Understanding, Protecting, and Augmenting Human Cognition with Generative AI:
A Synthesis of the CHI 2025 *Tools for Thought Workshop*

Lev Tankelevitch
lev.tankelevitch@microsoft.com
Microsoft Research
Cambridge, United Kingdom

Elena L. Glassman
glassman@seas.harvard.edu
Harvard University
USA

Jessica He
jessicahe@ibm.com
IBM Research
USA

Aniket Kittur
nkittur@andrew.cmu.edu
Carnegie Mellon University
USA

Mina Lee
mnlee@cs.uchicago.edu
University of Chicago
USA

Srishti Palani
srishti.palani@salesforce.com
Tableau Research
USA

Advait Sarkar
advait@microsoft.com
Microsoft Research
Cambridge, United Kingdom

Gonzalo Ramos
goramos@microsoft.com
Microsoft Research
Redmond, USA

Yvonne Rogers
y.rogers@ucl.ac.uk
University College London
London, United Kingdom

Hari Subramonyam
harihars@stanford.edu
Stanford University
USA

Figure 1: The CHI 2025 *Tools for Thought Workshop* was held in Yokohama, Japan, on April 26th, 2025.

Abstract

Generative AI (GenAI) radically expands the scope and capability of automation for work, education, and everyday tasks, a transformation posing both risks and opportunities for human cognition. How will human cognition change, and what opportunities are there for GenAI to augment it? Which theories, metrics, and other tools are needed to address these questions? The CHI 2025 workshop on Tools for Thought aimed to bridge an emerging science of how the use of GenAI affects human thought, from metacognition to critical thinking, memory, and creativity, with an emerging design practice for building GenAI tools that both protect and augment human thought. Fifty-six researchers, designers, and thinkers from across disciplines as well as industry and academia, along with 34 papers and portfolios, seeded a day of discussion, ideation, and community-building. We synthesize this material here to begin mapping the space of research and design opportunities and to catalyze a multidisciplinary community around this pressing area of research.

1 Introduction

With the development of Generative AI (GenAI), society is undergoing a radical expansion in the scope and capability of automation and mechanization of cognitive work, a transformation which poses important risks for human cognition. Yet, the current moment also

presents an expansion in the opportunities to *positively* transform how we think. How will people learn and think in the future, as AI is embedded into our work and lives as a 'co-worker', tool, or ubiquitous capability? How can we protect human cognition from the potential negative impacts of AI-driven automation? What opportunities are there for AI to augment thinking, much as older external tools such as writing have done? Which theories, metrics, and other tools can best equip researchers and designers to address these questions? The CHI 2025 workshop on *Tools for Thought* [123]—systems or methods that extend, support, or augment cognition [36, 126]—aimed to bridge an emerging science of how the use of AI affects human thought, from metacognition to critical thinking, memory, and creativity, with an emerging design practice for building AI tools that not only protect human thought but also augment it. Concretely, it aimed to map the space of research and design opportunities in this area, and catalyze a multidisciplinary community interested in pursuing them.

Held in Yokohama in April 2025, the workshop convened 56 researchers, designers, and thinkers from industry and academia, covering fields such as human-computer interaction and other aspects of computer science, cognitive science, design, and education. Thirty-four accepted papers and portfolios (selected from over 70 submissions) set the stage for a day of discussion, ideation, and community-building.[1] Here we synthesize the workshop outputs to advance the conversation on this pressing area of research. We structure this article around three areas: (i) *understanding AI's impact on, and protecting, cognition*, (ii) *augmenting human cognition with AI*, and (iii) *formative research, theory, measurement, and evaluation*.

2 Understanding AI's impact on, and protecting, human cognition

AI systems are designed and commonly used to automate capabilities and processes such as writing and research, processes through which people think, and therefore learn, build skills, and grow expertise. The generative nature of these systems means that human workflows are shifting from production to 'critical integration' of material [97], i.e., decisions about when and how to use AI, how to frame tasks, and how to assess outputs. Thus, the quality and nature of thinking, such as where and what kind of thinking is applied, all change when knowledge workflows reorient themselves around AI [124]. As the next sections illustrate, these changes may manifest in critical thinking, learning, and creativity, among other cognitive processes. They have implications for workflows and our notions of expertise, as well as motivation, well-being, and other human values.

2.1 Critical thinking

As Singh et al. [115] observe in emerging research, AI may discourage people from engaging in critical thinking, as seen, for example, in the shift from active information seeking to more passive consumption of AI-generated information (e.g., see [24]). This is

exacerbated by GenAI systems' tendencies for agreeing with users' ideas and homogenizing outputs, features which may contribute to the known 'echo chamber' effects in contemporary information consumption. Beyond mere consumption, people may also inappropriately rely on inaccurate or misleading AI-generated information in everyday decisions [115], particularly as they may disengage from critical thinking when their self-confidence in a domain is low, or when their confidence in AI is high [66]. For example, biased writing assistants can shift users' views on key issues [130] and homogenize writing toward western styles [2]. More broadly, AI use has been shown to exert an aggregate influence on the intentions of knowledge workers [100], nudging the outputs of groups of workers towards smaller, less diverse sets of ideas, in the process of "mechanized convergence" [97].

How can we understand these changes from the perspective of human cognition? Singh et al. [115] point to Dewey's theory of reflective thinking [29] as a powerful explanatory lens. Reflective thinking, as an intentional and conscious evaluation of evidence, requires enduring a 'state of perplexity, confusion, or doubt prompting inquiry' [115], as well as a suspension of judgment during this period of inquiry. However, the polished, seemingly coherent, and sycophantic nature of AI-generated information may minimize this prerequisite state, discourage suspension of judgment, and create an 'illusion of comprehensive understanding' [115]. Empirically testing which aspects of AI output and interaction patterns causally impact critical thinking in different contexts is an important and ongoing research direction [61, 62, 66, 85]. How these effects play out after prolonged use is another important question. Worryingly, the potential impacts on critical thinking may be particularly acute in a learning context.

2.2 Learning

Unlike experts, students or novices are still developing the self-regulation strategies and schemas (i.e., mental models or sense-making frames) necessary for information work [84, 115, 116]. Thus, in addition to AI-generated output that can discourage a reflective state, students can also lack a developed skill set for prompting and critically evaluating AI output. For example, in a study of AI-assisted ML code debugging, Bo et al. [10] found that students who already had an understanding of the problem used AI intentionally by asking specific, planned questions, and ultimately performed better. In contrast, those who lacked such knowledge asked more open-ended questions and were led by the AI down unproductive paths, engaging in more over-reliance. As Prather and Reeves [89] conclude about AI assistance in CS education, "underprepared, unconfident, and underperforming students seem to benefit the least from these [AI] tools, resulting in less critical thinking and lower grades". Thus, those populations most likely to rely on AI due to existing performance disparities may benefit the least, or may even be harmed by using it.

By inappropriately relying on (non-pedagogical) AI tools to help solve their learning tasks, students may also impede their development of necessary metacognitive strategies and schemas [115]. Students access synthesized and tailor-made information without engaging in the cognitive processes typical of deliberate learning, e.g., remembering, understanding, and applying disparate

---

[1]Throughout the paper, each accepted submission is first introduced with a highlight like this. Subsequent citations follow ordinary formatting conventions. PDFs of all accepted submissions can be found [here].

sources of information, as outlined in Bloom's revised taxonomy of educational goals [63]. Indeed, in a study of AI-assisted reading, Fu and Hiniker [43] find that students may increasingly offload lower-level understanding to AI, consistent with more passive engagement. These observations naturally raise questions such as: how do we help people use GenAI effectively for *learning*, rather than, or alongside, productivity purposes? What can we gather from those already using GenAI successfully for this purpose?

2.3 Creativity

In the creative domain, Dalsgaard [20] argues that GenAI's facility for producing higher-fidelity outputs risks shifting designers' focus from understanding the core problem framing or design space to the mere tweaking of polished outputs, a form of design fixation [70]. There is also a risk of GenAI homogenizing creative outputs [3, 32], an effect potentially amplified by the challenge of working effectively with GenAI systems and grasping their subtle limitations [124] (see also [5, 64]).

Yet, Dalsgaard [20] positions GenAI not solely as a risk for creativity, but rather as providing 'instruments of inquiry': a set of tools that shape and guide creative thinking in myriad ways. Grounded in pragmatist philosophy, this perspective suggests that GenAI has a role in shaping people's perception about a design problem, their conceptions, how they externalize and manipulate ideas, and the mediation of the components of a design task. The nature of these effects depends on users' approach and skill set for working with GenAI. For example, designers can leverage GenAI's ability for rapid, high-fidelity prototyping to better engage in Schön's process of 'reflection-in-action' [110], where iterative interaction with a design material affords a rapid feedback loop. Likewise, designers aware of GenAI's limitations can consciously exploit these for more creative outcomes. Empirically disentangling helpful and harmful uses of GenAI in creative processes and workflows, and the roles of interaction patterns and system capabilities is an important research question (e.g., [5]).

Aspects of GenAI systems, such as the dominance of text-based prompting or their limited explainability, also raise questions about how to preserve practitioners' flow states and integrate GenAI into existing practices [86]. Naqvi et al. [81] highlight Sawyer's model of the creative process, which characterizes it according to eight features such as iteration, reflection, ambiguity, exploration, and emergence. GenAI may be shifting how these characteristics arise in creative workflows. For example, reflection may shift from a focus on one's own creative process and output to prompting and evaluating AI-generated output, whereas ambiguity may now be driven more by models' lack of explainability or other limitations, rather than the problem space or iterative workflow itself. Quintana and Quintana [94] find evidence of this in AI-assisted learning experience design, where attention may shift from a focus on the design workflow itself to prompt engineering AI to get the desired output. What do these workflow changes mean for students who are still developing the schemas necessary for design work? How can AI become a 'ready-to-hand tool', available for almost unconscious use, rather than a 'present-at-hand' tool that requires focused effort [94]?[2]

2.4 Contextual and applied factors: Workflows, expertise, and human values

Beyond the impact on any one cognitive domain, several contextual and applied factors are central to understanding how GenAI may impact cognition in realistic settings. These include considerations of workflows, expertise, and human values around intrinsic motivation, well-being, and other aspects.

