evaluate when LLMs are effective tools for thought rather than substitutes for human cognition, advancing theoretical understanding and practical guidance for developing AI systems that protect and augment human cognitive capabilities.

CCS Concepts

• **Human-centered computing** → **Human computer interaction (HCI); HCI design and evaluation methods; Interactive systems and tools.**

Keywords

Human-LLM interaction, Human-AI collaboration, Tools for Thought, AI Reliance

1 Introduction

The emergence of Generative AI (GenAI) has fundamentally reshaped how humans approach cognitive tasks, from creative ideation to complex problem-solving [14, 19, 23]. Among these technologies, Large Language Models (LLMs) have become particularly influential due to their ability to engage in natural language dialogue and support a wide range of tasks [22]. While LLMs show remarkable potential for augmenting human thinking and reflection [19], they simultaneously risk cognitive erosion through uncritical reliance on their outputs [3, 10, 16, 21]. This risk is particularly pronounced as LLMs can generate seemingly high-quality outputs that may contain substantial flaws [15]: they can produce hallucinations—confidently stated but factually incorrect information—that can lead to misinformation and flawed mental models [20]. The fluent

articulation of these potentially problematic outputs makes them especially susceptible to uncritical acceptance by humans [23].

Current research proposes various design approaches to prevent cognitive erosion and promote meaningful engagement with LLMs. These approaches aim to facilitate reflection and critical thinking through different mechanisms, including specialized RAG-systems [14], learning environments [1], or hierarchical recommendations with prompt templates [26]. These design interventions are particularly relevant in open-ended tasks—such as strategic business choices, career development planning, and complex design decisions—where ambiguous goals and multiple solution paths require deep cognitive engagement to develop well-contextualized solutions [9]. Yet, unlike tasks with well-defined evaluation criteria, such as mathematical problem-solving [2] or code generation [1, 24], measuring the success in these open-ended tasks poses significant challenges.

While prior research has extensively studied the success of human-AI collaboration in traditional AI systems for classification [11, 13, 17, 18, 25] and regression tasks [7, 8], initial work has started to investigate human-LLM collaboration in structured domains like code generation or math [2, 24]. However, in contrast to “traditional” AI systems where an objective ground truth exists, dialogues between humans and LLMs in open-ended tasks create complex interaction patterns—such as iterative refinement of ideas, exploration of alternative solutions, and critical evaluation of LLM suggestions—that are inherently more difficult to evaluate [23]. These patterns emerge as humans engage with LLMs through multiple rounds of questioning, refinement, and synthesis, often involving both creative exploration and detailed elaboration of specific approaches. Yet, without an understanding of these interaction patterns, we lack the ability to design and validate LLM systems that genuinely enhance human cognitive processes rather than diminish them. This gap leads to our research question:

RQ: *How can we conceptualize and measure interaction patterns in human-LLM collaboration to distinguish cognitive augmentation from erosion in open-ended tasks?*

To answer our research question, we introduce a framework that conceptualizes human-AI collaboration patterns along two fundamental dimensions: **cognitive activity mode** (exploration vs. exploitation) and **cognitive engagement mode** (constructive vs. detrimental). By distinguishing between these dimensions, our framework provides a structured approach to understanding when LLMs effectively augment human cognition and when they risk passive engagement. This work contributes to both theoretical discourse and practical system design, offering measurable criteria for evaluating and promoting meaningful cognitive engagement in open-ended tasks.

2 Conceptualization

To analyze how humans engage with LLMs in open-ended tasks, we develop a framework building on established theories of exploration-exploitation trade-off [12] and cognitive engagement [5].

2.1 Cognitive Activity Mode: The Exploration-Exploitation Trade-off

Cognitive activity modes represent a fundamental tension in human learning and decision-making, characterized by the trade-off between exploration and exploitation [12]. This trade-off manifests as a continuous process of balancing exploring new possibilities with exploiting existing solutions. In open-ended tasks, where optimal solutions are unclear and context-dependent, this balance becomes particularly critical: Exploration enables discovering diverse solution paths and challenging assumptions, while exploitation allows deep development of promising approaches through iterative refinement.

