

# AI as Extraherics: Fostering Higher-order Thinking Skills in Human-AI Interaction

Koji Yatani

The University of Tokyo  
Bunkyo-ku, Tokyo, Japan  
koji@iis-lab.org

Zefan Sramek

The University of Tokyo  
Bunkyo-ku, Tokyo, Japan  
zefanS@iis-lab.org

Chi-Lan Yang

The University of Tokyo  
Bunkyo-ku, Tokyo, Japan  
chilan.yang@iii.u-tokyo.ac.jp

## ABSTRACT

As artificial intelligence (AI) technologies, including generative AI, continue to evolve, concerns have arisen about over-reliance on AI, which may lead to human deskilling and diminished cognitive engagement. Over-reliance on AI can also lead users to accept information given by AI without performing critical examinations, causing negative consequences, such as misleading users with hallucinated contents. This paper introduces *extraheric AI*, a human-AI interaction conceptual framework that fosters users' higher-order thinking skills, such as creativity, critical thinking, and problem-solving, during task completion. Unlike existing human-AI interaction designs, which replace or augment human cognition, extraheric AI fosters cognitive engagement by posing questions or providing alternative perspectives to users, rather than direct answers. We discuss interaction strategies, evaluation methods aligned with cognitive load theory and Bloom's taxonomy, and future research directions to ensure that human cognitive skills remain a crucial element in AI-integrated environments, promoting a balanced partnership between humans and AI.

## CCS CONCEPTS

- Human-centered computing → HCI theory, concepts and models; HCI design and evaluation methods; Interaction design theory, concepts and paradigms; • Computing methodologies → Philosophical/theoretical foundations of artificial intelligence.

## KEYWORDS

Human-AI interaction, higher-order thinking skills, germane cognitive load, Bloom's taxonomy

## 1 INTRODUCTION

Recent advances in artificial intelligence (AI), including generative AI, have shown strong potential to support human tasks, reduce workloads, and augment capabilities, but concerns have arisen about the over-use or over-reliance on AI technology [16, 22]. Such reliance on AI for cognitive tasks can lead to deskilling, where individuals lose opportunities for cognitive skill maintenance and development [143]. Over-reliance on AI for information-seeking may also sway users toward particular viewpoints or opinions presented by AI, as seen in writing and design explorations [40, 134], and can further exacerbate issues of misinformation and disinformation when users blindly trust erroneous or hallucinated AI outputs. This dependence may also diminish perceived ownership, sense of challenge, productivity, and accomplishment [77]. A fundamental issue

underlying these negative consequences is the focus of current human-AI interaction research on supporting human tasks by replacing or augmenting human cognitive abilities. Such AI design may enhance task efficiency but deprive users of opportunities for cognitive engagement and growth. With generative AI becoming increasingly capable of outperforming humans in many tasks, users may be more likely to trust AI without skepticism.

To address these challenges, researchers have explored redesigning human-AI interactions to promote cognitive engagement. Tankelevitch et al. discuss the potential of generative AI for expanding users' metacognitive capabilities [128]. Danny et al. showed that asking questions about a user's argument, rather than providing additional explanations, can stimulate users' critical thinking through "*human-AI co-reasoning*" [29]. These discussions and projects suggest a strong potential for human-AI interaction research to stimulate users' creativity, critical thinking, and problem-solving skills among the seven core 21st-century skills outlined by van Laar et al. [133].

In this paper, we introduce a novel conceptual framework for human-AI interaction: *extraheric AI*. We define "*extraherics*" as a mechanism that fosters users' higher-order thinking skills during the course of task completion. *Extraheric* is based on the Latin word "extrahō" (to draw forth or pull out), and we use this term to suggest that AI can draw forth people's higher-order thinking skills and thus promote their cognitive potential. Rather than replacing or augmenting human cognitive abilities, extraheric AI encourages users to engage in higher-order thinking during task completion. For instance, in writing, extraheric AI might prompt users to reflect on specific content or visualize how others have approached similar topics, rather than directly performing revisions or replacement. This process encourages users to examine, select, and synthesize information, creating implicit learning opportunities that foster higher-order thinking skills.

As users increasingly rely on intelligent systems, extraheric AI aligns with N. Katherine Hayles' vision of a positive *posthuman* future, where humans are not "*hopelessly compromised*" by machines but instead engage in a collaborative cognitive environment with them [59]. In this model, thinking is a shared process between human and nonhuman actors, ensuring that human higher-order thinking skills remain vital and are actively fostered [59].

This paper provides an overview of extraheric AI, detailing its interaction strategies, evaluation methods, design considerations, and future research directions. The contributions include:

- Defining the novel human-AI interaction conceptual framework, *extraheric AI*, and comparing it with other human-AI interaction designs using cognitive load theory,
- Identifying interaction strategies for extraheric AI based on recent HCI literature,
- Outlining evaluation methods for assessing users' cognitive, attitudinal, and behavioral change when using extraheric AI, and
- Exploring design considerations and proposing future research directions for extraheric AI.

## 2 RELATED WORK

### 2.1 Higher-Order Thinking Skills

Lewis and Smith define higher-order thinking as a cognitive process that “occurs when a person takes new information and information stored in memory and interrelates and/or rearranges and extends this information to achieve a purpose or find possible answers in perplexing situations” [85]. This definition is purposefully broad, and, as the authors note, higher-order thinking can be used for a variety of tasks, including: “deciding what to believe; deciding what to do; creating a new idea, a new object, or an artistic expression; making a prediction; and solving a nonroutine problem” [85]. *Critical thinking skills* and *creative thinking skills* are both characteristic features of higher-order thinking skills [28].

Beyond the acquisition of knowledge and skills, the development of the ability to think critically has become a key educational goal in the last half-century [96], and it is increasingly essential in the face of the development of advanced AI tools that are changing the ways we gather and interact with information. Broadly speaking, *critical thinking* describes “thinking that is purposeful, reasoned, and goal-directed—the kind of thinking involved in solving problems, formulating inferences, calculating likelihoods, and making decisions when the thinker is using skills that are thoughtful and effective for the particular context and type of thinking task” [57], as well as “the appropriate use of reflective skepticism within the problem area under consideration” [96].

Although often associated with artistic creation, *creative thinking* is also broadly defined as thinking applicable to numerous domains. Torrance defines creative thinking as “the process of sensing gaps or disturbing, missing elements; forming ideas or hypotheses concerning them; testing these hypotheses; and communicating the results, possibly modifying and retesting the hypotheses” [132]. Similarly to critical thinking, creative thinking has been of increasing interest to educators over the last several decades. Meador notes that “the ability to think creatively is essential in life for many reasons, including solving problems, producing meaningful and satisfying ideas and products, and developing works in art forms” [97].

Both critical and creative thinking are skills that can be developed by learning about and practicing reasoning, analysis, planning, and questioning [57, 79, 98]. Aside from such knowledge and practice, however, Halpern also emphasizes the importance of the *attitudes* of a critical thinker, characterized by willingness to plan, flexibility, persistence, willingness to self-correct, being mindful, and consensus-seeking [57]. Similarly, Michalko emphasizes the importance of cultivating a creative

attitude, which includes the *belief* that one is creative [98]. Encouraging and developing such attitudes among users is also a key goal for extraheric AI because these attitudes allow users to transfer the skills they have developed to other task domains.

Metacognition is another concept related to and yet distinct from higher-order thinking. Metacognition refers to “thinking about thinking,” being aware of one’s cognitive processes and regulating them to achieve specific goals [79]. Metacognitive skills include *metacognitive knowledge* (awareness of one’s own learning processes, strategies, and strengths or weaknesses) [47] and *metacognitive regulation* (organizing, monitoring, and assessing one’s own learning activities) [15]. Metacognitive skills may thus help different aspects of higher-order thinking; for example, when analyzing a complex problem, metacognitive skills help a person plan how to approach the problem, monitor their comprehension, and adjust their strategies as appropriate. In this sense, metacognition focuses on awareness and regulation of thinking processes, enabling effective use of higher-order thinking skills like analysis, evaluation, and creation. Tankelevitch et al. [128] have recently explored the role of metacognition in the use of generative AI, arguing that generative AI places substantial metacognitive demands on the user. They propose that generative AI systems can aid users by including metacognitive supports as well as reducing metacognitive demands through design, by “offload[ing] metacognitive processing from the user to the system” [128]. Extraheric AI proposes a different perspective on the division of cognitive labor between users and AI. It aims at stimulating cognitive activities to foster higher-order thinking skills through the use of AI, instead of offloading such activities.

### 2.2 Technology Support for Higher-Order Thinking Skills

Higher-order thinking skills in general have received little attention from the HCI community to date, though some scholars from other research fields have examined the use of technology to foster higher-order thinking skills [61]. That being said, there exists a long history of scholarship on the ways that technology can be harnessed to promote *critical thinking*, *creative thinking*, and support educational goals in general. Early research included discussions about the roles of computers in schools [118], and proposals, such as Jonassen’s concept of *Mindtools* [68].

HCI researchers have long been interested in creativity support tools [48], and have developed numerous methods for evaluating their impacts on creative thinking [113]. Recent research in HCI and related fields has also explored various techniques for promoting critical thinking in a variety of application domains, including web search [146], online collaboration [126], educational exhibitions [81], online learning [63], digital media literacy [108], data sensing [80], engineering research [6], and misinformation mitigation [14, 37]. With the power and flexibility of AI, extraheric AI has the potential to accelerate this research direction and play a substantial role in the development and promotion of higher-order thinking skills.

Interactive Tutoring Systems (ITS) is a domain related to interaction and interface designs for higher-order thinking skill

development<sup>1</sup>. ITS can foster students' higher-order thinking skills by providing personalized learning experiences, such as adaptive feedback and problem-solving exercises. ITS can also offer support for thinking activities by incorporating expert thinking models and metacognitive strategies [49]. While ITS is specifically designed to fit into educational contexts, extraheric AI can be integrated into users' existing non-educational activities. Extraheric AI thus has the potential to enhance users' higher-order thinking skills in a broader set of application domains. In an example of building abilities to counter mis/dis-information on the Internet, ITS can offer personalized teaching content based on students' preferences and behavior on online content. Extraheric AI can complement this by encouraging the execution of higher-order thinking skills when users read online content in a practical setting.

Effortful user interfaces are another interface design relevant to extraheric AI. Effortful user interfaces purposefully infuse cognitive load into interfaces in order to encourage users' learning of computing systems [27]. Effortful user interfaces typically introduce additional burdens during task completion, which serves as a motivator for users to learn. For example, Grossman et al. revealed that disabling the activation of a command in a pull-down menu can lead to accelerated learning of keyboard shortcuts [54]. Similar concepts and merits of introducing purposeful workload to interface designs have been discussed within HCI [115]. These interface designs take an approach to infuse deliberate intrinsic cognitive load to encourage users' learning and behavior. In a related way, extraheric AI instead aims to increase germane cognitive load to stimulate cognitive activities that can foster users' higher-order thinking skills, as described in Section 3.

### 3 EXTRAHERIC AI

Human-AI interaction systems are typically designed to directly support human tasks, such as by taking on subtasks, accelerating processes, or reducing input effort. In his book *Human-Centered AI*, Ben Schneiderman offers the following categorization for tools serving human needs [124]:

- *Orthotics*: Systems that enhance performance in specific tasks (e.g., auto-completion, FlashFill in Excel, and Copilot for coding).
- *Prosthetics*: Systems that replace missing capabilities (e.g., real-time captioning, and visual information verbalization).
- *Exoskeletons*: Systems that expand human capacities related to specific tasks (e.g., language translation, and information search assistants).