2.4.1 Workflows. Many of the changes driven by GenAI have been described as a shift to cognitive or metacognitive 'laziness' [39], yet such characterizations themselves require reflection. When does 'lazy' simply become 'resource reallocation'? Is cognitive offloading to GenAI instead better viewed as a form of 'cognitive miserliness' [120] or 'satisficing' [114]? Are there points in a workflow that benefit from cognitive offloading such that conserved cognitive energy can be deployed to other tasks?[3] Understanding *workflows*—complex, non-linear combinations of discrete tasks—may be one key to answering these questions. In programming, for example, whereas code generation and debugging may be increasingly solved by GenAI assistance, might design specification, refactoring, and code management need to remain essential human skills [89]? In design, using GenAI to rapidly generate many alternatives may facilitate ideation, while the task of sifting through, combining, and curating alternatives may remain a human endeavor [20]. Focusing on tasks at the expense of workflows risks missing interactions between tasks. For example, if GenAI is introduced too early in the ideation process, rather than augmenting ideation, it can stifle creativity by fixating users' attention onto specific ideas [70, 93]. Moving beyond discrete tasks, how can we better understand workflows, and GenAI's ability to reshape them?

2.4.2 The role of expertise. Interlinked with workflows is the role of people's *expertise*, another key factor in understanding the impact of AI on cognition. As Siu and Fok [116] summarize:

> *"experts develop sophisticated mental models and perceptual skills that enable them to maintain flexible cognitive frames throughout their sense-making process, allowing them to recognize meaningful patterns and adapt their understanding as new information emerges"*

Experts in a particular domain may therefore be better equipped to specify goals, decompose tasks, and ultimately prompt GenAI systems in that context. They are also better able to assess, interpret, and validate the varied outputs produced by GenAI [116]. As per §2.2, these differences are even observable among students, with more advanced students using GenAI in more effective ways [10, 75, 90].

Yet, although experts in domains such as research [116] and journalism [84] may be keen to offload routine, low-level tasks

---

[2]A distinction proposed by Heidegger [49], noted by Winograd and Flores [131], and echoed in the 'Extended Mind' thesis [19].
[3]Conserved energy may equally be deployed across workflows, or to non-work tasks entirely, such as leisure activities, a topic broaching the role of human values, as per §2.4.3.

to AI, they hesitate to offload higher-level analysis and decision-making to systems that they suspect lack the necessary nuance and context. Partly, this hesitation may be explained by the limitations in the accuracy of current GenAI systems, particularly for the kind of complex work in which experts engage [8, 53, 113]. Part of it may also lie in workflow compatibility. Experts often have existing workflows and styles that they want AI to augment, rather than override [7]. This raises questions such as: how can we design AI systems that enable selective delegation, preserve agency, and support verification in a manner that respects varying levels and types of expertise [116]?

A provocation here is whether experts are always better suited to judging when and what to offload to GenAI. One critical stance suggests that experts' extensive accumulation of experience can lead to cognitive entrenchment that misses alternative or innovative approaches [116]. Is optimal offloading to GenAI therefore an emerging kind of expertise, or meta-expertise, one that involves delegation, navigating multiple domains, and orchestrating tools across workflows [124]? AI-assisted coding offers a model: success here depends on task decomposition, evaluation, and preference, not just rote instruction. More broadly, this hints at the fact that expertise comes in different flavours: domain expertise differs from expertise with AI, which also differs from managerial expertise (i.e., the ability to define and delegate tasks to others). How do we better understand the types of expertise involved in working optimally with GenAI, including delegation, orchestration, and navigation across multiple domains and workflows? Acknowledging and supporting different types of expertise will be crucial for effective use of AI.

An equally important question is how the use of GenAI and its focus on efficiency may change experts' experience of their work, particularly reducing the space for 'deep work'—i.e., a distraction-free focus on cognitively demanding and meaningful tasks, aimed to achieve meaningful progress [83]. For example, as Nishal et al. [84] have found, journalists emphasize the importance of developing one's writerly voice and the freedom to explore ideas, both of which may be curtailed by an AI-driven focus on productivity. How do we protect and incentivize time for 'deep work' [83]? This raises broader considerations of motivation, well-being and other human values that sit alongside the impact on thinking and learning per se.

2.4.3 Motivation, well-being, and other human values. How will GenAI, including its capability to automate cognitive tasks, affect our well-being? Citing Self-Determination Theory, Langer [65] argues that ongoing human engagement in epistemic labor, i.e., actively acquiring, processing, and evaluating knowledge, is vital for sustaining intrinsic motivation and psychological well-being. Such epistemic labor fulfills the basic human needs for autonomy, competence, and relatedness, all of which may be threatened by excessive dependence on GenAI tools. Hence, the perceived sacredness of a writer's voice [84] or an artist's style and its centrality to professional and personal identity, and attendant societal phenomena that discourage AI use on these grounds [101]. How does using AI affect people's intrinsic motivation for tasks and other aspects of well-being [48, 107, 108]?

This ties into the notion of 'dialectical activities', pursuits such as parenting and art, that are valued for their intrinsic nature and sustained engagement [139]. In creative domains, where taste and intentionality matter, the perceived interchangeability of GenAI outputs can feel particularly demotivating. For instance, journalists may appreciate the assistive value of GenAI, yet strongly resist AI generating the core ideas of their work, a perceived degradation of journalism's artistic value [84]. Yet, individuals differ in which tasks they find intrinsically motivating. Some may approach their work like sculptors, directly shaping material, while others may see themselves as miners, sifting through material in search of valuable insights. GenAI's ability to produce material and reshape workflows carries distinct implications for each of these working styles which we need to understand [97].

Lastly, considerations pertinent to well-being such as emotional vulnerability, dependency, and mental health in the context of GenAI interactions were only briefly touched upon at the workshop, yet are emerging as important issues [1, 28, 40].

3 Augmenting human cognition with AI

By understanding the interplay between GenAI and human cognition, we can not only mitigate any potential negative impacts of GenAI, but also design systems that ultimately augment our thinking, leading to better cognitive and workflow outcomes in the short- and long- term.

As evidenced by the range of our workshop submissions, approaches to GenAI-driven cognitive augmentation have taken, and should take, many forms. One dimension along which to organize current approaches is the *directness* by which AI systems are designed to augment our cognition. Sitting at one extreme are designs intended to directly challenge our thinking through provocation [14, 33] or questioning [18, 23], discussed below in §3.1. Less directly, some systems may not explicitly challenge us, but rather provide formal structure or scaffolding for processes such as sense making [129] and ideation [135] (§3.2). Other approaches include using AI to transform the representation of information across modalities, fidelities, and other dimensions that may alter our thinking in subtle but potentially powerful ways (§3.3). Finally, some approaches harness emotional and motivational pathways, or other 'System 1' processes [37], to indirectly augment aspects such as learning, decision making, and creativity (§3.4).

A common thread across these approaches is their focus on what Zhang and Reicherts [140] refer to as *process-oriented* support, where AI assists users in identifying and addressing challenges throughout a task, ultimately enabling them to solve it themselves, rather than attempting to solve it for them (i.e., task automation, or 'end-to-end' support; see also [133]). A key distinction of process-oriented support is its focus on helping users reason *forward* towards their task solution—thereby supporting human understanding, verification, and agency—rather than reason backward from an AI-generated solution, which removes the user from the process (i.e., reducing situational awareness [113]) and increases the risk of over-reliance [96]. Khurana and Chilana [58] identify that a semi-automated (or 'process-oriented') approach—automating routine steps, while also guiding users through more complex decision-making—is more appropriate when users are less familiar with an

application or task, when a task involves complex decision making or creativity (rather than solely repetitive or straightforward steps), or when users have learning goals (rather than solely performance goals; see also [59]).

The next sections explore the different approaches to augmenting cognition. Cutting across these approaches are considerations about interacting and interfacing with GenAI (§3.5), as well as the overarching workflows that encompass discrete tasks (§3.6).

3.1 Provoking thinking through challenge

One form of cognitive augmentation applies metaphors for AI as an active entity that challenges us during tasks: provocateur [99, 105], antagonist [12], coach [50]; disrupting our default modes of thinking, and prompting reflection, exploration, and learning. In the design domain, for example, Gmeiner et al. [45] explore the use of metacognitive support agents that ask reflective questions and proactively support task planning and decomposition during the AI-assisted design process. They find that such interventions have the potential to support intent formulation, problem exploration, and outcome evaluation (see also [44, 46]). Castañeda et al. [14] take this approach one level higher: using AI to support meta-decision making (i.e., deciding how to decide), specifically helping users to iteratively define, refine, and prioritize their decision criteria. They propose a set of design goals for such assistance, instantiated in their prototype, including using concrete decision options or examples as reference points for users, systematically varying these options to help users discover decision dimensions of interest, and enabling iteration on all aspects of a decision.

Cheung [18] proposes another metaphor for AI-driven cognitive augmentation: *the ignorant co-learner*, a deliberately 'artificially ignorant' system that fosters 'moments of uncertainty, dissonance, or pause that compel users to think critically and reflexively'. For example, systems may offer 'multiple, conflicting perspectives or highlight areas of uncertainty', or prompts that call attention to the interpretive process in information consumption. While appealing in theory, all such approaches raise the question of how we can *motivate* people to engage with systems that deliberately question or otherwise slow them down. Indeed, Cheung argues that AI provocations 'must be clearly linked to pedagogical goals such as critical thinking, metacognition, and epistemic agency'. To this end, Ayyappan and Joyner [4] explore the value of explicitly clarifying the role of AI assistance during learning, allowing students to toggle between *'Tell me the answer'* and *'Ask me a question'* modes of AI support.