Exploration describes a cognitive process of investigating unknown solutions, characterized by experimentation and discovery under uncertainty [6, 12]. In the context of human-LLM interaction for open-ended tasks, exploration manifests in activities where humans actively seek new perspectives, challenge assumptions, and generate novel ideas. Humans typically engage in divergent thinking [6], using the LLM as a tool for generating novel ideas. This might involve probing the system with open-ended questions, deliberately seeking alternative viewpoints, or using the LLM’s responses as prompts for further creative thinking. Exploration is particularly valuable in the early stages of open-ended tasks, where the problem space needs to be thoroughly understood and various solution paths considered.

Exploitation represents a cognitive process focused on utilizing and refining existing knowledge, characterized by the systematic improvement and application of identified solutions [6, 12]. In the context of human-LLM interaction for open-ended tasks, exploitation involves leveraging the system’s capabilities to refine and implement existing ideas or solutions. Humans in this mode typically engage with the LLM to deepen their understanding of specific solutions, seek detailed implementation guidance, or validate and improve already identified approaches. This might include requesting clarification of complex concepts, exploring the implications of chosen solutions, or refining specific aspects of an approach. Exploitation becomes particularly important when promising directions are identified and must be further developed to create practical, implementable solutions.

These two cognitive activities—exploration and exploitation—form a spectrum along which humans navigate when engaging with LLMs for problem-solving and the balance between them shapes the overall quality and depth of human-LLM collaboration. While exploration demands higher cognitive effort through active engagement with novel solutions, exploitation offers a potentially less demanding path of iteratively refining existing ones. When interacting with LLMs, this dynamics become particularly critical—the system’s ability to generate seemingly good and coherent responses might unconsciously incline humans toward exploitation [15], potentially limiting both their cognitive engagement and the depth of the developed solution.

2.2 Cognitive Engagement Mode: The Constructive-Detrimental Quality

The mode of cognitive engagement [5] represents our second dimension in understanding how humans interact with LLMs, particularly in open-ended tasks where deep thinking and critical reflection are essential to developing well-contextualized solutions.

Constructive Engagement describes cognitive processes where humans invest cognitive effort to maintain active agency and critical thinking [5]. In open-ended tasks, constructive engagement manifests when humans thoughtfully evaluate LLM outputs, integrate system suggestions with their own knowledge and experience, and strategically use the LLM to extend their thinking capabilities. For instance, when an LLM suggests a solution approach, constructively engaged humans actively contextualize this suggestion within their specific problem context, critically examining its assumptions and limitations. They might follow up with probing questions to understand the reasoning behind suggestions, deliberately challenge the system’s responses, or use them as starting points for developing more detailed solutions that incorporate their domain expertise. This type of engagement ensures that the LLM serves as a thought partner rather than a replacement for human judgment.

Detrimental Engagement represents patterns of shallow, mechanical interactions where humans’ cognitive processes are diminished rather than enhanced [5]. In open-ended tasks, this typically occurs when humans uncritically accept LLM outputs, overrely on the system for cognitive tasks, or engage in passive consumption of information without processing or integrating them into their knowledge and experience or tasks at hand. This form of engagement is particularly concerning as it may create an illusion of productivity while actually undermining the development of robust problem-solving capabilities and domain understanding, potentially even leading to a decrease in the quality of generated outputs.

2.3 Framework Integration: A Two-Dimensional View of Human-LLM Interaction

By integrating the cognitive mode with the quality of cognitive engagement, we develop a comprehensive framework for analyzing human-LLM interaction patterns in open-ended tasks (see Figure 1). This integration creates four distinct interaction patterns, each with unique implications for cognitive development and task outcomes.

		Cognitive Activity	
		Exploration	Exploitation
Cognitive Engagement	Constructive	Constructive Exploration	Constructive Exploitation
	Detrimental	Detrimental Exploration	Detrimental Exploitation

Figure 1: We distinguish between four types of human-LLM interaction across the two dimensions of cognitive activity mode and cognitive engagement mode.