Many existing human-AI interaction systems fit one or more of these categories, although their classification can vary depending on context and user capabilities. For example, a language translation application serves as prosthetics for users with no background in a language, but acts as exoskeleton for those with some proficiency.

Unlike these categories, extraheric AI focuses on fostering users' higher-order thinking skills through interaction. Extraheric AI agents help users explore different information and perspectives

<sup>1</sup>ITS is also known as Intelligent Computer-Aided Instructions (ICAI) in the domain of educational systems using AI, but the term of ITS has been adopted most widely [104].

while completing tasks, instead of providing direct support. Extraheric AI thus introduces a new research direction for human-AI interaction: increasing germane cognitive load. According to John Sweller's cognitive load theory, there are three types of cognitive load: *intrinsic*, *extraneous*, and *germane* [127]. Intrinsic load is the mental effort required by the inherent complexity of the material. Extraneous load is the additional mental effort caused by how information is presented, which does not directly aid learning (e.g., poorly-designed interfaces). Germane load is the mental effort involved in constructing and automating *schemas* (cognitive frameworks that help organize information efficiently). Increasing germane load is thus desirable because it fosters the development and retention of higher-order thinking skills.

Extraheric AI aims to increase germane cognitive load while the other AI types (orthotics, prosthetics, and exoskeleton) primarily focus on reducing intrinsic or extraneous cognitive load. This also suggests that extraheric AI can coexist with these supporting AI designs. For instance, a writing support AI system might automatically correct minor errors in a document while also prompting users with questions to deepen their thinking on the content. While exploring extraheric AI designs that allow the coexistence with other AI types is beyond the scope of this paper, it presents a valuable opportunity for future research in human-AI interaction.

Extraheric AI is particularly applicable to domains where higher-order thinking skills determine the execution and outcome quality of a given task. While Section 4 illustrates applications demonstrated in the existing literature, examples of these applications would be: intellectual activities that involve opinion formation, decision making, and creative thinking (e.g., brainstorming, market analysis, and design proposals), resolving complex problems (e.g., programming and debugging a large project), and behavioral changes supported by perspective and perception changes (e.g., personal informatics for a person's physical and mental health). For example, as Robins et al. note, "*novice programmers must learn to develop models of the problem domain, the notional machine, and the desired program, and also develop tracking and debugging skills so as to model and correct their programs*" [117]. This combination of both conceptual and practical skills illustrates the essential nature of developing higher-order thinking skills for such complex task domains. Future extraheric AI development should actively explore new application domains because higher-order thinking skills can be critical in a broad range of intellectual activities.

### 4 EXTRAHERIC AI INTERACTION STRATEGIES

While Section 3 clarifies the definition of extraheric AI, it is necessary to understand how researchers and developers can design and implement it. As this paper emphasizes the HCI perspective of extraheric AI rather than machine learning models and architecture for it, this section discusses interaction strategies relevant to extraheric AI. To explore these strategies in a comprehensive and bottom-up manner, we conducted an extensive literature survey, as outlined below.

We systematically collected full-paper publications at CHI 2023 and 2024 using the keywords: “AI,” “artificial intelligence,” “LLM,” and “large language models” in the ACM Digital Library. We manually performed the initial screening based on the following criteria.

- The paper demonstrates a prototype system utilizing AI technology (not limited to generative AI) or emulating AI (e.g., the wizard-of-Oz method), instead of executing purely observational studies. This criterion was essential as we aimed to focus on actual demonstrations of extraheretic AI.
- The paper targets cognitive activities involving users’ higher-order thinking. This criterion was chosen because such activities are the main focus of extraheretic AI.
- The paper demonstrates extraheretic characteristics in its prototype design (i.e., promoting users’ critical thinking). This criterion was included because it is the essential component of extraheretic AI. The work rather categorized as pure supporting AI systems (i.e., orthotics, prosthetics, and exoskeleton) was excluded accordingly.

We then conducted a detailed review of the papers that satisfied all the three criteria to determine interaction strategies their prototypes as well as the application domains. To ensure objectivity and reliability, two of the authors independently reviewed these papers and annotated the interaction strategies they employed. In cases of disagreement, we re-reviewed these papers to reach a consensus. We also consolidated annotations that were closely related to enhance clarity and accuracy.

Through this process, we categorized 50 papers and identified eight distinct interaction strategies, which are discussed in detail below. Table 1 illustrates our categorization against the interaction strategies and application domains.

#### 4.1 Suggesting & Recommending

Suggesting and recommending is an extraheretic AI interaction strategy that involves proposing ideas, viewpoints, solutions, or actions to the user, without necessarily detailing the rationales behind them. With this strategy, users’ cognitive engagement comes in the form of evaluating and deciding whether or not to incorporate the AI’s suggestions or recommendations into their thinking or tasks. For example, in the context of news article reading, extraheretic AI may recommend related articles with similar or different perspectives to encourage the user to explore multiple viewpoints. In the context of technical tasks like software development, extraheretic AI may suggest multiple implementations of a particular method and allow the user to choose the one they determine to be most appropriate. An example system using this strategy is CoArgue, developed by Liu et al. [88], where the user can overview claims from different stances to construct their own arguments. In all such cases, it is critical that the AI makes *multiple* suggestions to allow the user to evaluate and choose among them.

#### 4.2 Explaining

Explaining is a strategy in which extraheretic AI offers explanations of information related to the task which the user currently engages in. Unlike suggestions or recommendations, this strategy

emphasizes providing details on the ‘why’ and ‘how’ of a particular piece of information. In the context of news article reading for opinion formation, extraheretic AI with this strategy may visualize additional background or contextual explanations about a particular component of the article the user is currently reading. In this manner, extraheretic AI allows the user to confirm their understanding and situate the article correctly. ClarifAI, developed by Zavolokina et al. [83], demonstrates such an explaining strategy to alert users to potential propaganda content. In the context of software development, extraheretic AI may provide explanations of methods or API calls to allow the user to better understand the code’s function. It is important that extraheretic AI offers explanations that allow the user to deepen their understanding of the task at hand. It thus should aim to provide additional context or information rather than step-by-step instructions that the user may blindly follow.

#### 4.3 Nudging

A nudge is an approach to subtly encourage or influence behavior through indirect suggestions and reinforcements without preventing alternative choices [131]. Although the concept originates in the field of behavioral economics, HCI research has extensively explored its applications and confirmed effects on decision-making and behavioral change [10, 121]. Extraheretic AI using this strategy may indirectly show information that can persuade the user in particular directions while still offering them freedom to choose. In the context of news article reading, extraheretic AI may visualize a conceptual space of relevant articles in a side view, nudging the user to explore other opinions. In the context of software development, a system could populate a dynamic list of relevant packages or libraries that may be of use to the user, but without specifically recommending any of them. Wu et al. use this strategy in designing interventions for problematic smartphone use [144], utilizing large language models for constructing personalized and dynamic persuasive messages with different persuasion strategies. It is important that systems using this strategy make it easy for the user to find information that helps them explore different directions and perspectives, without presenting their options as being necessary or exhaustive.

#### 4.4 Debating & Discussing

In this mode, users debate or discuss a given topic and exchange their thoughts and opinions with AI agents. In the context of news article reading, extraheretic AI using this strategy could offer an online discussion thread where the user may discuss their thoughts with AI agents holding various different opinions. In the context of software development, the user may engage in paired programming with an AI peer, and discuss the use of different libraries, code structures, or algorithms. Zhang et al. develop a system where the user can discuss with multiple AI agents exhibiting different viewpoints in order to deepen their understanding of different opinions and overcome filter bubble effects [153]. When using this strategy, it is important that debates and discussions focus on presenting different perspectives and ideas rather than simply disagreeing with or asking the user to justify their opinion.

	Suggesting & recommending	Explaining	Nudging	Debating & Discussing	Questioning	Scaffolding	Simulating	Demonstrating
<b>Information seeking, information search</b>	[82], [122]	[70], [83]				[83]		
<b>Dis/mis-information checking</b>		[24], [64], [84], [150]		[129], [153]				
<b>Textual content creation (e.g., writing)</b>	[26], [35], [51], [92], [109]					[35], [100], [114]		
<b>Discussions, argument construction</b>	[71], [88]			[52]	[29]			
<b>Personal informatics, behavioral changes</b>		[125]	[144]			[11]		
<b>Programming</b>	[20], [73]	[36], [73], [147]			[66]	[36]		
<b>Decision-making, sense-making, data analysis</b>	[21], [55], [111], [123]	[89], [111], [138], [145], [149], [152]						
<b>Education</b>	[151]	[19]			[19], [23], [60], [151]		[81]	[91]
<b>Design/idea/prototype explorations</b>	[25], [31], [76], [90], [137]	[76], [139]		[154]				

Table 1: Categorization of existing literature related to extraheric AI published as full papers at CHI 2023 and 2024 by interaction strategy and application domain. Note that some papers demonstrate multiple interaction strategies.

#### 4.5 Questioning

In this mode, extraheric AI asks questions about particular parts of what the user currently engages in. Such questioning is not supposed to validate the correctness of opinions and perspectives, but rather stimulate users' cognitive activities to expand their thoughts or consider different perspectives. In the context of news article reading, extraheric AI using this strategy may ask questions about a particular portion of the content, such as "How do you think people in other countries perceive this news? What consequence could occur in their countries?". In the context of software development, extraheric AI may ask the user to explain how a particular code block functions, or why they chose to implement an algorithm in the way they did. Danry et al. demonstrate the positive effects of this strategy on the critical thinking task of evaluating the logical soundness of statements that can create social division [29]. This strategy encompasses interactive environments based on a well-known effective pedagogical approach called *learning by teaching* [42]. Through users' active engagement with questions from AI agents, this process effectively stimulates users' higher-order thinking.

#### 4.6 Scaffolding

Scaffolding is a learning approach where teachers offer temporary customized support to help students learn new concepts and skills, and gradually remove this help as students become more capable on their own [142]. Scaffolding allows a learner to effectively complete tasks beyond their current ability level, but also, as Wood et al. note, results in "*development of task competence by the learner at a pace that would far outstrip [their] unassisted efforts*" [142]. Extraheric AI can serve as a scaffold for users by taking on part of a task and allowing them to focus only on particular portions at a time. In the context of software development, extraheric AI using this strategy may help the user focus on program structure by allowing them to write pseudo-code or use visual programming methods before later translating these into functional code. Lee et al. demonstrate DAPIE [83], where the AI agent offers step-by-step explanations while encouraging children to actively interact with it and assess their understanding. However, as Wood et al. note, "*comprehension of the solution must precede production. That is to say, the learner must be able to recognize a solution to a particular class of problems before [they are themselves] able to produce the steps leading to it without assistance*" [142]. As a result, it is critical

that extraheric AI using this strategy focuses on developing the user's fundamental understanding of a task rather than simply allowing them to offload task decomposition.

#### 4.7 Simulating

In this mode, extraheric AI simulates a circumstance where the user experiences a situation from a standpoint other than their own or develops skills that would be difficult to otherwise practice. For example, AI agents could simulate audience members of different opinions and perspectives, allowing users to practice public speaking and responding to audience questions. Extraheric AI could be tuned to different levels of aggressiveness to develop resilience and abilities for handling different types of audiences. Simulations can also allow users to experience situations from a different standpoint. For example, the user could take the role of an interviewer tasked with interviewing an AI agent playing the role of a job candidate. By asking a variety of questions and observing the agent's responses, the user can think about how they may answer such questions as interviewees in actual job interviews. The user will also have the opportunity to empathize with their interviewers, potentially giving them valuable insight into how to best communicate their ideas. A series of art exhibitions by Lee et al. allowed students to immersively experience facets of AI (e.g., artwork generated using their facial data), stimulating critical thinking through an informal learning experience [81]. As with other strategies, it is important that such simulations be designed to present a variety of viewpoints to encourage users to consider different perspectives and think critically about their own positionality. This strategy could be particularly valuable for helping users understand their own and others' implicit biases.