Yet GenAI is also being rapidly deployed across workplaces, where the primary aim is *productivity* rather than learning. In this context, how can systems demonstrate the value of 'extra' cognitive work to justify the time and effort taken, and thereby build intrinsic motivation to engage with them? Or, to question the doctrine of 'desirable difficulty' [9], must reflection or critical thinking always be *difficult* or otherwise high-friction to be meaningful? What is the appropriate level and timing of friction for a given task?

3.2 Structuring and supporting tasks

Not all cognitive augmentation is designed to challenge thinking or provide reflective friction. Indeed, several of the above examples combine reflective prompts or other provocations with scaffolding intended to guide users through task completion. Prime examples of this latter approach often harness the flexibility and pattern recognition abilities of large language models (LLMs) to identify and surface structure within data to support sense making (e.g., [22]). Wang and Chilton [129] develop a human-AI complementary workflow that augments human schema induction—the process of identifying patterns from examples, such as articles or videos, to support sense making and learning. AI clusters examples by similarities, extracts structural patterns within each cluster, and refines clusters by generating contrastive examples. Humans evaluate and refine the generated schemas in an iterative loop. Yang et al. [135] target human design ideation. Their prototype uses AI to first organize ideation stimuli into a functional hierarchy (i.e., a mechanism tree), and then generate analogical cues for these stimuli—e.g., how might a mechanism in context A be used in context B—to support users in identifying novel applications across domains (see also [56]). Finally, Schnizer and Mayer [109] offer a vision for a GenAI-driven system for data visualization workflows that assesses and adapts to users' expertise, and provides dynamic suggestions, guidance, and reflective prompts.

These explorations raise pertinent questions for understanding the impact on cognition. In contrast to the risk of AI systems deterring users via friction (§3.1), one key concern here is the risk of *over*-scaffolding tasks, with distinct implications for different cognitive processes. During sense-making, for example, which degree or type of cognitive effort is necessary for users to be helpfully close to the data and which can be bypassed? During the creative process, how do we protect the delicate, initial period of human idea incubation from the undue influence of AI? How does AI support during the divergence phase of ideation impact convergence? What is the appropriate balance of AI support between these two phases of ideation? Here, we can turn to the learning sciences, which define 'learning goals' as a principled way to designing appropriate scaffolding [92]. Alongside deliberate learning, how might such goals be defined in a productivity context?

Overall, there is an interesting design tension between approaches that induce cognitive friction and those that scaffold tasks—finding leverage points and 'sweet spots' in AI systems that get the best of both worlds is likely a fruitful area of future study.

3.3 Transforming information representations

Stepping outside of specific tasks, one thread of work explores how AI can be used to dynamically and interactively transform information between different representations, thus contributing to a known form of cognitive enhancement [122]. Liu and Li [71] explore how transforming content across modalities—e.g., from speech to text, or from text to visuals—can augment cognitive workflows. For example, how might flexibly translating between textual and graphical modalities enable the manipulation of concepts, rather than words or pixels? Their work raises key research

questions for this space, such as how to best visualize the evolving semantic space of spoken content, or how to balance semantic abstraction with familiarity to support cognition.

Rather than modalities, Yen et al. [136] explore using AI to transform information between levels of *formality*. In particular, they explore an approach to semi-formal programming—the integration of formal representations (e.g., defined variables in a piece of code) with informal ones (e.g., sketches). They propose three principles that work together: loosening strictly typed programs (e.g., dynamically converting unknown variables into placeholders), gradually enriching informal representations as meaning emerges through user interaction and context, and, rather than prespecifying everything in advance, using AI to dynamically interpret or query fuzzy attributes as needed.

Echoing across these projects is a model of human-AI partnership proposed by Suh [121] in which AI expands humans' multidimensional conceptual environments (spanning knowledge, representations, design possibilities, among other aspects); humans further expand these environments via creativity, critical thinking, and contextual understanding; and together, they explore these environments via interfaces and interactions that thoughtfully support human-AI collaboration.

GenAI has surfaced many questions about the malleability of text—its ability to be flexibly summarized, expanded, and re-written on demand—and the implications for information processing, knowledge, and understanding. What are the affordances of text and other media as malleable projections of knowledge? How can systems that consume and render text support users in manipulating it to suit their cognitive goals and preferences? Beyond text, how do we effectively represent information in different output modalities to augment cognition?

3.4 Augmenting emotions and other 'System 1' processes

An implicit assumption underlying a lot of work on GenAI-driven cognitive augmentation is the prioritization of conscious reflection, deliberate reasoning, and other processes commonly associated with 'System 2' thinking [38]. Can (or should) we augment people's 'System 1' processes—e.g., emotions, heuristics—to progress users towards desired cognitive outcomes [76]? This may be welcomed, and indeed feel natural, in the creative domain. For example, Pilcher and Tütüncü [88] propose an approach to fostering creativity by welcoming and amplifying GenAI's inaccuracies or fabrications ('hallucinations'). Their mixed-reality prototype envelopes users in surreal and occasionally unsettling multi-modal outputs to provoke 'surprise, reflection, and creative risk-taking'.

What would augmenting 'System 1' processes look like in other tasks? In learning, for example, Leong [67]'s work demonstrates how AI can help change users' perspectives through emotional and motivational pathways. For instance, AI can personalize experiences to users' interests, such as examples in a vocabulary learning app [68], or the identity of a virtual instructor [87], thereby increasing their motivation. Similarly, by (privately) applying augmented reality filters to one's audience, it is possible to reframe the situation and thereby reduce public speaking anxiety [69]. More provocative

interventions can be imagined, such as gamifying cognitive effort, or leveraging the common feeling of being 'unsettled' during decision making to provide timely provocations. Ethical implications emerge: for example, where does human agency fit in, and what is the responsibility of systems, particularly those that deliberately induce uncertainty, to provide emotional support to users? More fundamentally, what is the scope for GenAI-driven training, real-time support, or other approaches to augment 'System 1' cognition, e.g., intuition [11, 15], as an end in itself?

3.5 Interacting and interfacing with AI

Behind the various approaches to augmenting cognition are considerations about how we interact and interface with AI. Although turn-based conversational prompting has been a dominant interaction pattern for LLMs, other paradigms are emerging. Reviewing recent work, Zindulka et al. [142] argue that direct manipulation—prompting by manipulating relevant on-screen objects—can ease cognitive load associated with prompt formulation, more accurately express intent, and reduce the workflow disruptions that occur when navigating between content and the prompt interface. In turn, they argue that direct manipulation has the potential to free cognitive resources for higher-level thinking, and support the continuous expression of users' thought processes without being overwhelmed by AI output.

Other work has explored the dynamic between prompting and evaluating AI outputs. Min and Xia [78] propose 'feedforward' information about the prospective AI output as a fundamental design component of GenAI systems. Alongside foreshadowing AI output, effective feedforward should enable users to efficiently disambiguate and reflect on prompts, and engage with feedforward content as a form of prompting. Through exploratory work, Min and Xia [78] identify three initial dimensions for feedforward design: representation (e.g., topic lists), level of detail, and manipulability (e.g., selecting feedforward elements). These can be viewed as a form of 'dynamic prompt middleware' [35] for improving critical reflection during the prompting process.

Feedforward is one way to speed up the interactive loop between humans and AI. Liu et al. [74] approach this from a more radical angle, proposing a paradigm in which systems continuously generate internal and partly intrinsically-driven responses ('thoughts'), and selectively yet proactively communicate those responses to users throughout an interaction: full-duplex communication, in which proactive AI responses are gated by human heuristics for contributing to conversations. An instantiation of this paradigm shows promise in conversational engagement, and offers the possibility of a shared cognitive space in which both human and AI can build on fragments of ideas [73].

Other interaction modalities are also being explored. For example, the recent rapid improvement in speech recognition and generation has led to increasing adoption of Voice User Interfaces [112]. Speech interaction may encourage more externalization of thoughts [95, 138], whereas nonverbal aspects of speech like tone, loudness, and pauses can convey complementary information (e.g., emotional state [112]). What are other approaches to expressing human intent that have yet to be explored?

How might different interaction approaches augment users' expression of intent, and might there be unexpected knock-on effects on cognition? For example, Zindulka et al. [142] surmise that direct manipulation could also *negatively* impact cognition. Given that writing *is* thinking [42], the replacement of text prompting with direct manipulation may remove opportunities for users to think through their intents and decompose their tasks (see also [141]). Moreover, the blending of user- and AI- generated content in the interface may encourage analogous blending in users' minds, including any potential biases of LLMs [130]. Likewise, how does the anthropomimetic slant of approaches like those in Liu et al. [74] trade off in terms of improved conversational flow, users' mental models, over-reliance, emotional impact, and other dimensions? Future work should map the cognitive affordances of different modalities, including how they might express or transform knowledge and other internal states, and how such information might be represented back to users to augment cognition.

3.6 Augmenting workflows

High-level intents rarely remain confined to discrete tasks. Rather, they ladder up to workflows (as per §2.4.1). To truly augment cognition therefore involves thinking beyond the task to consider the broader workflows which encompass them [86, 124, 125]. Workflows increasingly interleave manual and AI-assisted processes, human collaborators, and AI agents that coordinate to complete tasks. GenAI's flexibility and scale at handling context invite us to re-imagine workflows from first principles.

One possibility is a move away from application-centric computing, where workflows are fragmented across siloed applications, to *activity-centric computing* [6], where application boundaries are dissolved in favor of coherently supporting users' broader workflow goals and activities. To this end, Xia [132] proposes a new AI-powered paradigm in which interfaces are dynamically generated to suit a user's task context, needs, and preferences (see also [13]). Their approach uses AI to infer the underlying task-driven models from users' prompts, representing the information entities, relationships, and data within information tasks. These task-driven models map to UI specifications, which are then used to dynamically generate interfaces suitable to users' tasks. Such approaches raise implications for cognition: for example, how can systems accurately determine users' higher- and lower- level goals across workflows, while minimizing the cognitive burden on users to correct inaccurate goal inferences and manually configure interfaces?