- **Constructive Exploration** characterizes interactions where humans maintain high cognitive engagement while exploring new possibilities. Humans leverage the LLM to enhance creative thinking and problem-solving while maintaining critical judgment and synthesizing multiple perspectives with their own knowledge. This pattern is particularly valuable when the problem space needs a thorough understanding or existing approaches prove insufficient.
- **Constructive Exploitation** occurs when humans maintain thoughtful engagement while refining specific solutions. humans leverage the LLM to deepen their understanding while actively contextualizing and evaluating suggestions, engaging in detailed dialogue about specific aspects while ensuring alignment with their broader goals. This pattern becomes valuable when promising directions require further careful contextualization and development.
- **Detrimental Exploration** manifests as superficial brainstorming or unfocused idea generation without deep processing or critical evaluation. humans might accumulate numerous suggestions from the LLM without meaningfully engaging with any of them, leading to scattered thinking and shallow understanding. This pattern creates an illusion of productive exploration while failing to develop genuine insights or viable solutions.
- **Detrimental Exploitation** represents the most concerning pattern, characterized by excessive reliance on the LLM and minimal cognitive contribution from the user. In this mode, humans might uncritically implement the LLM’s suggestions, delegate core thinking tasks to the system, or focus solely on refining LLM-generated solutions without meaningful evaluation or integration with their knowledge. This pattern risks producing suboptimal solutions and eroding humans’ cognitive capabilities over time.

Through understanding these interaction patterns, we aim to design LLM systems that foster constructive engagement while preventing cognitive erosion

3 Framework Application

To move from theoretical conceptualization to practical application, we develop a systematic approach for analyzing human-LLM interactions in open-ended tasks. This section presents our measurement methodology and demonstrates its application through an illustrative example.

3.1 Measurement Approach

Building on our conceptualization of cognitive activity and cognitive engagement mode, we systematically measure human-LLM interaction patterns in open-ended tasks.

Qualitative Analysis of Cognitive Activity Mode. We first segment dialogues based on transitions between exploration and exploitation modes. A new segment begins when humans switch their cognitive mode, for instance, from exploring alternative solutions (e.g., seeking unconsidered perspectives) to exploiting a specific approach (e.g., detailed elaboration or implementation guidance). This segmentation enables us to analyze both the frequency and duration of different cognitive modes throughout the dialogue. By calculating the ratio of exploration segments to total segments, we

derive a score between 0 and 1 that quantifies the balance between exploration and exploitation, where 0 represents pure exploitation (all segments focused on refining specific solutions) and 1 represents pure exploration (all segments dedicated to discovering new possibilities). This measurement approach allows us to evaluate how different interface designs or interventions might influence the sequence and duration of exploration and exploitation phases.

Analysis of Cognitive Engagement Mode. Within each segment, we assess cognitive engagement mode through three key indicators: (1) the presence and quality of follow-up questions that build on previous responses (e.g., probing for deeper explanations or requesting clarification of assumptions), (2) explicit integration of personal context or experience (e.g., relating LLM suggestions to specific situations or domain knowledge), and (3) critical examination of LLM suggestions (e.g., challenging proposed solutions or identifying potential limitations). These indicators help classify each segment as either constructive or detrimental engagement. Constructive engagement is characterized by active processing and thoughtful integration of LLM outputs, while detrimental engagement manifests as passive consumption without meaningful evaluation or contextualization. Similar to our cognitive activity mode analysis, we calculate the ratio of constructively engaged segments to total segments, resulting in a score between 0 and 1, where 0 represents purely detrimental engagement and 1 represents consistently constructive engagement throughout the dialogue. This measurement approach enables us to identify patterns of cognitive engagement and evaluate how different interaction designs might promote more constructive engagement patterns.