#### 4.8 Demonstrating

Demonstrating is a strategy where users simply observe the behavior or interaction of AI agents and learn implicitly through these observations. In this case, there is no direct information flow from extraheric AI to users. Users thus would have the largest freedom in how they interpret the behavior or interaction of AI agents and internalize take-aways through vicarious learning [116]. In the context of news article reading for opinion formation, extraheric AI using this strategy may take the role of a peer, demonstrating their reading process and sharing opinions. The user can review these demonstrations and construct their own opinions by integrating what they have observed with their own reading. Liu et al. create a classmate AI agent in a virtual reality classroom that plays the role of an active student [91]. Students in the same virtual classroom observe its behavior, which can stimulate their active class engagement. In addition to employing such a role model, extraheric AI using this strategy may include multiple AI agents to offer the demonstrations of diverse perspective or approaches to a task or topic.

### 5 EXTRAHERIC AI EVALUATION APPROACHES

User evaluation is a critical component in HCI research. Qualitative evaluation approaches, such as journaling, questionnaires using open-ended questions, and interviews,

can be widely applicable to extraheric AI research to examine its user experience and effects. In this section, we present aspects of user experience and abilities that may convey the effects of extraheric AI, and discuss other possible approaches associated with germane cognitive load and Bloom's taxonomy [4, 13]. We, however, note that extraheric AI research can be highly dependent on its use contexts and applications, and thus the objectives of extraheric AI may be diverse. As such, this section intends to lay a foundation to help researchers and developers explore how they can combine these approaches, as well as qualitative methods, to obtain a holistic perspective on user experience and the effects of their extraheric AI systems. We encourage the HCI research community to explore evaluation approaches together to acquire deep understanding of how extraheric AI influences users. Table 2 summarizes our proposed areas of evaluation and related metrics.

#### 5.1 Evaluations on Germane Load

As discussed in Section 3, extraheric AI can be interpreted as a mechanism to increase users' germane load. It is thus important to assess how users' germane load would change with and without the presence of extraheric AI for completing a given task. Although cognitive theory and educational psychology research have explored different approaches to evaluate the three types of cognitive load [33, 105], HCI research has not adopted a standardized way of measuring cognitive load yet. NASA-TLX [58] is a well-established metric for evaluating subjective workload and commonly used in HCI, but it is not considered associated with cognitive load theory. To address this issue, Gerjets et al. revised NASA-TLX by introducing *task demands*, *navigational demands*, and *efforts*, which are explicitly connected with intrinsic, extraneous, and germane cognitive load, respectively [50]. These statements are: *"how much mental and physical activity was required to accomplish the learning task (e.g., thinking, deciding, calculating, remembering, looking, searching etc.)"* for task demands, *"how much effort the participant had to invest to navigate the learning environment"* for navigational demands, and *"how hard the participant had to work to understand the contents of the learning environment"* for efforts. As the NASA-TLX has already been widely adopted in the field of HCI, we suggest that this modified version may be immediately adapted to HCI research to quantitatively measure users' perceived cognitive load with extraheric AI. However, in addition to the revised Gerjets et al.'s revised NASA-TLX, there exist other scales for measuring cognitive workloads [105]. We encourage the HCI research community to discuss and explore how HCI research can adapt these scales to better assess users' germane load.

#### 5.2 Evaluations on Higher-order Thinking Skills

Several theoretical frameworks exist in the fields of educational psychology and cognitive psychology to assess individuals' higher-order thinking skills. For instance, Bloom's taxonomy [13], the SOLO Taxonomy [12], Fink's Taxonomy of Significant Learning [46], and the Paul-Elder Critical Thinking Framework [107] can be potential frameworks for evaluating higher-order thinking skills. Here, we apply Bloom's taxonomy

Evaluation Area	Potential Evaluation Method	Reference
Germane Load	Revised NASA-TLX	[50]
Cognitive Activity (lower level)	Knowledge and comprehension tests	[38, 69]
Cognitive Activity (inter. level)	Reflections; Concept maps	[41, 140]
Cognitive Activity (higher level)	Performance-based assessment; Prototyping	[1]
Sense of Agency	<i>Sense of Agency Scale</i>	[130]
Self-Efficacy	<i>New General Self-Efficacy Scale</i>	[18]
Task Motivation	<i>Motivation Source Inventory</i>	[67]
AI Use Motivation	Measures of likeability and trust of agents	[112]
Attribution of Credit and Blame	Measures of credit and blame	[74]

Table 2: Areas of evaluation for extratheric AI and their potential metrics.

as an example to guide us on how to evaluate individuals' higher-order thinking skills with extratheric AI.

Bloom's taxonomy is a well-recognized framework that hierarchically organizes the stages of cognitive activities, and has already been widely adopted by HCI and computer science education research [93]. The initial taxonomy was proposed by Bloom et al. in 1958 [13], and includes six stages of cognitive development: *knowledge* (remembering facts, terms, and basic concepts); *comprehension* (understanding information); *application* (using knowledge in new situations); *analysis* (breaking information into components to understand its structure); *synthesis* (combining different information and knowledge together to form a new solution); and *evaluation* (making judgments about information and ideas). Anderson and Krathwohl revised it to emphasize creativity, application, and higher-order thinking skills [4]: *remember* (equivalent to *knowledge* in the original Bloom's taxonomy), *understand* (*comprehension*), *apply* (*application*), *analyze* (*analysis*), *evaluate* (*evaluation*), and *create* (*synthesis*; it is placed as the highest level in the revised taxonomy). While the order of the stages is slightly different, both taxonomies cover the same set of cognitive activities around higher-order thinking skills. In this section, we discuss possible evaluation approaches following the categorization of the six stages of these taxonomies into three consolidated levels, following Jones et al.'s categorization [69].

**5.2.1 Assessing Users' Knowledge and Comprehension Ability (Lower Level).** *Knowledge and comprehension* lie in the most basic levels in Bloom's taxonomy and its revised version. They involve correct recall of facts, terms, and basic concepts, and the understanding of information by interpreting, summarizing, explaining, or translating it into one's own words. Typical methods to evaluate individuals' understanding and knowledge involve assessing their recall of facts, terms, and basic concepts, as well as subject matter tests utilizing true or false questions, multiple-choice questions, fill-in-the-blank questions, summarizing, paraphrasing, and providing examples. For instance, Jones et al. employed knowledge tests with questions phrased with "define," "list," "state," "identify," and "label" to gauge students' knowledge capability [69]. Rubrics developed based on Bloom's taxonomy are another common approach for computer science educators to assess students' cognitive development [38].

We note that extratheric AI research should not overemphasize supporting or assessing participants' lower-level cognitive

abilities. Knowledge is a fundamental tool for higher-order thinking but is not generally recognized as a higher-order thinking skill. Over-weighting changes in *knowledge and comprehension* abilities thus may not capture the true effects and benefits of extratheric AI.

**5.2.2 Assessing Users' Application and Analysis Ability (Intermediate Level).** *Application and analysis* involve using knowledge and concepts in new situations, solving problems by implementing learned procedures or techniques, and being able to break down information into smaller components to understand its structure, relationships, and patterns.

*Application and analysis* abilities can be assessed using a number of methods, including evaluating users' performance in a problem-solving context, and analyzing their reflections through reflective journals and workshops [41, 110]. Accuracy of concept maps, a visual tool that organizes and represents relationships between concepts, can also indicate the degree of individuals' *application and analysis* abilities [103]. There exist various different approaches to evaluate concept maps [95, 140], and choice of these evaluation methods may depend on research objectives and contexts. Future researcher is encouraged to explore how extratheric AI may utilize these and other methods.

**5.2.3 Assessing Users' Evaluation and Creation Ability (Higher Level).** *Evaluation and creation* involve making judgments or decisions based on criteria and standards by assessing the value or quality of ideas, methods, or materials, as well as combining elements in new ways to form a coherent outcome, generating novel ideas, or creating original products.

Performance-based assessment, such as analyzing the process and outcome of users' discussion, debate, and decision-making tasks is appropriate for assessing their higher-level abilities [1]. By analyzing arguments provided in a discussion or written report, researchers can gauge users' higher-level cognitive abilities with regard to identifying the strengths and weaknesses of arguments and counterarguments based on evidence and logic. Inviting users to outline alternative solutions or methods to complete the given task is also effective in assessing their *evaluation and creation* abilities. We suggest that the HCI research community, which has a long history of engaging participants through interactive and participatory interface prototyping approaches, can uniquely contribute to assessing *evaluation and creation* abilities by

exploiting these existing approaches and their associated body of knowledge.

### 5.3 Attitudinal and Behavioral Metrics

In addition to the direct examination of higher-order thinking skills discussed in the previous section, changes in users' attitudes and behavior over time can also be an important indicator of the effects of extraheretic AI. For instance, users may exhibit a more open attitude toward diverse opinions or an increase in self-efficacy after they have developed stronger higher-order thinking skills. There exist established scales for quantitatively examining these changes. We point out key attitudinal and behavioral aspects that are relevant to the use of extraheretic AI, and invite HCI researchers to broaden the exploration of methods to assess them.

**5.3.1 Sense of Agency.** Sense of agency refers to the feeling or perception of having control over one's own actions, thoughts, and effects. We expect that the development of higher-order thinking skills may lead users to develop a stronger sense of control over their tasks and the information presented to them. Such perception would accordingly be reflected in their sense of agency. Accordingly, Xiao et al. included a sense of agency measurement in their extraheretic AI research to support users to achieve better-informed consent [145]. Beyond the effects of higher-order thinking skill development, we also hypothesize that extraheretic AI may have different impacts on users' sense of agency compared to orthotics, prosthetics, and exoskeleton. Systems that negatively impact human agency appear to have a higher chance of contributing to feelings of dehumanization [143], reduced meaningful human interaction [136], and reduced trust in AI [87]. This suggests that measuring and understanding users' sense of agency is critical for the success of future human-AI interaction paradigms.

Although measurement of agency remains a challenging task with considerable disagreement with regard to its methodology [17], there exist several well-established scales. The Sense of Agency Scale developed by Tapal et al. [130] is one of the most commonly-used instruments in HCI and human-AI interaction research [40, 145]. Exploring appropriate adaption of existing sense of agency scales to extraheretic AI research is an important open research direction.

**5.3.2 Self-efficacy.** Self-efficacy refers to one's perception of the ability to effectively utilize tools and environments to achieve desired outcomes [8]. While sense of agency is about the perception of control over given systems and information, self-efficacy is about confidence in the ability to complete tasks. In extraheretic AI research, the development of higher-order thinking skills may contribute to a deeper understanding of the process of completing tasks, potentially resulting in improved self-efficacy. As noted above, one of the major concerns with over-reliance on AI is the risk of deskilling and reduced cognitive engagement. These concerns relate directly to users' self-efficacy, and assessing changes in self-efficacy is therefore critical for understanding the effectiveness of extraheretic AI systems. Similarly to sense of agency, there exist several scales for self-efficacy. The New General

Self-Efficacy Scale developed and further validated by Chen et al. is one of the established scales [18], though it has not yet been widely adopted by HCI research. The HCI community is encouraged to further explore how existing self-efficacy scales can be adapted to extraheretic AI research.