Workflows extend not only across tasks, but also across people. The modern workplace is characterized by frequent collaboration across large networks. Yet, despite our inherently social nature, achieving efficient collaboration remains an ongoing challenge [111, 134]. How can AI support more effective communication and coordination of tasks? Mukhopadhyay and Luther [80] explore how AI could augment leadership behaviours for collaborative project development (see also [79]). Although their conversational AI system could support aspects of planning (e.g., by generating templates) or problem-solving (e.g., by providing feedback), they found that its impact "remained constrained to specific tasks rather than seamlessly integrating into the overall workflow". Indeed, like most modern collaborative work, participants' communication and

workflows were distributed across interfaces which limited the AI system's reach and contextual understanding [137]. Integrating AI support into realistic workflows is an open challenge, touching upon interaction design questions such as alternatives to the chat-based paradigm (§3.5), and the role of proactive AI assistance [16, 17, 27, 52, 73, 91, 128].

Not only does AI have potential to assist collaborative workflows, Johnson and Rick [55] propose the inverse: collaborative, multi-user interactions with AI can mitigate over-reliance observed with individual interactions. They posit that group interactions can encourage active engagement with AI assistance, for example, because the responsibility of critically evaluating AI contributions (and the impact of doing so) is distributed across members (e.g., see [25]). Groups' tendency for collaborative boundary regulation—the "process of establishing limits and ground rules to maintain a healthy group environment"—can similarly regulate the influence of AI on group processes and outputs. Yet, AI may also plausibly exacerbate unproductive group dynamics like social loafing and group-think. Designing collaborative AI-assisted workflows to harness the benefits of both group work and AI assistance is an open challenge [54]. For example, which aspects of collaborative processes can be offloaded to AI and when? How can we detect and mitigate inappropriate reliance in group settings? How can we flexibly adapt AI's role in a group setting to suit different needs, and to shift between individual and group support?

The rise of "agentic" AI—systems that attempt to execute workflows autonomously—raises further questions for collaborative work. If humans will increasingly hand over the production of code [98, 102], writing [21], and other tasks to autonomous agents, then how will human roles change? Will roles shift to management, including the definition of goals, outputs, guardrails, and verification procedures? How can people preserve and grow the expertise needed to assess the correctness of a workflow? As AI agents begin to work across people and teams, their management, including 'context engineering' [77] and verification, will only increase in complexity. What will AI-augmented workflows look like in this context? How can AI support humans in the cognitive management of these complex workflows?

4 Formative research, theory, measurement, and evaluation

Understanding the impact of GenAI and designing systems that both protect and augment cognition will require formative research on GenAI users, workflows, and expertise, as well as a coherent approach and toolkit for running rigorous research studies, including theories, construct definitions and metrics [127], tasks, and protocols. Establishing a shared foundation and methodological approach will enable cross-study comparability, cumulative progress, and more intentional design of GenAI systems that support rather than undermine human thought.

4.1 Formative research: Users, workflows, and expertise

We need a deeper understanding of distinct sets of GenAI tool users, their workflows, and how their expertise—in their domains, in GenAI, and in the management of the two (§2.4.2)—is instantiated

within them. How are cognitive tasks and behaviors structured and sequenced? This is necessary to understand how GenAI changes or eliminates many of the micro-decisions people make in their workflows. To this end, Nishal et al. [84] explore the writing workflows of science journalists, whereas Siu and Fok [116] describe the research survey practices of CS researchers, and the document review workflows of knowledge workers. Beyond work, Kim et al. [60] explore how 'heavy' LLM users employ these tools for decision-making in their daily lives, defining such users as those who use LLMs for a substantial (and specifically defined) subset of concrete tasks that cover both work and personal domains, and spanning the range from purely rational to purely intuitive decision making, according to Cognitive Continuum Theory. Alongside the interview and survey methods deployed in these studies, Schnizer and Mayer [109] propose that expertise can also be inferred via complementary approaches such as user interaction patterns, like prompting or eye-tracking. Without a clear understanding of who the 'users' are, any inferences and generalizations we make about the impact of AI will remain too vague or too muddled to be useful. Equally important, such understanding unlocks the ability to design systems tailored to users' expertise levels and needs [109].

4.2 Theories and constructs

What do we specifically mean by terms like "critical thinking", "creativity", "agency", "learning"? Emerging research has drawn on a range of existing theories, models, and constructs from across disciplines: Bloom's Taxonomy [63], Dewey's theory of reflective thinking [29], Schön's model of reflection [110], Sawyer's model of creativity [106], Hammond's Cognitive Continuum Theory [47], dual-process theories of cognition [38], theories of metacognition [124], and pragmatism [31], among others. How far can these theories and associated constructs carry us today? Which theories and constructs are we missing? Do we need new or updated theories? For example, Felten [41] argues that further theory development is needed to understand how human and AI biases can either amplify or diminish each other, i.e., moving away from separate treatment of these sources of bias and instead considering *compound human-AI bias.* This interactionist perspective can help us develop strategies to mitigate the negative impacts of biases on human cognition.

Slovak [118] offers a concrete avenue for how theories can be used effectively. If GenAI systems are viewed as psychological interventions, then theories can inform *theories-of-change* [26] for these interventions, which prescribe "a particular set of experience trajectories that are expected to lead to the psychological effects for the participant(s)" (see also [119]). If theories and the resulting theories-of-change can prescribe (or describe) in detail what a user *should* or *would* experience when interacting with an AI system, then it becomes easier to translate these theories into design briefs that inform system design for cognitive augmentation, or, for existing systems, help uncover their positive or negative impact.

4.3 Measurement and evaluation

Beyond defining constructs, how do we reliably measure them? Here, the disciplines with which HCI commonly intersects—cognitive,

learning, information, and organizational sciences, and others—will become even more vital as sources of foundational research into cognition [124, 127].

Yet, other disciplines may be insufficient on their own, as GenAI introduces new challenges and opportunities for evaluation. For example, human-AI interactions increasingly involve qualitative, open-ended tasks that may lack ground truth and often require iterative interactions, whereas much research in cognitive science (and human-AI interaction) has relied on simple, closed-ended tasks for experimental control. Therefore, there is a need for new methods to measure these interactions and task outcomes effectively. Prompting offers one such window into cognition, including users' level of expertise, and their cognitive strategies, such as task decomposition [57] and abstraction [72, 104], as well as their biases or gaps in understanding. These can be studied through user-centric approaches such as participatory prompting [34, 103]. They can also be approached in an automated manner, leveraging GenAI's natural language capabilities. For example, Holstein et al. [51] present a promising novel framework and automated pipeline for analyzing prompting patterns along two dimensions: cognitive activity mode (exploration vs. exploitation) and cognitive engagement mode (constructive vs. detrimental).

Another challenge is that many constructs central to this space, like critical thinking and agency, are inherently introspective and subjective. Collecting reliable self-report data on them requires research participants to have self-awareness, and even well-defined theoretical constructs may not align with people's self-understanding. How do we measure something like "critical thinking" if people differ in their views of the term and of their own thinking? Despite the challenges, a turn to theory here can provide a starting point. For example, Dewey's theory of reflective thought defines it as "definite units that are linked together so that there is a sustained movement to a common end" [30]. This quality of 'traceability' offers an intriguing avenue to characterising thinking with and without GenAI assistance. Similarly, Dervin's sense-making theory [82], including the concept of cognitive gaps and bridges, offers another promising direction. These conceptualizations can inform the elicitation and analysis of people's varied experiences [15], which can now be conducted at scale using GenAI [117]. Ultimately, although the misalignment between theoretical definitions and subjective introspection may remain unresolved, it nevertheless deserves consideration in future work.

Finally, many potential changes in workflows and impacts on cognition may take time to emerge as users adopt tools and adapt their practices. How does prolonged use of GenAI affect users' ability to reason independently, sustain attention, or engage in reflective thinking? Longitudinal studies are needed to track and understand potential changes.

5 Conclusion

GenAI poses a myriad of risks and opportunities for human cognition that present a large research and design space to address (and one that cannot be comprehensively addressed in one workshop). Building on the breadth of our workshop, we encourage future workshops and initiatives to take a focused approach, targeting specific areas to make concerted progress. We also encourage the HCI

community to further expand its awareness of, and collaboration with, other communities and disciplines so as to share insights and accelerate progress. Finally, we emphasize that ongoing developments in task automation should not distract from the protection of core aspects of human cognition, and equally importantly, the innovation of approaches to augment and transform human cognition for the better.

Acknowledgments

References

[1] David Adam. 2025. Supportive? Addictive? Abusive? How AI companions affect our mental health. *Nature* 641, 8062 (May 2025), 296–298. doi:10.1038/d41586-025-01349-9 Bandiera_abtest: a Cg_type: News Feature Publisher: Nature Publishing Group Subject_term: Society, Software, Computer science, Public health, Psychology.

[2] Dhruv Agarwal, Mor Naaman, and Aditya Vashistha. 2025. AI Suggestions Homogenize Writing Toward Western Styles and Diminish Cultural Nuances. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–21. doi:10.1145/3706598.3713564

[3] Barrett R Anderson, Jash Hemant Shah, and Max Kreminski. 2024. Homogenization Effects of Large Language Models on Human Creative Ideation. In *Proceedings of the 16th Conference on Creativity & Cognition (C&C '24)*. Association for Computing Machinery, New York, NY, USA, 413–425. doi:10.1145/3635636.3656204

[4] Dinesh Ayyappan and David A. Joyner. 2025. Building Self-Reliance in an Era of Accessible Intelligent Tutors. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI*. Yokohama, Japan.