Automated Analysis Through LLMs. To enable systematic analysis of larger datasets, we propose to leverage LLMs as analytical tools [4]. First, we prompt LLMs to identify segment transitions by recognizing shifts between exploration and exploitation modes. Second, we prompt them to assess cognitive engagement mode within each segment. This automation enables consistent analysis across diverse interaction logs while reducing the resource intensity of manual coding

3.2 Demonstration and Analysis

To illustrate how our framework can be applied to analyze human-LLM interactions, consider a dialogue where a user explores building an AI-driven market analysis tool (see Figure 2). The dialogue structure was designed to simulate authentic user interactions, incorporating both exploratory and exploitative approaches while demonstrating different patterns of engagement. To code the respective interactions, we prompted GPT-4o (version of November 20, 2024) to classify the interactions based on our definitions.

The interaction begins with constructive exploration as the user seeks to understand fundamental challenges (“What are the key challenges SMBs face when using market data?”) and follows up with contextual questions about industry-specific issues. This exploration phase maintains a high cognitive engagement through strategic questioning and building upon previous responses. A transition to constructive exploitation occurs when the user narrows focus to pricing strategy (“A freemium model sounds promising, but how can I design it?”). Here, the user actively engages with a specific solution while maintaining critical thinking through targeted

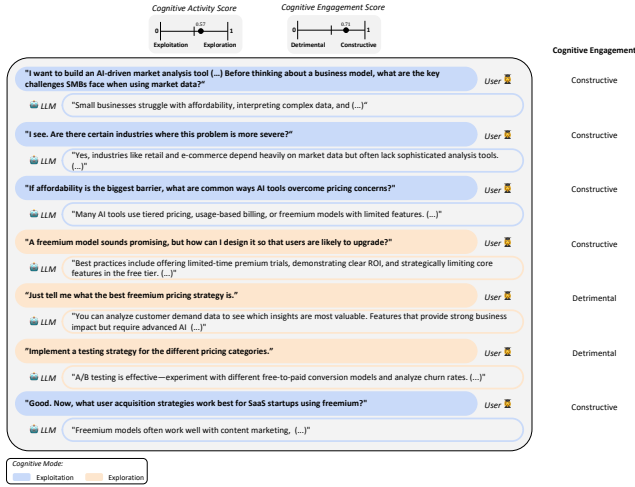


Figure 2: An exemplary chat history, where an LLM coded the user interaction along our two dimensions: cognitive activity and cognitive engagement mode.

questions. However, the interaction quality deteriorates into detrimental exploitation when the user shifts to passive consumption ("Just tell me what the best freemium pricing strategy is") and directive statements ("Implement a testing strategy"), showing minimal cognitive contribution. The dialogue concludes with a return to constructive exploration as the user broadens the scope to consider user acquisition strategies, demonstrating how interaction patterns can shift throughout a conversation. Using our measurement approach, we can quantify these patterns: the exploration-exploitation ratio indicates a balanced distribution with 4 out of 7 prompts being exploratory (57.14% exploration), while the cognitive engagement mode score (71.43% constructive) reveals predominantly meaningful cognitive engagement despite periods of passive consumption.

This example illustrates how our framework can identify and measure different interaction patterns, providing insights into when and how humans maintain constructive engagement versus falling into more passive interaction modes. The relatively balanced exploration-exploitation ratio (57.14%) suggests a healthy dynamic between discovering new possibilities and deepening specific solutions, while the high proportion of constructive engagement (71.43%) indicates that the user maintained meaningful cognitive contribution throughout most of the dialogue. However, the observed transition into detrimental exploitation during the pricing strategy discussion highlights how easily interaction quality can deteriorate when users seek quick solutions rather than maintaining thoughtful engagement. These measurements enable systematic evaluation of interaction quality and could inform the design of interventions that help users maintain constructive engagement patterns, particularly during transitions between exploration and exploitation phases.