**5.3.3 User Motivations and Willingness.** As extraheretic AI may increase germane cognitive load, users may experience a cognitive burden, particularly when first using a system. While this may potentially decrease their motivation, extraheretic AI may also strengthen their motivation when they feel the sense of growth from the cultivation of their higher-order thinking skills. Understanding how users' motivational changes over time can help researchers better adapt the design of extraheretic AI systems. Several theories and measurements have been widely used in HCI research to investigate users' motivation for using new technology. These include Self-Determination Theory (SDT) [34], the Technology Acceptance Model (TAM) [30, 62], the Motivation Source Inventory [67], or the Intrinsic Motivation Inventory [94]. As the fluctuation of users' motivation when working with extraheretic AI remains unexplored, we encourage researchers to consider this aspect when assessing the effectiveness of extraheretic AI.

Another important evaluation metric related to motivation is user willingness to use extraheretic AI. As extraheretic AI may not offer direct support for a given task, users may exhibit a lower willingness to continue to use. However, a decrease in users' willingness may also occur when they have become proficient in the task without the use of AI support, which can indicate a move towards positive disengagement. Understanding how user willingness changes alongside developments in their higher-order thinking skills is critical for understanding the effects on extraheretic AI on users' attitudinal and behavioral changes.

**5.3.4 Responsibility and Ownership Attribution with Extraheretic AI.** The magnitude of effort and autonomy users perceive in interaction with extraheretic AI may influence their perceived attribution of responsibility and ownership of final outcomes to AI agents. Past studies have shown that users attributed more blame and responsibility when encountering failed collaborative outcomes to social actors that they conceptualized as more intelligent [7, 102]. Accordingly, Kadoma et al. used the perceived ownership between users and AI in co-writing process as one of the evaluation metrics [71]. As extraheretic AI aims at stimulating users' high-order thinking skills during task completion, it is essential to know how users assign responsibility or credit for the collaborative outcome to their extraheretic AI agents, especially when both a successful and failed outcome can occur. A common approach to assess responsibility attribution is to ask users to indicate the amount of blame and credit each stakeholder should receive for a specific task [45, 74]. A question asking the extent to which users feel that the outcome is attributed to them can also measure perceived ownership [40, 71]. As users may engage in many activities that require higher-order thinking skills, understanding how users attribute the collective responsibility and credit over time is critical for extraheretic AI research.

## 6 EXTRAHERIC AI DESIGN CONSIDERATIONS

Existing literature has identified and validated guidelines and design considerations for human-AI interaction research [3, 86, 148]. While these guidelines and design considerations also apply to extraheric AI, its unique characteristics introduce additional design considerations researchers and developers should take into account.

### 6.1 Explicitly Considering The Social Roles of Extraheric AI Agents

As users are expected to collaboratively explore various perspectives and information with extraheric AI, AI agents may play different social roles similar to what is seen in human-human communication and collaboration. Köbis et al. described four social roles of AI [78]: role model, advisor, partner, and delegatee<sup>2</sup>, and discuss how different roles may cause different possible risks of negatively influencing users' ethical behavior. With extraheric AI, it is crucial to explicitly consider agents' expected or perceived social roles, as users may interpret the same output differently depending on the roles these AI agents have. There may also be other possible social roles in the context of extraheric AI; for example, a competitor role where an AI agent competes against users may contribute to users' active thinking. Kim et al. examined the effects of a social bot playing a role of a depressed peer, which displays depressive symptoms to urge users to offer support and encouragement [75]. The interactions with the bot helped their study participants reframe their own negative experiences. Future extraheric AI research should consider both positive and negative effects amplified by these social roles on the effectiveness and acceptance of the developed systems more broadly.

Uncovering the design space of extraheric AI social roles is therefore an important research agenda. The effect of such social roles on user outcomes needs further examination through empirical studies. For example, students may benefit more from interacting with extraheric AI agents playing the role of their peers rather than teachers, as Liu et al. demonstrated that an AI agent that simulates an active student peer in a virtual classroom can promote students' class participation [91]. Future research on extraheric AI should consider not only the interaction strategies by which AI can promote higher-order thinking skills, but also how the social roles of AI agents can enhance or degrade this process.

### 6.2 Generating Diversified Outputs

To promote higher-order thinking skill development, the output of extraheric AI should be designed to encourage users' diverse interpretations instead of constraining users to particular directions or perspectives. In particular, suggestions from AI have already been found to produce idea fixation and discourage divergent thinking [135]. Presenting multiple diversified outputs is thus critical, particularly when extraheric AI employs the *suggesting & recommending* interaction strategy. This design consideration is also crucial for the *debating & discussing*

<sup>2</sup>Their original article [78] used the term of "delegate" to represent AI use where people outsource their tasks to AI. We decided to use "delegatee" to clarify the relationship between users and AI.

interaction strategy. Only employing an extraheric AI agent with a very similar stance to users may create undesirable echo-chamber effects, whereas multiple AI agents with diverse opinions have already been shown to help mitigate filter bubble effects [153]. It is essential that future extraheric AI research explores how a system can produce diversified responses and perspectives, and present them in ways that push users to develop and use their higher-order thinking skills. However, when designing for diversity, it is also critical that designers and developers be aware of their positionality and potential blindspots [120], as well as continue to stay abreast of challenges with AI bias and other ethical issues [65].

### 6.3 Maintaining Non-judgmental Attitudes and Behavior

Extraheric AI should encourage users' intellectual exploration. And in many cases, there are no clear right or wrong directions for users to explore. It is thus critical that extraheric AI systems remain non-judgmental, and in particular, avoid providing simplified evaluations (e.g., numerical scores) of users' outcomes. Non-judgmental behavior is a fundamental principle of motivational interviewing, a communication approach designed to facilitate and engage intrinsic motivation within users to promote positive behavioral change [99]. Extraheric AI should embrace how users integrate different perspectives into their tasks, including their decisions to not make use of AI output. Non-judgmental behavior thus can be beneficial for promoting honesty and trust toward extraheric AI.

Researchers and developers may want to perform some quantification of users' outcomes to evaluate their extraheric AI systems. However, we maintain that such quantification should only be used for evaluating system performance (e.g., effects of different interaction strategies or roles on users' higher-order thinking skills), rather than for scoring users or otherwise directly rating their 'performance'. In general, extraherics is not primarily about directly accelerating task completion, but rather about expanding users' cognitive capabilities.

### 6.4 Aligning with Users' Workflows and Contexts for Task Completion

Integration with existing workflows has been identified as the key to adoption of a variety of technologies, ranging from creativity support tools [106] to generative AI [119]. While extraheric AI aims at fostering users' higher-order thinking skills, it should also support the tasks users wish to complete. It thus must take into account how it can provide users with opportunities to expand their higher-order thinking skills while minimizing interference with their task workflows. This constitutes an important difference from learning systems including ITS, where the primary focus lies in skill training and development. Understanding users' existing workflows and contexts is critical for designing appropriate extraheric AI that can maintain a balance between supporting users' task completion and cognitive development. Formative studies can help to establish such understandings while iterative design explorations or co-design processes are effective

approaches for ensuring that extraheretic AI systems match users' workflows and contexts.

## 6.5 Embracing User Disengagement from Extraheretic AI

As users engage with extraheretic AI over time, they may develop sufficient higher-order thinking skills, and as Lewis and Smith note, may become sufficiently adept at the task at hand such as to no longer require the use of these skills to complete it [85]. For example, in the context of writing, users may initially obtain help from extraheretic AI to engage in critical thinking by being questioned about the content of a particular sentence or paragraph, but may become able to do so independently later. In such cases, users may discontinue their use of extraheretic AI as they feel stronger self-efficacy in the given task. While adaptively changing the interaction strategies or social roles of extraheretic AI can be a valid approach to encourage continued engagement, disengagement can also be considered indicative of sufficient higher-order thinking skill development. As a result, future research should not overemphasize the frequency or occurrences of extraheretic AI use in its evaluations. Section 5 discusses evaluation approaches for extraheretic AI, and advocates for a holistic perspective for measuring the different aspects of how people use extraheretic AI systems. While the frequency of extraheretic AI use is one potential metric to understand user behavior, researchers and developers should carefully interpret and thoroughly discuss its contextual significance, instead of simply observing its magnitude. We therefore encourage applying mixed methods using diverse data sources to understand users' motivations for both engagement and disengagement.

## 7 RESEARCH OPPORTUNITIES

The introduction of extraheretic AI has the potential to contribute to and extend existing understanding and theory about human-AI interaction, and the use of extraheretic AI may change the way people internalize knowledge and develop higher-order thinking skills. Here, we outline several open directions for studying extraheretic AI, from interaction and interface design, and cognitive load theory, to new evaluation metrics, and interpersonal and social implications.

### 7.1 Technology to Present Multiple Responses and Perspectives

A first set of research questions concerns how to design technology that could present multiple responses and perspectives from extraheretic AI. In addition to explorations of interface designs that would not overwhelm users or interfere with their current tasks, diversifying responses and perspectives from extraheretic AI is also an important technical research challenge. Information retrieval research has introduced the concept of search result diversification [2], which refers to the process of delivering a set of relevant searches that also cover a broad range of aspects or perspectives related to the search query. This concept is particularly important in scenarios where the search query may be ambiguous, or where different users may have different information needs. While the goal of search result diversification

lies in maximizing the likelihood that users can find what they are looking for, output diversification of extraheretic AI aims at increasing the likelihood of users engaging in higher-order thinking activities. Unlike search results, which themselves are not directly controllable by search engines, researchers and developers have the freedom to control the output from extraheretic AI agents by configuring their characters, personality, social roles, and behavior. Employing multiple extraheretic AI agents can be beneficial with respect to generating multiple perspectives and opinions, contributing to output diversification as a whole. Future research is encouraged to explore creative approaches for extraheretic AI output diversification by exploring various AI agent designs.

### 7.2 HCI Perspectives on Germane Cognitive Load

In Section 3, we discussed the difference between extraheretic AI and other supportive AI through the lens of the cognitive load theory. Germane cognitive load is a relatively new concept, and remains controversial within the field of cognitive science. Kalyuga, for example, critiques the three-way model of cognitive load and proposes that it is more of an aspect of intrinsic load [72]. Similarly, de Jong points out the blurred nature of the boundaries between these types of cognitive load and discusses the challenging nature of separating them [32]. We argue that the HCI research community can contribute to further understanding these types of cognitive load by employing actual interactive systems and performing evaluations (e.g., comparative studies between extraheretic and non-extraheretic AI systems). However, we also echo de Jong's discussion on the challenges of evaluating solely one of these cognitive load types, and note that careful consideration is needed for future research. We see an opportunity for the HCI research community to contribute to the discussion of cognitive load theory in the context of extraheretic AI by incorporating holistic evaluation metrics that assess both human and system performance. We believe that extraheretic AI research can create a new bridge between computer science and cognitive science, calling for new interdisciplinary collaboration.

### 7.3 Explorations of Alternative Evaluation Approaches for Extraheretic AI

In Section 5, we discussed possible evaluation approaches based on cognitive load theory and Bloom's taxonomy, as well as attitudinal and behavioral scales. As extraheretic AI is a new concept, we have borrowed existing evaluation metrics from the field of education research. However, we also suggest that there are unique opportunities for the HCI research community to develop alternative frameworks and evaluation approaches to capture changes in users' high-order thinking skills over time when using extraheretic AI.