[5] Mia Magdalena Bangerl, Leonie Disch, Tamara David, and Viktoria Pammer-Schindler. 2025. CreAltive Collaboration? Users' Misjudgment of AI-Creativity Affects Their Collaborative Performance. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–17. doi:10.1145/3706598.3713886

[6] Jakob E. Bardram, Steven Jeuris, Paolo Tell, Steven Houben, and Stephen Voida. 2019. Activity-centric computing systems. *Commun. ACM* 62, 8 (July 2019), 72–81. doi:10.1145/3325901

[7] Gordon Baxter, John Rooksby, Yuanzhi Wang, and Ali Khajeh-Hosseini. 2012. The ironies of automation: still going strong at 30?. In *Proceedings of the 30th European Conference on Cognitive Ergonomics*. ACM, Edinburgh United Kingdom, 65–71. doi:10.1145/2448136.2448149

[8] Joel Becker, Nate Rush, Elizabeth Barnes, and David Rein. 2025. Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity. doi:10.48550/arXiv.2507.09089 arXiv:2507.09089 [cs].

[9] Robert A Bjork. 1994. Institutional impediments to effective training. *Learning, remembering, believing: Enhancing human performance* (1994), 295–306.

[10] Jessica Y. Bo, Majeed Kazemitabaar, Emma Zhuang, and Ashton Anderson. 2025. Who's the Leader? Analyzing Novice Workflows in LLM-Assisted Debugging of Machine Learning Code. doi:10.48550/arXiv.2505.08063 arXiv:2505.08063 [cs].

[11] Esther Boissin, Serge Caparos, Matthieu Raoelison, and Wim De Neys. 2021. From bias to sound intuiting: Boosting correct intuitive reasoning. *Cognition* 211 (June 2021), 104645. doi:10.1016/j.cognition.2021.104645

[12] Alice Cai, Ian Arawjo, and Elena L. Glassman. 2024. Antagonistic AI. doi:10.48550/arXiv.2402.07350 arXiv:2402.07350 [cs] version: 1.

[13] Yining Cao, Peiling Jiang, and Haijun Xia. 2025. Generative and Malleable User Interfaces with Generative and Evolving Task-Driven Data Model. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–20. doi:10.1145/3706598.3713285

[14] Chance Castañeda, Jessica Mindel, Will Page, Hayden Stec, Manqing Yu, and Kenneth Holstein. 2025. Supporting AI-Augmented Meta-Decision Making with InDecision. doi:10.48550/arXiv.2504.12433 arXiv:2504.12433 [cs].

[15] 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 (Sept. 2023), 1–32. doi:10.1145/3610219

[16] Valerie Chen, Alan Zhu, Sebastian Zhao, Hussein Mozannar, David Sontag, and Ameet Talwalkar. 2025. Need Help? Designing Proactive AI Assistants for Programming. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–18. doi:10.1145/3706598.3714002

[17] Xinyue Chen, Lev Tankelevitch, Rishi Vanukuru, Ava Elizabeth Scott, Payod Panda, and Sean Rintel. 2025. Are We On Track? AI-Assisted Active and Passive Goal Reflection During Meetings. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–22. doi:10.1145/3706598.3714052

[18] Manni Cheung. 2025. From Answer Machines to Ignorant Co-Learners: Designing AI to Augment Rather than Replace Human Thinking. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI*. Yokohama, Japan.

[19] Andy Clark and David Chalmers. 1998. The Extended Mind. *Analysis* 58, 1 (Jan. 1998), 7–19. doi:10.1093/analys/58.1.7

[20] Peter Dalsgaard. 2025. Tools for Creative Cognition: Generative AI in Design Thinking. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI*. Yokohama, Japan. https://github.com/ai-tools-for-thought/workshop/blob/5a2aadd00942f89fc1b1c97fbbedd7ed4e0d92c7/documents/Dalsgaard%20-%202025%20-%20Generative_AI_as_a_Tool_for_Designerly_Thinking.pdf

[21] Hai Dang, Sven Goller, Florian Lehmann, and Daniel Buschek. 2023. Choice Over Control: How Users Write with Large Language Models using Diegetic and Non-Diegetic Prompting. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–17. doi:10.1145/3544548.3580969

[22] Hai Dang, Chelse Swoopes, Daniel Buschek, and Elena L. Glassman. 2025. CorpusStudio: Surfacing Emergent Patterns In A Corpus Of Prior Work While Writing. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–19. doi:10.1145/3706598.3713974

[23] Valdemar Danry, Pat Pataranutaporn, Yaoli Mao, and Pattie Maes. 2023. Don't Just Tell Me, Ask Me: AI Systems that Intelligently Frame Explanations as Questions Improve Human Logical Discernment Accuracy over Causal AI explanations. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3544548.3580672

[24] Valdemar Danry, vdanry@mit.edu, https://orcid.org/0000-0001-5225-0077, View Profile, Pat Pataranutaporn, patpat@mit.edu, https://orcid.org/0000-0002-1879-7340, View Profile, Matthew Groh, matthew.groh@kellogg.northwestern.edu, https://orcid.org/0000-0002-9029-0157, View Profile, Ziv Epstein, zive@stanford.edu, https://orcid.org/0000-0002-5831-5756, and View Profile. 2025. Deceptive Explanations by Large Language Models Lead People to Change their Beliefs About Misinformation More Often than Honest

Explanations. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–31. doi:10.1145/3706598.3713408

[25] Regina de Brito Duarte, Mónica Costa Abreu, Joana Campos, and Ana Paiva. 2025. The Amplifying Effect of Explainability in AI-assisted Decision-making in Groups. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–15. doi:10.1145/3706598.3713534

[26] Mary J. De Silva, Erica Breuer, Lucy Lee, Laura Asher, Neerja Chowdhary, Crick Lund, and Vikram Patel. 2014. Theory of Change: a theory-driven approach to enhance the Medical Research Council's framework for complex interventions. *Trials* 15, 1 (July 2014), 267. doi:10.1186/1745-6215-15-267

[27] Yang Deng, Lizi Liao, Zhonghua Zheng, Grace Hui Yang, and Tat-Seng Chua. 2024. Towards Human-centered Proactive Conversational Agents. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, Washington DC USA, 807–818. doi:10.1145/3626772.3657843

[28] Alicia DeVrio, Myra Cheng, Lisa Egede, Alexandra Olteanu, and Su Lin Blodgett. 2025. A Taxonomy of Linguistic Expressions That Contribute To Anthropomorphism of Language Technologies. doi:10.1145/3706598.3714038 arXiv:2502.09870 [cs].

[29] John Dewey. 1910. *How we think*. D C Heath, Lexington, MA, US. doi:10.1037/10903-000 Pages: vi, 228.

[30] J. Dewey. 1933. *How we think*. Heath, Oxford, England. Pages: x, 301.

[31] John Dewey. 1938. *Logic: The Theory of Inquiry*. Henry Holt, New York, NY, USA.

[32] Anil R. Doshi and Oliver P. Hauser. 2024. Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances* 10, 28 (July 2024), eadn5290. doi:10.1126/sciadv.adn5290 Publisher: American Association for the Advancement of Science.

[33] Ian Drosos, Advait Sarkar, Xiaotong, Xu, and Neil Toronto. 2025. "It makes you think": Provocations Help Restore Critical Thinking to AI-Assisted Knowledge Work. doi:10.48550/arXiv.2501.17247 arXiv:2501.17247 [cs].

[34] Ian Drosos, Advait Sarkar, Xiaotong Xu, Carina Negreanu, Sean Rintel, and Lev Tankelevitch. 2024. "It's like a rubber duck that talks back": Understanding Generative AI-Assisted Data Analysis Workflows through a Participatory Prompting Study. In *Proceedings of the 3rd Annual Meeting of the Symposium on Human-Computer Interaction for Work* (Newcastle upon Tyne, United Kingdom) *(CHIWORK '24)*. Association for Computing Machinery, New York, NY, USA, Article 16, 21 pages. doi:10.1145/3663384.3663389

[35] Ian Drosos, Jack Williams, Advait Sarkar, Nicholas Wilson, Sean Rintel, and Payod Panda. 2025. Dynamic Prompt Middleware: Contextual Prompt Refinement Controls for Comprehension Tasks. In *Proceedings of the 4th Annual Symposium on Human-Computer Interaction for Work (CHIWORK '25)*. Association for Computing Machinery, New York, NY, USA, Article 24, 23 pages. doi:10.1145/3729176.3729203

[36] Douglas C. Engelbart. 1962. *Augmenting Human Intellect: A Conceptual Framework*. Technical Report AFOSR-3223. Stanford Research Institute, Menlo Park, CA. https://www.dougengelbart.org/content/view/138/000/ Prepared for the Air Force Office of Scientific Research, Stanford Research Institute.

[37] Jonathan St B. T. Evans. 2017. Dual Process Theory: Perspectives and problems. In *Dual Process Theory 2.0*. Routledge. Num Pages: 19.

[38] Jonathan St. B. T. Evans and Keith E. Stanovich. 2013. Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science* 8, 3 (May 2013), 223–241. doi:10.1177/1745691612460685 Publisher: SAGE Publications Inc.

[39] Yizhou Fan, Luzhen Tang, Huixiao Le, Kejie Shen, Shufang Tan, Yueying Zhao, Yuan Shen, Xinyu Li, and Dragan Gašević. 2025. Beware of Metacognitive Laziness: Effects of Generative Artificial Intelligence on Learning Motivation, Processes, and Performance. *British Journal of Educational Technology* 56, 2 (March 2025), 489–530. doi:10.1111/bjet.13544 arXiv:2412.09315 [cs].

[40] Cathy Mengying Fang, Auren R. Liu, Valdemar Danry, Eunhae Lee, Samantha W. T. Chan, Pat Pataranutaporn, Pattie Maes, Jason Phang, Michael Lampe, Lama Ahmad, and Sandhini Agarwal. 2025. How AI and Human Behaviors Shape Psychosocial Effects of Chatbot Use: A Longitudinal Randomized Controlled Study. doi:10.48550/arXiv.2503.17473 arXiv:2503.17473 [cs].

[41] Nick von Felten. 2025. Beyond Isolation: Towards an Interactionist Perspective on Human Cognitive Bias and AI Bias. doi:10.48550/arXiv.2504.18759 arXiv:2504.18759 [cs].