4 Discussion

Our framework contributes to the theoretical understanding of human-LLM interaction as generative AI becomes increasingly

prevalent in knowledge work. By systematically distinguishing between cognitive activity modes and engagement patterns, we extend beyond traditional effectiveness metrics to examine the underlying mechanisms of human-LLM interaction. While LLMs demonstrate significant potential for augmenting human thinking [19], our framework reveals how their cognitive impact fundamentally depends on the quality of user engagement: the same system can either enhance or diminish human cognitive capabilities based on whether it facilitates active reasoning or enables passive consumption [21]. Through automated analysis, our framework enables real-time detection of engagement patterns and shifts toward detrimental interactions, opening avenues for timely interventions that ensure LLMs augment rather than diminish critical thinking.

The practical implications of our framework inform the design of AI systems that protect and augment human thinking. Our measurement approach enables systematic evaluation of how interface elements influence cognitive engagement, allowing for adaptive interventions when detrimental patterns are detected. Such interventions could prompt humans to articulate reasoning, consider alternative perspectives during exploration, or deepen reflection during exploitation phases. These design principles become increasingly important as AI systems integrate into professional and educational cognitive workflows.

Our framework's primary limitation is its need for empirical validation across different domains and contexts. Future studies should investigate how interface designs influence both cognitive activity modes and engagement patterns in open-ended tasks, providing evidence-based guidelines for LLM interfaces that enhance rather than diminish human cognitive capabilities. Finally, empirical validation must also verify the reliability and validity of our proposed measurement approach, particularly its ability to identify and distinguish between different interaction patterns reliably and whether these patterns translate in varying output quality.

5 Conclusion

Understanding how to design generative AI systems that genuinely augment human cognition while preventing cognitive erosion is a critical challenge of our time. Our framework offers a systematic approach to analyzing and measuring human-LLM interaction patterns in open-ended tasks, distinguishing between constructive and detrimental engagement forms across exploration and exploitation modes. By providing measures and design implications, this work contributes to the development of generative AI systems that enhance rather than diminish human cognitive capabilities. As LLMs become increasingly integrated into cognitive workflows, such frameworks become essential for ensuring these powerful tools augment human thinking rather than replace it.

References

- [1] John Chen, Xi Lu, Yuzhou Du, Michael Rejtig, Ruth Bagley, Mike Horn, and Uri Wilensky. 2024. Learning agent-based modeling with LLM companions: Experiences of novices and experts using ChatGPT & NetLogo chat. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–18.
- [2] Katherine M Collins, Albert Q Jiang, Simon Frieder, Lionel Wong, Miri Zilka, Umang Bhatt, Thomas Lukasiewicz, Yuhuai Wu, Joshua B Tenenbaum, William Hart, et al. 2024. Evaluating language models for mathematics through interactions. *Proceedings of the National Academy of Sciences* 121, 24 (2024), e2318124121.
- [3] Ian Drosos, Advait Sarkar, Neil Toronto, et al. 2025. "It makes you think": Provocations Help Restore Critical Thinking to AI-Assisted Knowledge Work.