As the development of higher-order thinking skills may need long-term effort, future extraheretic AI research will need to devise tracking methods for cognitive development and evaluate their validity. Periodical evaluations employing approaches discussed in Section 5 may be a possible design, but researchers and developers should also consider the effort required of participants

for repeatedly completing such evaluation tasks. Observing the evolution of reflection journals or concept maps can be an interesting approach to gauging the development of higher-order thinking skills. For example, researchers and developers may quantify the correctness and level of detail of concept maps developed by participants and observe how these metrics transition over time [39]. The HCI research community has extensive experience in in-the-wild, long-term user evaluations, and we argue that this knowledge can contribute to establishing evaluation methods for longer-term development of higher-order thinking skills in a realistic setting.

#### 7.4 Effects Caused by Different Social Roles of Extraheretic AI

As discussed in Section 6.1, the social roles of extraheretic AI agents may have impacts on user perception and behavior of these agents, and, accordingly, have effects on encouraging cognitive activities related to higher-order thinking skills. As demonstrated through studies of interpersonal interaction, the perceived social roles of interlocutors can influence behavior [155], power dynamics [44], and communication patterns [141]. However, it is still unclear how interacting with extraheretic AI designed with varying social roles may influence these aspects and further users' higher-order thinking skills. Examining the effects of different social roles of extraheretic AI is therefore a key open research direction.

Multiple AI agents acting in different social roles have the potential to enable social learning for users as well as produce diversified outputs, as suggested in Section 6.2. Social learning involves acquiring knowledge, skills, or behavior by observing, imitating, and interacting with others in a social context [9]. Individuals may gain higher-order thinking skills through collaborative interactions, discussions, and the exchange of diverse perspectives with other people. However, whether and how a similar social learning effect can be attained when users are interacting with multiple extraheretic AI agents in a simulated social context is largely unknown. Therefore, future research efforts can investigate the effects caused by interacting with multiple extraheretic AI agents that display different social roles.

#### 7.5 Ethical Considerations and Guidelines for Extraheretic AI

Recent years have seen substantial investigations and reflections on the ethical issues of machine learning [101], large language models [65], and other forms of machine ethics [5]. However, as Elliott et al. note, these numerous guidelines can be practically counterproductive and require harmonization to be usable [43]. The development of extraheretic AI and its application to various domains can and must consider such ethical issues and guidelines in its implementation, but, as a new proposal, it will likely introduce new ethical considerations as well. For example, extraheretic AI may introduce competing incentives for users and developers as it potentially leads to user mastery and thus disengagement from AI. Many existing AI products rely on a business model of continuous use, potentially creating a trade-off between users' learning outcomes and providers' financial outcomes in the case of extraheretic AI. In the past, the development

of many AI technologies has preceded the investigation of ethical concerns, which has been largely relegated to fields outside computer science. We urge researchers exploring extraheretic AI to proactively integrate ethical considerations into their research programs from the start, in order to develop and integrate design and ethical best practices simultaneously.

A set of related open questions for extraheretic AI design concerns user populations that may need extra caution and care. For instance, in light of existing childhood development research, extra caution may be necessary in the application of extraheretic AI to educational contexts involving children or adolescents. Research on children's free play and its impacts on mental and physical development suggests that undirected play undertaken freely by children is essential for the development of physical and cognitive abilities and mental well-being [53]. During group and collaborative activities with others, children develop numerous competencies beyond the task at hand, which allow them to "*tolerate bruises, handle their emotions, read other children's emotions, take turns, resolve conflicts, and play fair*" [56]. However, these educationally-beneficial interactions may not occur effectively through interaction with AI agents, which can be easily ignored, shut off, or reset without any consequences, potentially counteracting the educational benefit that extraheretic AI would otherwise provide. Similarly, other specific user populations and application domains may require extra caution to design ethical extraheretic AI, and future research is encouraged to explore domain-specific considerations and guidelines.

### 8 CONCLUSION

This paper presents *extraheretic AI*, a novel human-AI interaction conceptual framework aimed at mitigating the risks of over-reliance on AI, which can lead to human deskillling and reduced cognitive engagement. Unlike traditional AI designs that replace or augment human cognition, extraheretic AI fosters higher-order thinking skills by engaging users through questions and alternative perspectives rather than providing direct answers and support to given tasks. This paper illustrates HCI research components of extraheretic AI: interaction strategies, evaluation approaches, and design considerations. As discussed above, extraheretic AI opens up several research opportunities on which the HCI research community can take a strong initiative. We hope that this work will serve as a catalyst for deeper discourse and further research on human-AI interaction that prioritizes a balanced partnership between humans and interactive intelligent systems.

### ACKNOWLEDGMENTS

We appreciate all the support we received in preparing this paper. In particular, Barnaby Ralph participated in our exploration of a new term for our conceptual framework and eventually coined *extraherics*. He also contributed to our discussions of extraherics by bringing in posthuman literature. Kakeru Miyazaki assisted us with conducting a literature survey to identify extraheretic AI strategies and application domains. We greatly appreciate their

dedicated help. This project was partly supported by JST PRESTO (Grant Number: JPMJPR23IB).

## REFERENCES

- [1] Yousef Abosalem. 2016. Assessment techniques and students' higher-order thinking skills. *International Journal of Secondary Education* 4, 1 (2016), 1–11.
- [2] Rakesh Agrawal, Sreenivas Gollapudi, Alan Halverson, and Samuel Jeong. 2009. Diversifying search results. In *Proceedings of the Second ACM International Conference on Web Search and Data Mining* (Barcelona, Spain) (*WSDM '09*). Association for Computing Machinery, New York, NY, USA, 5–14. <https://doi.org/10.1145/1498759.1498766>
- [3] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collison, Jina Suh, Shamsi Iqbal, Paul N. Bennett, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz. 2019. Guidelines for Human-AI Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) (*CHI '19*). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3290605.3300233>
- [4] Lorin W. Anderson and David R. Krathwohl. 2001. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Longman, New York.
- [5] Michael Anderson and Susan Leigh Anderson. 2011. *Machine Ethics*. Cambridge University Press.
- [6] Yousef A Asiri, David E Millard, and Mark J Weal. 2021. Assessing the impact of engagement and real-time feedback in a mobile behavior change intervention for supporting critical thinking in engineering research projects. *IEEE Transactions on Learning Technologies* 14, 4 (2021), 445–459.
- [7] Edmond Awad, Sydney Levine, Max Kleiman-Weiner, Sohan Dsouza, Joshua B Tenenbaum, Azim Shariff, Jean-François Bonnefon, and Iyad Rahwan. 2020. Drivers are blamed more than their automated cars when both make mistakes. *Nature human behaviour* 4, 2 (2020), 134–143.
- [8] Albert Bandura. 1977. Self-Efficacy: Toward a Unifying Theory of Behavioral Change. *Psychological Review* 84, 2 (1977), 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>
- [9] Albert Bandura. 1977. *Social Learning Theory*. Prentice-Hall.
- [10] Kristoffer Bergram, Marija Djokovic, Valéry Bezençon, and Adrian Holzer. 2022. The digital landscape of nudging: A systematic literature review of empirical research on digital nudges. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–16.
- [11] Ananya Bhattacharjee, Yuchen Zeng, Sarah Yi Xu, Dana Kulzhabayeva, Miny Ma, Rachel Kornfield, Syed Ishitaque Ahmed, Alex Mariakakis, Mary P Czerwinski, Anastasia Kuzminykh, Michael Liut, and Joseph Jay Williams. 2024. Understanding the Role of Large Language Models in Personalizing and Scaffolding Strategies to Combat Academic Procrastination. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 15, 18 pages. <https://doi.org/10.1145/3613904.3642081>
- [12] John B. Biggs and Kevin F. Collis. 1982. *Evaluating the Quality of Learning: The SOLO Taxonomy (Structure of the Observed Learning Outcome)*. Academic Press, New York.
- [13] 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*. David McKay Company, New York.
- [14] Nattapat Boonprakong, Benjamin Tag, and Tilman Dingler. 2023. Designing Technologies to Support Critical Thinking in an Age of Misinformation. *IEEE Pervasive Computing* 22, 3 (2023), 8–17.
- [15] Ann L. Brown. 1978. Knowing When, Where and How to Remember: A Problem of Metacognition. In *Advances in Instructional Psychology*, Robert Glaser (Ed.). Vol. 1. Lawrence Erlbaum Associates, Hillsdale, NJ, 77–165.
- [16] 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.
- [17] Federica Cavazzoni, Alec Fiorini, and Guido Veronese. 2022. How do we assess how agentive we are? A literature review of existing instruments to evaluate and measure individuals' agency. *Social Indicators Research* 159, 3 (2022), 1125–1153.
- [18] Gilad Chen, Stanley M. Gully, and Dov Eden. 2001. Validation of a New General Self-Efficacy Scale. *Organizational Research Methods* 4, 1 (2001), 62–83. <https://doi.org/10.1177/109442810141004>
- [19] Liuqing Chen, Zhaojun Jiang, Duowei Xia, Zebin Cai, Lingyun Sun, Peter Childs, and Haoyu Zuo. 2024. BIDTrainer: An LLMs-driven Education Tool for Enhancing the Understanding and Reasoning in Bio-inspired Design. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 676, 20 pages. <https://doi.org/10.1145/3613904.3642887>
- [20] Liuqing Chen, Shuhong Xiao, Yunlong Chen, Yaxuan Song, Ruoyu Wu, and Lingyun Sun. 2024. ChatScratch: An AI-Augmented System Toward Autonomous Visual Programming Learning for Children Aged 6–12. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 649, 19 pages. <https://doi.org/10.1145/3613904.3642229>
- [21] Qing Chen, Wei Shuai, Jiayao Zhang, Zhida Sun, and Nan Cao. 2024. Beyond Numbers: Creating Analogies to Enhance Data Comprehension and Communication with Generative AI. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 377, 14 pages. <https://doi.org/10.1145/3613904.3642480>
- [22] Valerie Chen, Q Vera Liao, Jennifer Wortman Vaughan, and Gagan Bansal. 2023. Understanding the role of human intuition on reliance in human-AI decision-making with explanations. *Proceedings of the ACM on Human-computer Interaction* 7, CSCW2 (2023), 1–32.
- [23] Alan Y. Cheng, Meng Guo, Melissa Ran, Arpit Ranasaria, Arjun Sharma, Anthony Xie, Khuyen N. Le, Bala Vinaithirthan, Shihe (Tracy) Luan, David Thomas Henry Wright, Andrea Cuadra, Roy Pea, and James A. Landay. 2024. Scientific and Fantastical: Creating Immersive, Culturally Relevant Learning Experiences with Augmented Reality and Large Language Models. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 275, 23 pages. <https://doi.org/10.1145/3613904.3642041>
- [24] Furui Cheng, Vilém Zouhar, Simran Arora, Mrinmaya Sachan, Hendrik Strobelt, and Mennatallah El-Assady. 2024. RELIC: Investigating Large Language Model Responses using Self-Consistency. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 647, 18 pages. <https://doi.org/10.1145/3613904.3641904>
- [25] DaEun Choi, Sumin Hong, Jeongeon Park, John Joon Young Chung, and Juho Kim. 2024. CreativeConnect: Supporting Reference Recombination for Graphic Design Ideation with Generative AI. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 1055, 25 pages. <https://doi.org/10.1145/3613904.3642794>
- [26] Seulgi Choi, Hyewon Lee, Yoonjoo Lee, and Juho Kim. 2024. VIVID: Human-AI Collaborative Authoring of Vicarious Dialogues from Lecture Videos. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 277, 26 pages. <https://doi.org/10.1145/3613904.3642867>
- [27] Andy Cockburn, Per Ola Kristensson, Jason Alexander, and Shumin Zhai. 2007. Hard lessons: effort-inducing interfaces benefit spatial learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (*CHI '07*). Association for Computing Machinery, New York, NY, USA, 1571–1580. <https://doi.org/10.1145/1240624.1240863>
- [28] Wendy Conklin. 2011. *Higher-order thinking skills to develop 21st century learners*. Teacher Created Materials.
- [29] Valdemar Danny, Pat Pataranaporn, 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. <https://doi.org/10.1145/3544548.3580672>
- [30] Fred D Davis, RP Bagozzi, and PR Warshaw. 1989. Technology acceptance model. *J Manag Sci* 35, 8 (1989), 982–1003.
- [31] Richard Lee Davis, Thimo Wambgsanss, Wei Jiang, Kevin Gonyop Kim, Tanja Käser, and Pierre Dillenbourg. 2024. Fashioning Creative Expertise with Generative AI: Graphical Interfaces for Design Space Exploration Better Support Ideation Than Text Prompts. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 167, 26 pages. <https://doi.org/10.1145/3613904.3642908>
- [32] Ton de Jong. 2010. Cognitive Load Theory, Educational Research, and Instructional Design: Some Food for Thought. *Instructional Science* 38, 2 (2010), 105–134.
- [33] Nicolas Debeau and Cécile van de Leemput. 2014. What does germane load mean? An empirical contribution to the cognitive load theory. *Frontiers in Psychology* 5 (Oct 2014), 1099. <https://doi.org/10.3389/fpsyg.2014.01099> Prepublished online 2014 Mar 6.
- [34] Edward L Deci and Richard M Ryan. 2000. The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry* 11, 4 (2000), 227–268.
- [35] Paramveer S. Dhillon, Somayeh Molaei, Jiaqi Li, Maximilian Golub, Shaochun Zheng, and Lionel Peter Robert. 2024. Shaping Human-AI Collaboration:

- Varied Scaffolding Levels in Co-writing with Language Models. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 1044, 18 pages. <https://doi.org/10.1145/3613904.3642134>
- [36] Griffin Dietz, Nadir Tamer, Carina Ly, Jimmy K Le, and James A. Landay. 2023. Visual StoryCoder: A Multimodal Programming Environment for Children's Creation of Stories. 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 96, 16 pages. <https://doi.org/10.1145/3544548.3580981>
- [37] Tilman Dingler, Benjamin Tag, Philipp Lorenz-Spreen, Andrew W Vargo, Simon Knight, and Stephan Lewandowsky. 2021. Workshop on technologies to support critical thinking in an age of misinformation. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–5.
- [38] Mohsen Dorodchi, Nasrin Dehbozorgi, and Tonya K Frevert. 2017. I wish I could rank my exam's challenge level!: An algorithm of Bloom's taxonomy in teaching CS1. In *2017 IEEE Frontiers in Education Conference (FIE)*. IEEE, 1–5.
- [39] Emma EH Doyle, Sara E Harrison, Stephen R Hill, Matt Williams, Douglas Paton, and Ann Bostrom. 2022. Eliciting mental models of science and risk for disaster communication: A scoping review of methodologies. *International Journal of Disaster Risk Reduction* 77 (2022), 103084.
- [40] Fiona Draxler, Anna Werner, Florian Lehmann, Matthias Hoppe, Albrecht Schmidt, Daniel Buschek, and Robin Welsch. 2024. The AI Ghostwriter Effect: When Users do not Perceive Ownership of AI-Generated Text but Self-Declare as Authors. *ACM Trans. Comput.-Hum. Interact.* 31, 2, Article 25 (feb 2024), 40 pages. <https://doi.org/10.1145/3637875>
- [41] Heidi Dunfee, Aaron Rindfussch, Maryanne Driscoll, John Hollman, and Margaret M. Plack. 2008. Assessing Reflection and Higher-order Thinking in the Clinical Setting Using Electronic Discussion Threads. *Journal of Physical Therapy Education* 22, 2 (Fall 2008), 60–67.
- [42] David Duran and Keith Topping. 2017. *Learning by teaching: Evidence-based strategies to enhance learning in the classroom*. Routledge.
- [43] Karen Elliott, Rob Price, Patricia Shaw, Tasos Spiliotopoulos, Magdalene Ng, Kovila Coopamootoo, and Aad Van Moorsel. 2021. Towards an equitable digital society: artificial intelligence (AI) and corporate digital responsibility (CDR). *Society* 58, 3 (2021), 179–188.
- [44] Robin J Ely. 1995. The power in demography: Women's social constructions of gender identity at work. *Academy of Management journal* 38, 3 (1995), 589–634.
- [45] Norman T Feather. 1969. Attribution of responsibility and valence of success and failure in relation to initial confidence and task performance. *Journal of Personality and Social Psychology* 13, 2 (1969), 129.
- [46] L. Dee Fink. 2003. *Creating Significant Learning Experiences: An Integrated Approach to Designing College Courses*. Jossey-Bass, San Francisco, CA.
- [47] John H. Flavell. 1979. Metacognition and Cognitive Monitoring: A New Area of Cognitive-Developmental Inquiry. *American Psychologist* 34 (1979), 906–911. <https://doi.org/10.1037/0003-066X.34.10.906>
- [48] Jonas Frich, Lindsay MacDonald Vermeulen, Christian Remy, Michael Mose Biskjaer, and Peter Dalsgaard. 2019. Mapping the landscape of creativity support tools in HCI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–18.
- [49] Claudia Gama. 2004. Metacognition in interactive learning environments: The reflection assistant model. In *International conference on intelligent tutoring systems*. Springer, 668–677.
- [50] Peter Gerjets, Katharina Scheiter, and Richard Catrambone. 2006. Can learning from molar and modular worked examples be enhanced by providing instructional explanations and prompting self-explanations? *Learning and Instruction* 16, 2 (2006), 104–121. <https://doi.org/10.1016/j.learninstruc.2006.02.007>
- [51] Andreas Göldi, Thiendo Wambganss, Seyed Parsa Neshaei, and Roman Rietsche. 2024. Intelligent Support Engages Writers Through Relevant Cognitive Processes. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 1047, 12 pages. <https://doi.org/10.1145/3613904.3642549>
- [52] Jarod Govers, Eduardo Veloso, Vassilis Kostakos, and Jorge Goncalves. 2024. AI-Driven Mediation Strategies for Audience Depolarisation in Online Debates. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 803, 18 pages. <https://doi.org/10.1145/3613904.3642322>
- [53] Peter Gray. 2011. The decline of play and the rise of psychopathology in children and adolescents. *American journal of play* 3, 4 (2011), 443–463.
- [54] Tovi Grossman, Pierre Dragicevic, and Ravin Balakrishnan. 2007. Strategies for accelerating on-line learning of hotkeys. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (*CHI '07*). Association for Computing Machinery, New York, NY, USA, 1591–1600. <https://doi.org/10.1145/1240624.1240865>
- [55] Ken Gu, Madeleine Grunde-McLaughlin, Andrew McNutt, Jeffrey Heer, and Tim Althoff. 2024. How Do Data Analysts Respond to AI Assistance? A Wizard-of-Oz Study. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 1015, 22 pages. <https://doi.org/10.1145/3613904.3641891>
- [56] Jonathan Haidt. 2024. *The Anxious Generation*. Allen Lane.
- [57] Diane F. Halpern. 1996. *Thought & Knowledge: An Introduction to Critical Thinking* (3rd ed.). Lawrence Erlbaum Associates.
- [58] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (task load index): Results of empirical and theoretical research. In *Advances in Psychology*. Vol. 52. Elsevier, 139–183. [https://doi.org/10.1016/s0166-4115\(08\)62386-9](https://doi.org/10.1016/s0166-4115(08)62386-9)
- [59] N. Katherine Hayles. 1999. *How We Became Posthuman*. The University of Chicago Press.
- [60] Hui-Ru Ho, Edward M. Hubbard, and Bilge Mutlu. 2024. "It's Not a Replacement:" Enabling Parent-Robot Collaboration to Support In-Home Learning Experiences of Young Children. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 122, 18 pages. <https://doi.org/10.1145/3613904.3642806>
- [61] Michael H Hopson, Richard L Simms, and Gerald A Knezek. 2001. Using a technology-enriched environment to improve higher-order thinking skills. *Journal of Research on Technology in education* 34, 2 (2001), 109–119.
- [62] Kasper Hornbæk and Morten Hertzum. 2017. Technology acceptance and user experience: A review of the experiential component in HCI. *ACM Transactions on Computer-Human Interaction (TOCHI)* 24, 5 (2017), 1–30.
- [63] Saki Inoue, Yuanyuan Wang, Yukiko Kawai, and Kazutoshi Sumiya. 2023. Encouraging Critical Thinking Support System: Question Generation and Lecture Slide Recommendations. In *Proceedings of the Tenth ACM Conference on Learning@ Scale*. 287–291.
- [64] Farnaz Jahanbakhsh, Yannis Katsis, Dakuo Wang, Lucian Popa, and Michael Muller. 2023. Exploring the Use of Personalized AI for Identifying Misinformation on Social Media. 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 105, 27 pages. <https://doi.org/10.1145/3544548.3581219>
- [65] Junfeng Jiao, Saleh Afroogh, Yiming Xu, and Connor Phillips. 2024. Navigating llm ethics: Advancements, challenges, and future directions. *arXiv preprint arXiv:2406.18841* (2024).
- [66] Hyoungwook Jin, Seonghee Lee, Hyungyu Shin, and Juho Kim. 2024. Teach AI How to Code: Using Large Language Models as Teachable Agents for Programming Education. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 652, 28 pages. <https://doi.org/10.1145/3613904.3642349>
- [67] Jr. John E. Barbuto and Richard W. Scholl. 1998. Motivation Sources Inventory: Development and Validation of New Scales to Measure an Integrative Taxonomy of Motivation. *Psychological Reports* 82, 3 (1998), 1011–1022. <https://doi.org/10.2466/pr0.1998.82.3.1011>
- [68] David H. Jonassen. 2000. *Computers as Mindtools for Schools: Engaging Critical Thinking* (2nd ed.). Merrill.
- [69] Karl O. Jones, Janice Harland, Juliet M.V. Reid, and Rebecca Bartlett. 2009. Relationship between examination questions and bloom's taxonomy. In *2009 39th IEEE Frontiers in Education Conference*. 1–6. <https://doi.org/10.1109/FIE.2009.5350598>
- [70] Prerna Juneja, Wenjuan Zhang, Alison Marie Smith-Renner, Hemank Lamba, Joel Tetreault, and Alex Jaimes. 2024. Dissecting users' needs for search result explanations. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 841, 17 pages. <https://doi.org/10.1145/3613904.3642059>
- [71] Kowe Kodama, Marianne Aubin Le Quere, Xiyu Jenny Fu, Christin Munsch, Danaë Metaxa, and Mor Naaman. 2024. The Role of Inclusion, Control, and Ownership in Workplace AI-Mediated Communication. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 1016, 10 pages. <https://doi.org/10.1145/3613904.3642650>
- [72] Slava Kalyuga. 2011. Cognitive Load Theory: How Many Types of Load Does It Really Need? *Educational Psychology Review* 23, 1 (2011), 1–19.
- [73] Majeed Kazemitaabar, Runlong Ye, Xiaoning Wang, Austin Zachary Henley, Paul Denny, Michelle Craig, and Tovi Grossman. 2024. CodeAid: Evaluating a Classroom Deployment of an LLM-based Programming Assistant that Balances Student and Educator Needs. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 650, 20 pages. <https://doi.org/10.1145/3613904.3642773>
- [74] Taemie Kim and Pamela Hinds. 2006. Who should I blame? Effects of autonomy and transparency on attributions in human-robot interaction. In *ROMAN*