[42] Linda Flower and John R. Hayes. 1981. A Cognitive Process Theory of Writing. *College Composition and Communication* 32, 4 (1981), 365–387. doi:10.2307/356600 Publisher: National Council of Teachers of English.

[43] Yue Fu and Alexis Hiniker. 2025. Supporting Students' Reading and Cognition with AI. doi:10.48550/arXiv.2504.13900 arXiv:2504.13900 [cs].

[44] Frederic Gmeiner, Kaitao Luo, Ye Wang, Kenneth Holstein, and Nikolas Martelaro. 2025. Exploring the Potential of Metacognitive Support Agents for Human-AI Co-Creation. In *Proceedings of the 2025 ACM Designing Interactive Systems Conference (DIS '25)*. Association for Computing Machinery, New York, NY, USA, 1244–1269. doi:10.1145/3715336.3735785

[45] Frederic Gmeiner, Nikolas Martelaro, and Kenneth Holstein. 2025. Designing Metacognitive Support Interactions to Augment People's Thinking in Complex (Co-)Creative Tasks. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI*. Yokohama, Japan.

[46] Frederic Gmeiner, Humphrey Yang, Lining Yao, Kenneth Holstein, and Nikolas Martelaro. 2023. Exploring Challenges and Opportunities to Support Designers in Learning to Co-create with AI-based Manufacturing Design Tools. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–20. doi:10.1145/3544548.3580999

[47] Kenneth R Hammond. 2000. *Human judgment and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice*. Oxford University Press.

[48] Jessica He, Stephanie Houde, and Justin D. Weisz. 2025. Which Contributions Deserve Credit? Perceptions of Attribution in Human-AI Co-Creation. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–18. doi:10.1145/3706598.3713522

[49] Martin Heidegger. 1927. *Being and Time*. HarperCollins, New York, NY, USA.

[50] Jake M. Hofman, Daniel G. Goldstein, and David M. Rothschild. 2023. A Sports Analogy for Understanding Different Ways to Use AI. *Harvard Business Review* (Dec. 2023). https://hbr.org/2023/12/a-sports-analogy-for-understanding-different-ways-to-use-ai

[51] Joshua Holstein, Moritz Diener, and Philipp Spitzer. 2025. From Consumption to Collaboration: Measuring Interaction Patterns to Augment Human Cognition in Open-Ended Tasks. doi:10.48550/arXiv.2504.02780 arXiv:2504.02780 [cs].

[52] Mo Houtti, Moyan Zhou, Loren Terveen, and Stevie Chancellor. 2025. Observe, Ask, Intervene: Designing AI Agents for More Inclusive Meetings. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–18. doi:10.1145/3706598.3713838

[53] Sonia Jaffe, Neha Parikh Shah, Jenna Butler, Alex Farach, Alexia Cambon, Brent Hecht, Michael Schwarz, and Jaime Teevan. 2024. *Generative AI in Real-World Workplaces*. Technical Report. https://www.microsoft.com/en-us/research/publication/generative-ai-in-real-world-workplaces/

[54] Janet G. Johnson, Macarena Peralta, Mansanjam Kaur, Ruijie Sophia Huang, Sheng Zhao, Ruijia Guan, Shwetha Rajaram, and Michael Nebeling. 2025. Exploring Collaborative GenAI Agents in Synchronous Group Settings: Eliciting Team Perceptions and Design Considerations for the Future of Work. doi:10.48550/arXiv.2504.14779 arXiv:2504.14779 [cs].

[55] Janet G. Johnson and Steven R. Rick. 2025. The Promise and Peril of Collaboration: Fostering Appropriate Reliance When Problem-Solving with GenAI. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI*. Yokohama, Japan.

[56] Hyeonsu B Kang, David Chuan-En Lin, Yan-Ying Chen, Matthew K. Hong, Nikolas Martelaro, and Aniket Kittur. 2025. BioSpark: Beyond Analogical Inspiration to LLM-augmented Transfer. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–29. doi:10.1145/3706598.3714053

[57] Majeed Kazemitabaar, Jack Williams, Ian Drosos, Tovi Grossman, Austin Zachary Henley, Carina Negreanu, and Advait Sarkar. 2024. Improving Steering and Verification in AI-Assisted Data Analysis with Interactive Task Decomposition. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology* (Pittsburgh, PA, USA) *(UIST '24)*. Association for Computing Machinery, New York, NY, USA, Article 92, 19 pages. doi:10.1145/3654777.3676345

[58] Anjali Khurana and Parmit K. Chilana. 2025. Designing Semi-Automated Copilots: Balancing User Control and Guidance for Effective Human-Agent Collaboration. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI*. Yokohama, Japan.

[59] Anjali Khurana, Xiaotian Su, April Yi Wang, and Parmit K Chilana. 2025. Do It For Me vs. Do It With Me: Investigating User Perceptions of Different Paradigms of Automation in Copilots for Feature-Rich Software. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–18. doi:10.1145/3706598.3713431

[60] Eunhye Kim, Kiroong Choe, Minju Yoo, Sadat Shams Chowdhury, and Jinwook Seo. 2025. Beyond Tools: Understanding How Heavy Users Integrate LLMs into Everyday Tasks and Decision-Making. doi:10.48550/arXiv.2502.15395 arXiv:2502.15395 [cs].

[61] Sunnie S. Y. Kim, Q. Vera Liao, Mihaela Vorvoreanu, Stephanie Ballard, and Jennifer Wortman Vaughan. 2024. "I'm Not Sure, But...": Examining the Impact of Large Language Models' Uncertainty Expression on User Reliance and Trust. doi:10.1145/3630106.3658941 arXiv:2405.00623 [cs] TLDR: The findings suggest that using natural language expressions of uncertainty may be an effective approach for reducing overreliance on LLMs, but that the precise language used

matters, and highlights the importance of user testing before deploying LLMs at scale..

[62] Sunnie S. Y. Kim, Jennifer Wortman Vaughan, Q. Vera Liao, Tania Lombrozo, and Olga Russakovsky. 2025. Fostering Appropriate Reliance on Large Language Models: The Role of Explanations, Sources, and Inconsistencies. doi:10.1145/3706598.3714020 arXiv:2502.08554 [cs].

[63] David R. Krathwohl. 2002. A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice* 41, 4 (Nov. 2002), 212–218. doi:10.1207/s15430421tip4104_2 Publisher: Routledge _eprint: https://doi.org/10.1207/s15430421tip4104_2.

[64] Harsh Kumar, Jonathan Vincentius, Ewan Jordan, and Ashton Anderson. 2025. Human Creativity in the Age of LLMs: Randomized Experiments on Divergent and Convergent Thinking. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–18. doi:10.1145/3706598.3714198

[65] Markus Langer. 2025. Why Should We Invest Epistemic Labor in a World of Generative AI?. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI.* Yokohama, Japan.

[66] Hao-Ping (Hank) Lee, Advait Sarkar, Lev Tankelevitch, Ian Drosos, Sean Rintel, Richard Banks, and Nicholas Wilson. 2025. The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers. https://www.microsoft.com/en-us/research/publication/the-impact-of-generative-ai-on-critical-thinking-self-reported-reductions-in-cognitive-effort-and-confidence-effects-from-a-survey-of-knowledge-workers/

[67] Joanne Leong. 2025. Designing Transformative Lenses to Aid Thinking. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI.* Yokohama, Japan.

[68] Joanne Leong, Pat Pataranutaporn, Valdemar Danry, Florian Perteneder, Yaoli Mao, and Pattie Maes. 2024. Putting Things into Context: Generative AI-Enabled Context Personalization for Vocabulary Learning Improves Learning Motivation. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–15. doi:10.1145/3613904.3642393

[69] Joanne Leong, Florian Perteneder, Muhender Raj Rajvee, and Pattie Maes. 2023. "Picture the Audience...": Exploring Private AR Face Filters for Online Public Speaking. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3544548.3581039

[70] David Chuan-En Lin, Hyeonsu B. Kang, Nikolas Martelaro, Aniket Kittur, Yan-Ying Chen, and Matthew K. Hong. 2025. Inkspire: Supporting Design Exploration with Generative AI through Analogical Sketching. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–18. doi:10.1145/3706598.3713397

[71] Can Liu and Wengxi Li. 2025. Bridging the Cognitive Gap: Cross-Modality Perception in AI-Generated Content. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI.* Yokohama, Japan.

[72] Michael Xieyang Liu, Advait Sarkar, Carina Negreanu, Benjamin Zorn, Jack Williams, Neil Toronto, and Andrew D. Gordon. 2023. "What It Wants Me To Say": Bridging the Abstraction Gap Between End-User Programmers and Code-Generating 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 598, 31 pages. doi:10.1145/3544548.3580817

[73] Xingyu Bruce Liu, Shitao Fang, Weiyan Shi, Chien-Sheng Wu, Takeo Igarashi, and Xiang 'Anthony' Chen. 2025. Proactive Conversational Agents with Inner Thoughts. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–19. doi:10.1145/3706598.3713760

[74] Xingyu Bruce Liu, Haijun Xia, and Xiang Anthony Chen. 2025. Interacting with Thoughtful AI. doi:10.48550/arXiv.2502.18676 arXiv:2502.18676 [cs].

[75] Lauren E. Margulieux, James Prather, Brent N. Reeves, Brett A. Becker, Gozde Cetin Uzun, Dastyni Loksa, Juho Leinonen, and Paul Denny. 2024. Self-Regulation, Self-Efficacy, and Fear of Failure Interactions with How Novices Use LLMs to Solve Programming Problems. In *Proceedings of the 2024 on Innovation and Technology in Computer Science Education V. 1 (ITiCSE 2024)*. Association for Computing Machinery, New York, NY, USA, 276–282. doi:10.1145/3649217.3653621

[76] Andy Matuschak and Michael Nielsen. 2019. How can we develop transformative tools for thought? https://numinous.productions/ttft

[77] Lingrui Mei, Jiayu Yao, Yuyao Ge, Yiwei Wang, Baolong Bi, Yujun Cai, Jiazhi Liu, Mingyu Li, Zhong-Zhi Li, Duzhen Zhang, Chenlin Zhou, Jiayi Mao, Tianze Xia, Jiafeng Guo, and Shenghua Liu. 2025. A Survey of Context Engineering for Large Language Models. doi:10.48550/arXiv.2507.13334 arXiv:2507.13334 [cs].