- arXiv preprint arXiv:2501.17247* (2025).
- [4] Stefan Feuerriegel, Abdurahman Maarouf, Dominik Bär, Dominique Geissler, Jonas Schweisthal, Nicolas Pröllochs, Claire E Robertson, Steve Rathje, Jochen Hartmann, Saif M Mohammad, et al. 2025. Using natural language processing to analyse text data in behavioural science. *Nature Reviews Psychology* (2025), 1–16.
- [5] Barbara A Greene. 2015. Measuring cognitive engagement with self-report scales: Reflections from over 20 years of research. *Educational Psychologist* 50, 1 (2015), 14–30.
- [6] Tsutomu Harada. 2023. Exploring the effects of risk-taking, exploitation, and exploration on divergent thinking under group dynamics. *Frontiers in psychology* 13 (2023), 1063525.
- [7] Patrick Hemmer, Max Schemmer, Niklas K"uhl, Michael V"ossing, and Gerhard Satzger. 2022. On the effect of information asymmetry in human-AI teams. In *CHI Conference on Human Factors in Computing Systems (CHI '22) Workshop on Human-Centered Explainable AI* (New Orleans, LA, USA). ACM.
- [8] Patricia K Kahr, Gerrit Rooks, Martijn C Willemsen, and Chris CP Snijders. 2024. Understanding Trust and Reliance Development in AI Advice: Assessing Model Accuracy, Model Explanations, and Experiences from Previous Interactions. *ACM Transactions on Interactive Intelligent Systems* 14, 4 (2024), 1–30.
- [9] Emily R Lai. 2011. Critical thinking: A literature review. *Pearson's Research Reports* 6, 1 (2011), 40–41.
- [10] 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. (2025).
- [11] Zhuoran Lu and Ming Yin. 2021. Human reliance on machine learning models when performance feedback is limited: Heuristics and risks. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–16.
- [12] James G March. 1991. Exploration and exploitation in organizational learning. *Organization science* 2, 1 (1991), 71–87.
- [13] Katelyn Morrison, Philipp Spitzer, Violet Turri, Michelle Feng, Niklas K"uhl, and Adam Perer. 2024. The impact of imperfect XAI on human-AI decision-making. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (2024), 1–39.
- [14] Marissa Radensky, Simra Shahid, Raymond Fok, Pao Siangliulue, Tom Hope, and Daniel S Weld. 2024. Scideator: Human-LLM Scientific Idea Generation Grounded in Research-Paper Facet Recombination. *arXiv preprint arXiv:2409.14634* (2024).
- [15] Daniel M Russell, Laura Koesten, Aniket Kittur, Nitesh Goyal, and Michael Xieyang Liu. 2024. Sensemaking: What is it today?. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–5.
- [16] Advait Sarkar, Neil Toronto, Ian Drosos, Christian Poelitz, et al. 2024. When Copilot Becomes Autopilot: Generative AI's Critical Risk to Knowledge Work and a Critical Solution. *arXiv preprint arXiv:2412.15030* (2024).
- [17] Max Schemmer, Niklas Kuehl, Carina Benz, Andrea Bartos, and Gerhard Satzger. 2023. Appropriate Reliance on AI Advice: Conceptualization and the Effect of Explanations. In *Proceedings of the 28th International Conference on Intelligent User Interfaces* (Sydney, NSW, Australia) (*IUI '23*). Association for Computing Machinery, New York, NY, USA, 410–422.
- [18] Jakob Schoeffer, Maria De-Arteaga, and Niklas Kuehl. 2024. Explanations, Fairness, and Appropriate Reliance in Human-AI Decision-Making. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–18.
- [19] Auste Simkute, Viktor Kewenig, Abigail Sellen, Sean Rintel, and Lev Tankelevitch. 2025. The New Calculator? Practices, Norms, and Implications of Generative AI in Higher Education. *arXiv preprint arXiv:2501.08864* (2025).
- [20] Philipp Spitzer, Joshua Holstein, Katelyn Morrison, Kenneth Holstein, Gerhard Satzger, and Niklas K"uhl. 2024. Don't be Fooled: The Misinformation Effect of Explanations in Human-AI Collaboration. *arXiv preprint arXiv:2409.12809* (2024).
- [21] Matthias Stadler, Maria Bannert, and Michael Sailer. 2024. Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior* 160 (2024), 108386.
- [22] Hari Subramonyam, Roy Pea, Christopher Pondoc, Maneesh Agrawala, and Colleen Seifert. 2024. Bridging the Gulf of Envisioning: Cognitive Challenges in Prompt Based Interactions with LLMs. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–19.
- [23] 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* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 680, 24 pages.
- [24] Helena Vasconcelos, Gagan Bansal, Adam Fourney, Q. Vera Liao, and Jennifer Wortman Vaughan. 2024. Generation Probabilities Are Not Enough: Uncertainty Highlighting in AI Code Completions. *ACM Trans. Comput.-Hum. Interact.* (2024).
- [25] Xinru Wang and Ming Yin. 2023. Watch out for updates: Understanding the effects of model explanation updates in ai-assisted decision making. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19.
- [26] 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.