- 2006-The 15th IEEE international symposium on robot and human interactive communication*. IEEE, 80–85.
- [75] Taewan Kim, Mintra Ruensuk, and Hwajung Hong. 2020. In Helping a Vulnerable Bot, You Help Yourself: Designing a Social Bot as a Care-Receiver to Promote Mental Health and Reduce Stigma. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3313831.3376743>
- [76] Tae Soo Kim, Yoonjoo Lee, Jamin Shin, Young-Ho Kim, and Juho Kim. 2024. EvalLM: Interactive Evaluation of Large Language Model Prompts on User-Defined Criteria. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 306, 21 pages. <https://doi.org/10.1145/3613904.3642216>
- [77] Charlotte Kobiella, Yarhy Said Flores López, Franz Waltenberger, Fiona Draxler, and Albrecht Schmidt. 2024. “If the Machine Is As Good As Me, Then What Use Am I?” – How the Use of ChatGPT Changes Young Professionals’ Perception of Productivity and Accomplishment. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 1018, 16 pages. <https://doi.org/10.1145/3613904.3641964>
- [78] Nils Köbis, Jean-François Bonnefon, and Iyad Rahwan. 2021. Bad machines corrupt good morals. *Nature Human Behaviour* 5 (June 2021), 679–685. <https://doi.org/10.1038/s41562-021-01070-8>
- [79] Emily R Lai. 2011. Metacognition: A literature review. *Always learning: Pearson research report* 24 (2011), 1–40.
- [80] Susan Lechelt, Yvonne Rogers, and Nicolai Marquardt. 2020. Coming to your senses: promoting critical thinking about sensors through playful interaction in classrooms. In *Proceedings of the interaction design and children conference*. 11–22.
- [81] Sunok Lee, Dasom Choi, Minha Lee, Jonghak Choi, and Sangsu Lee. 2023. Fostering Youth’s Critical Thinking Competency About AI through Exhibition. 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 451, 22 pages. <https://doi.org/10.1145/3544548.3581159>
- [82] Yoonjoo Lee, Hyeyoung B Kang, Matt Latzke, Juho Kim, Jonathan Bragg, Joseph Chee Chang, and Pao Siangliuhue. 2024. PaperWeaver: Enriching Topical Paper Alerts by Contextualizing Recommended Papers with User-collected Papers. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 19, 19 pages. <https://doi.org/10.1145/3613904.3642196>
- [83] Yoonjoo Lee, Tae Soo Kim, Sungdong Kim, Yohan Yun, and Juho Kim. 2023. DAPIE: Interactive Step-by-Step Explanatory Dialogues to Answer Children’s Why and How Questions. 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 450, 22 pages. <https://doi.org/10.1145/3544548.3581369>
- [84] Florian Leiser, Sven Eckhardt, Valentin Leuthe, Merlin Knaebel, Alexander Mäadche, Gerhard Schwabe, and Ali Sunyaev. 2024. HILL: A Hallucination Identifier for Large Language Models. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 482, 13 pages. <https://doi.org/10.1145/3613904.3642428>
- [85] Arthur Lewis and David Smith. 1993. Defining higher order thinking. *Theory into practice* 32, 3 (1993), 131–137.
- [86] Tianyi Li, Mihaela Vorvoreanu, Derek Debellis, and Saleema Amershi. 2023. Assessing Human-AI Interaction Early through Factorial Surveys: A Study on the Guidelines for Human-AI Interaction. *ACM Trans. Comput.-Hum. Interact.* 30, 5, Article 69 (sep 2023), 45 pages. <https://doi.org/10.1145/3511605>
- [87] Bingjie Liu. 2021. In AI we trust? Effects of agency locus and transparency on uncertainty reduction in human–AI interaction. *Journal of computer-mediated communication* 26, 6 (2021), 384–402.
- [88] Chengzhong Liu, Shixu Zhou, Dingdong Liu, Junze Li, Zeyu Huang, and Xiaojian Ma. 2023. CoArgue: Fostering Lurkers’ Contribution to Collective Arguments in Community-based QA Platforms. 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 271, 17 pages. <https://doi.org/10.1145/3544548.3580932>
- [89] Michael Xieyang Liu, Tongshuang Wu, Tianying Chen, Franklin Mingzhe Li, Aniket Kittur, and Brad A Myers. 2024. Selenite: Scaffolding Online Sensemaking with Comprehensive Overviews Elicited from Large Language Models. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 837, 26 pages. <https://doi.org/10.1145/3613904.3642149>
- [90] Yiren Liu, Si Chen, Haocong Cheng, Mengxia Yu, Xiao Ran, Andrew Mo, Yiliu Tang, and Yun Huang. 2024. How AI Processing Delays Foster Creativity: Exploring Research Question Co-Creation with an LLM-based Agent. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 17, 25 pages. <https://doi.org/10.1145/3613904.3642698>
- [91] Ziyi Liu, Zhengzhe Zhu, Lijun Zhu, Enze Jiang, Xiyun Hu, Kylie A Pepple, and Karthik Ramani. 2024. ClassMeta: Designing Interactive Virtual Classmate to Promote VR Classroom Participation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 659, 17 pages. <https://doi.org/10.1145/3613904.3642947>
- [92] Xinyi Lu, Simin Fan, Jessica Houghton, Lu Wang, and Xu Wang. 2023. ReadingQuizMaker: A Human-NLP Collaborative System that Supports Instructors to Design High-Quality Reading Quiz Questions. 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 454, 18 pages. <https://doi.org/10.1145/3544548.3580957>
- [93] Susana Masapanta-Carrión and J. Ángel Velázquez-Iturbiide. 2018. A Systematic Review of the Use of Bloom’s Taxonomy in Computer Science Education. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education* (Baltimore, Maryland, USA) (SIGCSE '18). Association for Computing Machinery, New York, NY, USA, 441–446. <https://doi.org/10.1145/3159450.3159491>
- [94] Edward McAuley, Terry Duncan, and Vance V Tammem. 1989. Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research quarterly for exercise and sport* 60, 1 (1989), 48–58.
- [95] John R. McClure, Brian Sonak, and Hoi K. Suen. 1999. Concept Map Assessment of Classroom Learning: Reliability, Validity, and Logistical Practicality. *Journal of Research in Science Teaching* 36, 4 (1999), 475–492. [https://doi.org/10.1002/\(SICI\)1098-2736\(199904\)36:4<475::AID-TEA5>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1098-2736(199904)36:4<475::AID-TEA5>3.0.CO;2-O)
- [96] John E. McPeck. 1981. *Critical Thinking and Education*. Martin Robertson.
- [97] Karen Meador. 1997. *Creative thinking and problem solving for young learners*. Bloomsbury Publishing USA.
- [98] Michael Michalko. 2010. *Thinkertoys: A handbook of creative-thinking techniques*. Ten Speed Press.
- [99] William R Miller and Stephen Rollnick. 2012. *Motivational interviewing: Helping people change*. Guilford press.
- [100] Piotr Mirowski, Kory W. Mathewson, Jaylen Pittman, and Richard Evans. 2023. Co-Writing Screenplays and Theatre Scripts with Language Models: Evaluation by Industry Professionals. 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 355, 34 pages. <https://doi.org/10.1145/3544548.3581225>
- [101] Brent Daniel Mittelstadt, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. 2016. The ethics of algorithms: Mapping the debate. *Big Data & Society* 3, 2 (2016), 2053951716679679.
- [102] Tomohito Miyake, Yuji Kawai, Jihou Park, Jiro Shimaya, Hideyuki Takahashi, and Minoru Asada. 2019. Mind perception and causal attribution for failure in a game with a robot. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 1–6.
- [103] Joseph D. Novak and D. Bob Gowin. 1984. *Learning How to Learn*. Cambridge University Press, Cambridge, UK.
- [104] H. S. Nwana. 1990. Intelligent tutoring systems: an overview. *Artificial Intelligence Review* 4 (December 1990), 251–277. <https://doi.org/10.1007/BF00168958>
- [105] Giuliana Orru and Luca Longo. 2019. The Evolution of Cognitive Load Theory and the Measurement of Its Intrinsic, Extraneous and Germane Loads: A Review. In *Human Mental Workload: Models and Applications*, Luca Longo and Matteo Leva (Eds.). Communications in Computer and Information Science, Vol. 1012. Springer, Cham. [https://doi.org/10.1007/978-3-030-14273-5\\_3](https://doi.org/10.1007/978-3-030-14273-5_3) H-WORKLOAD 2018.
- [106] Srishtha Palani, David Ledo, George Fitzmaurice, and Fraser Anderson. 2022. “I don’t want to feel like I’m working in a 1960s factory”: The Practitioner Perspective on Creativity Support Tool Adoption. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–18.
- [107] Richard Paul and Linda Elder. 2013. *Critical Thinking: Tools for Taking Charge of Your Professional and Personal Life*. Pearson Education.
- [108] Ana Pérez-Escoda, Rosa García-Ruiz, Ana Castro-Zubizarreta, and Ignacio Aguaded. 2017. Media literacy and digital skills for enhancing critical thinking in networked society. In *Proceedings of the 5th International Conference on Technological Ecosystems for Enhancing Multiculturality*. 1–7.
- [109] Savvas Petridis, Nicholas Diakopoulos, Kevin Crowston, Mark Hansen, Keren Henderson, Stan Jastrzebski, Jeffrey V Nickerson, and Lydia B Chilton. 2023. AngleKindling: Supporting Journalistic Angle Ideation with Large Language Models. 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 225, 16 pages. <https://doi.org/10.1145/3544548.3580907>
- [110] Margaret M. Plack, Maryanne Driscoll, Maria Marquez, Lynn Cuppernull, Joyce Maring, and Larrie Greenberg. 2007. Assessing Reflective Writing on a Pediatric Clerkship by Using a Modified Bloom's Taxonomy. *Ambulatory Pediatrics* 7, 4 (2007), 285–291. <https://doi.org/10.1016/j.ambp.2007.04.006>
- [111] Niroop Channa Rajashekhar, Yeo Eun Shin, Yuan Pu, Sunny Chung, Kisung You, Mauro Giuffre, Colleen E Chan, Theo Saarinen, Allen Hsiao, Jasjeet Sekhon, Ambrose H Wong, Leigh V Evans, Rene F. Kizilcec, Loren Laine, Terika McCall, and Dennis Shung. 2024. Human-Algorithmic Interaction Using a Large Language Model-Augmented Artificial Intelligence Clinical Decision Support System. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 442, 20 pages. <https://doi.org/10.1145/3613904.3642024>
- [112] PL Patrick Rau, Ye Li, and Dingjun Li. 2009. Effects of communication style and culture on ability to accept recommendations from robots. *Computers in Human Behavior* 25, 2 (2009), 587–595.
- [113] Christian Remy, Lindsay MacDonald Vermeulen, Jonas Frich, Michael Mose Biskjaer, and Peter Dalsgaard. 2020. Evaluating creativity support tools in HCI research. In *Proceedings of the 2020 ACM designing interactive systems conference*, 457–476.
- [114] Mohit Reza, Nathan M Laundry, Ilya Musabirov, Peter Dushniku, Zhi Yuan "Michael" Yu, Kashish Mittal, Tovi Grossman, Michael Liut, Anastasia Kuzminykh, and Joseph Jay Williams. 2024. ABScript: Rapid Exploration & Organization of Multiple Writing Variations in Human-AI Co-Writing Tasks using Large Language Models. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 1042, 18 pages. <https://doi.org/10.1145/3613904.3641899>
- [115] Yann Riche, Nathalie Henry Riche, Petra Isenberg, and Anastasia Bezerianos. 2010. Hard-to-use interfaces considered beneficial (some of the time). In *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (Atlanta, Georgia, USA) (*CHI EA '10*). Association for Computing Machinery, New York, NY, USA, 2705–2714. <https://doi.org/10.1145/1753846.1753855>
- [116] Debbie Roberts. 2010. Vicarious learning: A review of the literature. *Nurse Education in practice* 10, 1 (2010), 13–16.
- [117] Anthony Robins, Janet Rountree, and Nathan Rountree. 2003. Learning and teaching programming: A review and discussion. *Computer science education* 13, 2 (2003), 137–172.
- [118] Jeremy M Roschelle, Roy D Pea, Christopher M Hoadley, Douglas N Gordin, and Barbara M Means. 2000. Changing how and what children learn in school with computer-based technologies. *The future of children* (2000), 76–101.
- [119] Daniel Russo. 2024. Navigating the complexity of generative ai adoption in software engineering. *ACM Transactions on Software Engineering and Methodology* (2024).
- [120] Morgan Klaus Scheuerman and Jed R Brubaker. 2024. Products of positionality: How tech workers shape identity concepts in computer vision. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–18.
- [121] Christoph Schneider, Markus Weinmann, and Jan Vom Brocke. 2018. Digital nudging: guiding online user choices through interface design. *Commun. ACM* 61, 7 (2018), 67–73.
- [122] Nikhil Sharma, Q. Vera Liao, and Ziang Xiao. 2024. Generative Echo Chamber? Effect of LLM-Powered Search Systems on Diverse Information Seeking. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 1033, 17 pages. <https://doi.org/10.1145/3613904.3642459>
- [123] Chuhan Shi, Yicheng Hu, Shenan Wang, Shuai Ma, Chengbo Zheng, Xiaojuan Ma, and Qiong Luo. 2023. RetroLens: A Human-AI Collaborative System for Multi-step Retrosynthetic Route Planning. 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 770, 20 pages. <https://doi.org/10.1145/3544548.3581469>
- [124] Ben Shneiderman. 2022. *Human-Centered AI* (1st. ed.). Oxford University Press.
- [125] Konstantin R Strömel, Stanislav Henry, Tim Johansson, Jasmin Niess, and Paweł W. Woźniak. 2024. Narrating Fitness: Leveraging Large Language Models for Reflective Fitness Tracker Data Interpretation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 646, 16 pages. <https://doi.org/10.1145/3613904.3642032>
- [126] Na Sun, Chien Wen Yuan, Mary Beth Rosson, Yu Wu, and Jack M Carroll. 2017. Critical thinking in collaboration: Talk less, perceive more. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 2944–2950.
- [127] John Sweller. 1988. Cognitive load during problem solving: Effects on learning. *Cognitive Science* 12, 2 (1988), 257–285. [https://doi.org/10.1207/s15516709cog1202\\_4](https://doi.org/10.1207/s15516709cog1202_4)
- [128] 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 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. <https://doi.org/10.1145/3613904.3642902>
- [129] Thitaree Tanprasert, Sidney S Fels, Luanne Sinnamon, and Dongwook Yoon. 2024. Debate Chatbots to Facilitate Critical Thinking on YouTube: Social Identity and Conversational Style Make A Difference. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 805, 24 pages. <https://doi.org/10.1145/3613904.3642513>
- [130] Adam Tapal, Ela Oren, Reuven Dar, and Baruch Eitam. 2017. The Sense of Agency Scale: A Measure of Consciously Perceived Control over One's Mind, Body, and the Immediate Environment. *Frontiers in Psychology* 8 (2017), 1552. <https://doi.org/10.3389/fpsyg.2017.01552>
- [131] Richard H. Thaler and Cass R. Sustein. 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Penguin.
- [132] E. Paul Torrance. 1964. *Guiding Creative Talent*. Prentice-Hall.
- [133] Ester van Laar, Alexander J.A.M. van Deursen, Jan A.G.M. van Dijk, and Jos de Haan. 2017. The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior* 72 (2017), 577–588. <https://doi.org/10.1016/j.chb.2017.03.010>
- [134] Samangi Wadinambiarchchi, Ryan M. Kelly, Saumya Pareek, Qiushi Zhou, and Eduardo Veloso. 2024. The Effects of Generative AI on Design Fixation and Divergent Thinking. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 380, 18 pages. <https://doi.org/10.1145/3613904.3642919>
- [135] Samangi Wadinambiarchchi, Ryan M Kelly, Saumya Pareek, Qiushi Zhou, and Eduardo Veloso. 2024. The Effects of Generative AI on Design Fixation and Divergent Thinking. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–18.
- [136] Ben Wagner. 2019. Liable, but not in control? Ensuring meaningful human agency in automated decision-making systems. *Policy & Internet* 11, 1 (2019), 104–122.
- [137] Shun-Yu Wang, Wei-Chung Su, Serena Chen, Ching-Yi Tsai, Marta Misztal, Katherine M. Cheng, Alwena Lin, Yu Chen, and Mike Y. Chen. 2024. RoomDreaming: Generative-AI Approach to Facilitating Iterative, Preliminary Interior Design Exploration. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 379, 20 pages. <https://doi.org/10.1145/3613904.3642901>
- [138] Xinru Wang, Hannah Kim, Sajjadur Rahman, Kushan Mitra, and Zhengjie Miao. 2024. Human-LLM Collaborative Annotation Through Effective Verification of LLM Labels. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 303, 21 pages. <https://doi.org/10.1145/3613904.3641960>
- [139] Zijie J. Wang, Chimay Kulkarni, Lauren Wilcox, Michael Terry, and Michael Madaio. 2024. Farsight: Fostering Responsible AI Awareness During AI Application Prototyping. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 976, 40 pages. <https://doi.org/10.1145/3613904.3642335>
- [140] Mary Katherine Watson, Joshua Pelkey, Caroline R. Noyes, and Michael O. Rodgers. 2015. Assessing Conceptual Knowledge Using Three Concept Map Scoring Methods. *Journal of Engineering Education* 104, 3 (2015), 395–411. <https://doi.org/10.1002/jee.20111>
- [141] Howard T Welser, Eric Gleave, Danyel Fisher, and Marc Smith. 2007. Visualizing the signatures of social roles in online discussion groups. *Journal of social structure* 8, 2 (2007), 1–32.
- [142] David Wood, Jerome S Bruner, and Gail Ross. 1976. The role of tutoring in problem solving. *Journal of child psychology and psychiatry* 17, 2 (1976), 89–100.
- [143] Allison Woodruff, Renee Shelby, Patrick Gage Kelley, Steven Rousso-Schindler, Jamila Smith-Loud, and Lauren Wilcox. 2024. How Knowledge Workers Think Generative AI Will (Not) Transform Their Industries. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 641, 26 pages. <https://doi.org/10.1145/3613904.3642700>
- [144] Ruolan Wu, Chun Yu, Xiaole Pan, Yujia Liu, Ningning Zhang, Yue Fu, Yuhuan Wang, Zhi Zheng, Li Chen, Qiaolei Jiang, Xuhai Xu, and Yuanchun Shi. 2024. MindShift: Leveraging Large Language Models for Mental-States-Based Problematic Smartphone Use Intervention. In *Proceedings of the CHI Conference*