[78] Bryan Min and Haijun Xia. 2025. Feedforward in Generative AI: Opportunities for a Design Space. doi:10.48550/arXiv.2502.14229 arXiv:2502.14229 [cs].

[79] Anirban Mukhopadhyay and Kurt Luther. 2025. OSINT Clinic: Co-designing AI-Augmented Collaborative OSINT Investigations for Vulnerability Assessment. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–22. doi:10.1145/3706598.3713283

[80] Anirban Mukhopadhyay and Kurt Luther. 2025. Tailoring Generative AI to Augment Creative Leadership in Capture-The-Flag Development. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI.* Yokohama, Japan.

[81] Syeda Masooma Naqvi, Malik Khadar, and Harmanpreet Kaur. 2025. Computational Reification of Creativity: How Generative AI Reshapes Creative Cognition. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI.* Yokohama, Japan. https://github.com/ai-tools-for-thought/workshop/blob/5a2aadd00942f89fc1b1c97fbbedd7ed4e0d92c7/documents/Naqvi%20et%20al%20-%20CHI2025_Workshop_Tools4Thought.pdf

[82] Charles Naumer, Karen Fisher, and Brenda Dervin. 2008. Sense-Making: a methodological perspective. In *Sensemaking workshop, CHI*, Vol. 8. 506–513.

[83] Cal Newport. 2016. *Deep work: Rules for focused success in a distracted world.* Hachette UK.

[84] Sachita Nishal, Mina Lee, Nicholas Diakopoulos, and Jennifer Wortman Vaughan. 2025. Designing for Agency in LLM-Infused Writing Support Tools for Science Journalism. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI.* Yokohama, Japan. https://github.com/ai-tools-for-thought/workshop/blob/5a2aadd00942f89fc1b1c97fbbedd7ed4e0d92c7/documents/Nishal%20et%20al%20-%20Tools%20for%20Thought%20-%20Camera-Ready.pdf

[85] Ayano Okoso, Mingzhe Yang, and Yukino Baba. 2025. Do Expressions Change Decisions? Exploring the Impact of AI's Explanation Tone on Decision-Making. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–22. doi:10.1145/3706598.3713744

[86] Srishti Palani and Gonzalo Ramos. 2024. Evolving Roles and Workflows of Creative Practitioners in the Age of Generative AI. In *Creativity and Cognition*. ACM, Chicago IL USA, 170–184. doi:10.1145/3635636.3656190

[87] Pat Pataranutaporn, Valdemar Danry, Joanne Leong, Parinya Punpongsanon, Dan Novy, Pattie Maes, and Misha Sra. 2021. AI-generated characters for supporting personalized learning and well-being. *Nature Machine Intelligence* 3, 12 (Dec. 2021), 1013–1022. doi:10.1038/s42256-021-00417-9 Publisher: Nature Publishing Group.

[88] Kris Pilcher and Esen K. Tütüncü. 2025. Purposefully Induced Psychosis (PIP): Embracing Hallucination as Imagination in Large Language Models. doi:10.48550/arXiv.2504.12012 arXiv:2504.12012 [cs].

[89] James Prather and Brent N. Reeves. 2025. Pedagogy for Critical Engagement with Generative AI. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI.* Yokohama, Japan. https://github.com/ai-tools-for-thought/workshop/blob/5a2aadd00942f89fc1b1c97fbbedd7ed4e0d92c7/documents/Prather%20and%20Reeves%20-%20CHI_2025___Workshop_for_Tools_on_Thought.pdf

[90] James Prather, Brent N Reeves, Juho Leinonen, Stephen MacNeil, Arisoa S Randrianasolo, Brett A. Becker, Bailey Kimmel, Jared Wright, and Ben Briggs. 2024. The Widening Gap: The Benefits and Harms of Generative AI for Novice Programmers. In *Proceedings of the 2024 ACM Conference on International Computing Education Research - Volume 1 (ICER '24, Vol. 1)*. Association for Computing Machinery, New York, NY, USA, 469–486. doi:10.1145/3632620.3671116

[91] Kevin Pu, kevin.pu@mail.utoronto.ca, https://orcid.org/0009-0000-4722-0631, View Profile, Daniel Lazaro, d.lazaro@mail.utoronto.ca, https://orcid.org/0009-0007-7964-7702, View Profile, Ian Arawjo, ian.arawjo@umontreal.ca, https://orcid.org/0000-0001-8910-0822, View Profile, Haijun Xia, haijunxia@ucsd.edu, https://orcid.org/0000-0002-9425-0881, View Profile, Ziang Xiao, ziang.xiao@jhu.edu, https://orcid.org/0000-0003-3368-0180, View Profile, Tovi Grossman, tovi@dgp.toronto.edu, https://orcid.org/0000-0002-0494-5373, View Profile, Yan Chen, ych@vt.edu, https://orcid.org/0000-0002-1646-6935, and View Profile. 2025. Assistance or Disruption? Exploring and Evaluating the Design and Trade-offs of Proactive AI Programming Support. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–21. doi:10.1145/3706598.3713357

[92] Sadhana Puntambekar and Roland Hubscher. 2005. Tools for Scaffolding Students in a Complex Learning Environment: What Have We Gained and What Have We Missed? *Educational Psychologist* 40, 1 (March 2005), 1–12. doi:10.1207/s15326985ep4001_1 Publisher: Routledge _eprint: https://doi.org/10.1207/s15326985ep4001_1.

[93] Peinuan Qin, Chi-Lan Yang, Jingshu Li, Jing Wen, and Yi-Chieh Lee. 2025. Timing Matters: How Using LLMs at Different Timings Influences Writers' Perceptions and Ideation Outcomes in AI-Assisted Ideation. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*.

Association for Computing Machinery, New York, NY, USA, 1–16. doi:10.1145/3706598.3713146

[94] Chris Quintana and Rebecca M. Quintana. 2025. Students' Experiences Using Generative AI Within Learning Experience Design Workflows. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI*. Yokohama, Japan.

[95] Leon Reicherts, Yvonne Rogers, Licia Capra, Ethan Wood, Tu Dinh Duong, and Neil Sebire. 2022. It's Good to Talk: A Comparison of Using Voice Versus Screen-Based Interactions for Agent-Assisted Tasks. *ACM Transactions on Computer-Human Interaction* 29, 3 (June 2022), 1–41. doi:10.1145/3484221

[96] Leon Reicherts, Zelun Tony Zhang, Elisabeth von Oswald, Yuanting Liu, Yvonne Rogers, and Mariam Hassib. 2025. AI, Help Me Think—but for Myself: Assisting People in Complex Decision-Making by Providing Different Kinds of Cognitive Support. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–19. doi:10.1145/3706598.3713295

[97] Advait Sarkar. 2023. Exploring Perspectives on the Impact of Artificial Intelligence on the Creativity of Knowledge Work: Beyond Mechanised Plagiarism and Stochastic Parrots. In *Proceedings of the 2nd Annual Meeting of the Symposium on Human-Computer Interaction for Work*. ACM, Oldenburg Germany, 1–17. doi:10.1145/3596671.3597650

[98] Advait Sarkar. 2023. Will Code Remain a Relevant User Interface for End-User Programming with Generative AI Models?. In *Proceedings of the 2023 ACM SIGPLAN International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software* (Cascais, Portugal) *(Onward! 2023)*. Association for Computing Machinery, New York, NY, USA, 153–167. doi:10.1145/3622758.3622882

[99] Advait Sarkar. 2024. AI Should Challenge, Not Obey. *Commun. ACM* 67, 10 (Oct. 2024), 18–21. doi:10.1145/3649404

[100] Advait Sarkar. 2024. Intention Is All You Need. In *Proceedings of the 35th Annual Conference of the Psychology of Programming Interest Group (PPIG 2024)*.

[101] Advait Sarkar. 2025. AI Could Have Written This: Birth of a Classist Slur in Knowledge Work. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25)*. Association for Computing Machinery, New York, NY, USA, Article 621, 12 pages. doi:10.1145/3706599.3716239

[102] Advait Sarkar and Ian Drosos. 2025. Vibe coding: programming through conversation with artificial intelligence. In *Proceedings of the 36th Annual Conference of the Psychology of Programming Interest Group (PPIG 2025)*.

[103] Advait Sarkar, Ian Drosos, Rob Deline, Andrew D. Gordon, Carina Negreanu, Sean Rintel, Jack Williams, and Ben Zorn. 2023. Participatory prompting: a user-centric research method for eliciting AI assistance opportunities in knowledge workflows. In *Proceedings of the 34th Annual Conference of the Psychology of Programming Interest Group (PPIG 2023)*.

[104] Advait Sarkar, Andrew D. Gordon, Carina Negreanu, Christian Poelitz, Sruti Srinivasa Ragavan, and Ben Zorn. 2022. What is it like to program with artificial intelligence?. In *Proceedings of the 33rd Annual Conference of the Psychology of Programming Interest Group (PPIG 2022)*.

[105] Advait Sarkar, Xiaotong (Tone) Xu, Neil Toronto, Ian Drosos, and Christian Poelitz. 2024. When Copilot Becomes Autopilot: Generative AI's Critical Risk to Knowledge Work and a Critical Solution. In *Proceedings of the Annual Conference of the European Spreadsheet Risks Interest Group (EuSpRIG 2024)*.

[106] R. Keith Sawyer. 2021. The iterative and improvisational nature of the creative process. *Journal of Creativity* 31 (Dec. 2021), 100002. doi:10.1016/j.yjoc.2021.100002

[107] Aaron Schecter and Benjamin Richardson. 2025. How the Role of Generative AI Shapes Perceptions of Value in Human-AI Collaborative Work. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–15. doi:10.1145/3706598.3713946

[108] Noam Scheiber. 2025. At Amazon, Some Coders Say Their Jobs Have Begun to Resemble Warehouse Work. *The New York Times* (May 2025). https://www.nytimes.com/2025/05/25/business/amazon-ai-coders.html

[109] Kathrin Schnizer and Sven Mayer. 2025. User-Centered AI for Data Exploration: Rethinking GenAI's Role in Visualization. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI*. Yokohama, Japan.

[110] Donald A Schön. 2017. *The reflective practitioner: How professionals think in action*. Routledge.

[111] Ava Elizabeth Scott, Lev Tankelevitch, and Sean Rintel. 2024. Mental Models of Meeting Goals: Supporting Intentionality in Meeting Technologies. https://www.microsoft.com/en-us/research/publication/mental-models-of-meeting-goals-supporting-intentionality-in-meeting-technologies/

[112] Katie Seaborn, Jacqueline Urakami, Peter Pennefather, and Norihisa P. Miyake. 2024. Qualitative Approaches to Voice UX. *Comput. Surveys* 56, 12 (Dec. 2024), 1–34. doi:10.1145/3658666 arXiv:2404.14736 [cs].

[113] Auste Simkute, Lev Tankelevitch, Viktor Kewenig, Ava Elizabeth Scott, Abigail Sellen, and Sean Rintel. 2024. Ironies of Generative AI: Understanding and mitigating productivity loss in human-AI interactions. doi:10.48550/arXiv.2402.11364 arXiv:2402.11364 [cs].

[114] Herbert A. Simon. 1956. Rational choice and the structure of the environment. *Psychological Review* 63, 2 (1956), 129–138. doi:10.1037/h0042769

[115] Anjali Singh, Karan Taneja, Zhitong Guan, and Avijit Ghosh. 2025. Protecting Human Cognition in the Age of AI. doi:10.48550/arXiv.2502.12447 arXiv:2502.12447 [cs].

[116] Alexa Siu and Raymond Fok. 2025. Augmenting Expert Cognition in the Age of Generative AI: Insights from Document-Centric Knowledge Work. doi:10.48550/arXiv.2503.24334 arXiv:2503.24334 [cs].

[117] Amira Skeggs, Ashish Mehta, Valerie Yap, Seray B Ibrahim, Charla Rhodes, James J. Gross, Sean A. Munson, Predrag Klasnja, Amy Orben, and Petr Slovak. 2025. Micro-narratives: A Scalable Method for Eliciting Stories of People's Lived Experience. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–20. doi:10.1145/3706598.3713999

[118] Petr Slovak. 2025. Tools for (Changing) Thought: Scaffolding Cognitive Skills Development within Mental Health Contexts. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI*. Yokohama, Japan.

[119] Petr Slovak and Sean A. Munson. 2024. HCI Contributions in Mental Health: A Modular Framework to Guide Psychosocial Intervention Design. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–21. doi:10.1145/3613904.3642624

[120] Keith E. Stanovich, Richard F. West, and Maggie E. Toplak. 2009. Cognitive complexity and the bias blind spot. *Journal of Personality and Social Psychology* 96, 3 (2009), 520–537. doi:10.1037/a0015557

[121] Sangho Suh. 2025. The Next Leap Forward in Human Intelligence: Expanding & Exploring Complex Spaces of Thought. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI*. Yokohama, Japan.

[122] Sangho Suh, Hai Dang, Ryan Yen, Josh M. Pollock, Ian Arawjo, Rubaiat Habib Kazi, Hariharan Subramonyam, Jingyi Li, Nazmus Saquib, and Arvind Satyanarayan. 2024. Dynamic Abstractions: Building the Next Generation of Cognitive Tools and Interfaces. In *Adjunct Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology (UIST Adjunct '24)*. Association for Computing Machinery, New York, NY, USA, 1–3. doi:10.1145/3672539.3686706

[123] Lev Tankelevitch, Elena L. Glassman, Jessica He, Majeed Kazemitabaar, Aniket Kittur, Mina Lee, Srishti Palani, Advait Sarkar, Gonzalo Ramos, Yvonne Rogers, and Hari Subramonyam. 2025. Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25)*. Association for Computing Machinery, New York, NY, USA, 1–8. doi:10.1145/3706599.3706745

[124] 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 (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–24. doi:10.1145/3613904.3642902

[125] Rishi Vanukuru, Payod Panda, Xinyue Chen, Ava Elizabeth Scott, Lev Tankelevitch, and Sean Rintel. 2025. Designing Interfaces that Support Temporal Work Across Meetings with Generative AI. In *Proceedings of the 2025 ACM Designing Interactive Systems Conference (DIS '25)*. Association for Computing Machinery, New York, NY, USA, 3600–3620. doi:10.1145/3715336.3735833

[126] Bret Victor. 2013. Media for Thinking the Unthinkable. https://vimeo.com/67076984. Talk presented at Worrydream.com.

[127] Hanna Wallach, Meera Desai, Nicholas Pangakis, A. Feder Cooper, Angelina Wang, Solon Barocas, Alexandra Chouldechova, Chad Atalla, Su Lin Blodgett, Emily Corvi, P. Alex Dow, Jean Garcia-Gathright, Alexandra Olteanu, Stefanie Reed, Emily Sheng, Dan Vann, Jennifer Wortman Vaughan, Matthew Vogel, Hannah Washington, and Abigail Z. Jacobs. 2024. Evaluating Generative AI Systems is a Social Science Measurement Challenge. doi:10.48550/arXiv.2411.10939 arXiv:2411.10939 [cs].

[128] Ruotong Wang, Xinyi Zhou, Lin Qiu, Joseph Chee Chang, Jonathan Bragg, and Amy X. Zhang. 2025. Social-RAG: Retrieving from Group Interactions to Socially Ground AI Generation. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–25. doi:10.1145/3706598.3713749

[129] Sitong Wang and Lydia B. Chilton. 2025. Schemex: Discovering Design Patterns from Examples through Iterative Abstraction and Refinement. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI*. Yokohama, Japan.

[130] Sterling Williams-Ceci, Maurice Jakesch, Advait Bhat, Kowe Kadoma, Lior Zalmanson, and Mor Naaman. 2025. Biased AI Writing Assistants Shift Users' Attitudes on Societal Issues. doi:10.31234/osf.io/mhjn6_v3

[131] Terry Winograd and Fernando Flores. 1987. *Understanding Computers and Cognition: A New Foundation for Design*. Addison-Wesley, Boston, MA, USA.

[132] Haijun Xia. 2025. Generative, Malleable, and Personal User Interfaces. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI*. Yokohama, Japan.

[133] Xiaotong (Tone) Xu, Arina Konnova, Bianca Gao, Cindy Peng, Dave Vo, and Steven P. Dow. 2025. Productive vs. Reflective: How Different Ways of Integrating AI into Design Workflows Affect Cognition and Motivation. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–15. doi:10.1145/3706598.3713649

[134] Longqi Yang, David Holtz, Sonia Jaffe, Siddharth Suri, Shilpi Sinha, Jeffrey Weston, Connor Joyce, Neha Shah, Kevin Sherman, Brent Hecht, and Jaime Teevan. 2022. The effects of remote work on collaboration among information workers. *Nature Human Behaviour* 6, 1 (Jan. 2022), 43–54. doi:10.1038/s41562-021-01196-4 Publisher: Nature Publishing Group.

[135] Yaqing Yang, Vikram Mohanty, Nikolas Martelaro, Aniket Kittur, Yan-Ying Chen, and Matthew K. Hong. 2025. From Overload to Insight: Scaffolding Creative Ideation through Structuring Inspiration. doi:10.48550/arXiv.2504.15482 arXiv:2504.15482 [cs].

[136] Ryan Yen, Josh Pollock, Caroline Berger, and Arvind Satyanarayan. 2025. Something In Between Formal Spec and Informal Representation. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI*. Yokohama, Japan.

[137] Bhada Yun, Dana Feng, Ace S. Chen, Afshin Nikzad, and Niloufar Salehi. 2025. Generative AI in Knowledge Work: Design Implications for Data Navigation and Decision-Making. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–19. doi:10.1145/3706598.3713337

[138] Angela Zavaleta Bernuy, Naaz Sibia, Pan Chen, Jessica Jia-Ni Xu, Elexandra Tran, Runlong Ye, Viktoria Pammer-Schindler, Andrew Petersen, Joseph Jay Williams, and Michael Liut. 2024. Does the Medium Matter? An Exploration of Voice-Interaction for Self-Explanations. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference (DIS '24)*. Association for Computing Machinery, New York, NY, USA, 86–101. doi:10.1145/3643834.3661596

[139] Haoqi Zhang. 2024. Searching for the Non-Consequential: Dialectical Activities in HCI and the Limits of Computers. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3613904.3641945

[140] Zelun Tony Zhang and Leon Reicherts. 2025. Augmenting Human Cognition With Generative AI: Lessons From AI-Assisted Decision-Making. doi:10.48550/arXiv.2504.03207 arXiv:2504.03207 [cs].

[141] Tim Zindulka, Sven Goller, Florian Lehmann, and Daniel Buschek. 2025. Content-Driven Local Response: Supporting Sentence-Level and Message-Level Mobile Email Replies With and Without AI. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–23. doi:10.1145/3706598.3713890

[142] Tim Zindulka, Jannek Sekowski, Florian Lehmann, and Daniel Buschek. 2025. Prompting by Doing: How Direct Manipulation can Protect and Augment Writers' Thoughts in AI Tools. In *CHI' 2025 Workshop on Tools for Thought: Research and Design for Understanding, Protecting, and Augmenting Human Cognition with Generative AI*. Yokohama, Japan.