- on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 248, 24 pages. <https://doi.org/10.1145/3613904.3642790>
- [145] Ziang Xiao, Tiffany Wenting Li, Karrie Karahalios, and Hari Sundaram. 2023. Inform the Uninformed: Improving Online Informed Consent Reading with an AI-Powered Chatbot. 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 112, 17 pages. <https://doi.org/10.1145/3544548.3581252>
- [146] Yusuke Yamamoto and Takehiro Yamamoto. 2018. Query priming for promoting critical thinking in web search. In *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval*. 12–21.
- [147] Litao Yan, Alyssa Hwang, Zhiyuan Wu, and Andrew Head. 2024. Ivie: Lightweight Anchored Explanations of Just-Generated Code. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 140, 15 pages. <https://doi.org/10.1145/3613904.3642239>
- [148] Nur Yildirim, Mahima Pushkarna, Nitesh Goyal, Martin Wattenberg, and Fernanda Viégas. 2023. Investigating How Practitioners Use Human-AI Guidelines: A Case Study on the People + AI Guidebook. 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 356, 13 pages. <https://doi.org/10.1145/3544548.3580900>
- [149] Nur Yildirim, Hannah Richardson, Maria Teodora Wetscherek, Junaid Bajwa, Joseph Jacob, Mark Ames Pinnock, Stephen Harris, Daniel Coelho De Castro, Shruthi Bannur, Stephanie Hyland, Pratik Ghosh, Mercy Ranjit, Kenza Bouzid, Anton Schwaighofer, Fernando Pérez-García, Harshita Sharma, Ozan Oktay, Matthew Lungren, Javier Alvarez-Valle, Aditya Nori, and Anja Thieme. 2024. Multimodal Healthcare AI: Identifying and Designing Clinically Relevant Vision-Language Applications for Radiology. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 444, 22 pages. <https://doi.org/10.1145/3613904.3642013>
- [150] Liudmila Zavolokina, Kilian Sprenkamp, Zoya Katazhinskaya, Daniel Gordon Jones, and Gerhard Schwabe. 2024. Think Fast, Think Slow, Think Critical: Designing an Automated Propaganda Detection Tool. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 491, 24 pages. <https://doi.org/10.1145/3613904.3642805>
- [151] Chao Zhang, Xuechen Liu, Katherine Ziska, Soobin Jeon, Chi-Lin Yu, and Ying Xu. 2024. Mathemyths: Leveraging Large Language Models to Teach Mathematical Language through Child-AI Co-Creative Storytelling. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 274, 23 pages. <https://doi.org/10.1145/3613904.3642647>
- [152] Xiaoyu Zhang, Jianping Li, Po-Wei Chi, Senthil Chandrasegaran, and Kwan-Liu Ma. 2023. ConceptEVA: Concept-Based Interactive Exploration and Customization of Document Summaries. 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 204, 16 pages. <https://doi.org/10.1145/3544548.3581260>
- [153] Yu Zhang, Jingwei Sun, Li Feng, Cen Yao, Mingming Fan, Liuxin Zhang, Qianying Wang, Xin Geng, and Yong Rui. 2024. See Widely, Think Wisely: Toward Designing a Generative Multi-agent System to Burst Filter Bubbles. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 484, 24 pages. <https://doi.org/10.1145/3613904.3642545>
- [154] Suifang Zhou, Latisha Besariani Hendra, Qinshi Zhang, Jussi Holopainen, and RAY LC. 2024. Eternagram: Probing Player Attitudes Towards Climate Change Using a ChatGPT-driven Text-based Adventure. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '24*). Association for Computing Machinery, New York, NY, USA, Article 281, 23 pages. <https://doi.org/10.1145/3613904.3642850>
- [155] Philip G Zimbardo. 1995. The psychology of evil: A situationist perspective on recruiting good people to engage in anti-social acts. *Japanese Journal of Social Psychology* 11, 2 (1995), 125–133.