

Learning by Chatting? Investigating the Impact of Generative AI on Information Seeking and Learning

Shravika Mittal^{1,*}, Su Lin Blodgett^{2,*}, Q. Vera Liao^{3,*}

¹Georgia Institute of Technology

²Mila - Québec AI Institute

³University of Michigan, Ann Arbor

smittal87@gatech.edu, sulin.blodgett@mila.quebec, veraliao@umich.edu

Abstract

Generative AI (GenAI) tools offer increasing opportunities for augmenting human cognitive tasks. Among these tasks, information seeking is being rapidly reshaped by GenAI tools, with potentially profound implications for learning and knowledge acquisition. To investigate these implications, we conducted a between-subjects field experiment in which participants pursued informal learning by seeking information through either ChatGPT or Google Search over a span of 8 days. Using a daily diary protocol, we gathered in-situ data on their information-seeking processes. Our findings show that participants in the ChatGPT group experienced diminished agency in their information-seeking processes, as they offloaded much of the information selection to AI, and consequently experienced greater meta-cognitive load arising from this reduced sense of control. We further highlight two sources of distortion in information access when using chat-based GenAI tools: biases in ChatGPT outputs, particularly towards providing solution-oriented artifacts over principled knowledge; and systematic shifts in users' information-seeking behaviors, whereby the conversational and socially-oriented interaction paradigm of current GenAI tools may inadvertently reduce exploration of the broader knowledge space. As a result, on average, participants in the ChatGPT group had worse learning outcomes than those using Google, especially for higher-order critical learning. Our work suggests inherent tensions between offloading information seeking to AI and meaningful learning, and provides broader implications for understanding AI's risks to human cognition.

Extended version — <https://arxiv.org/abs/2606.11669>

Introduction

In recent years, advances in Generative AI (GenAI) have shifted its role from primarily automating routine tasks to increasingly augmenting human capabilities (Wu et al. 2025). This shift is evident in its widespread adoption across domains of knowledge work (e.g., software engineering) (Suri et al. 2024), creative practices (Suri et al. 2024), and formal education (Nugroho et al. 2025). While GenAI tools can substantially improve productivity and efficiency, these

gains are often realized through cognitive offloading (Shen and Tamkin 2026)—the practice of shifting effortful cognitive tasks to external aids.

Among the various cognitive tasks, information seeking is one that has undergone a profound shift with GenAI-powered technologies. As users increasingly move away from the previously dominant paradigm of retrieval-based search to “ask ChatGPT,” they are offloading much of the mental labor involved in information seeking to chatbots and answer engines (Shah and Bender 2022). In particular, GenAI tools offer advanced information synthesis capabilities (such as summary generation), combined with natural language understanding capabilities, to help process flexible and highly specified queries. While these affordances can support rapid information look-up, they also offload critical cognitive processes, such as selecting, evaluating, and verifying information, to AI systems (Hirvonen et al. 2024).

Information seeking is foundational to many downstream cognitive tasks such as learning, decision making, problem solving, and creativity (Shah and Bender 2022; Spatharioti et al. 2025). Consequently, shifts in how users seek information through GenAI tools may influence subsequent cognitive outcomes. In particular, emerging evidence in education suggests that unconstrained use of popular GenAI tools such as ChatGPT can impair learning. For example, in classroom settings, heavier use of ChatGPT has been associated with poorer academic performance (Bastani et al. 2024), as GenAI tools can automate or accelerate answer-finding and task completion at the expense of deep conceptual understanding and genuine learning (Qu et al. 2025).

Unlike prior work focused on formal educational settings, we examine GenAI's impact on *informal learning*, through which individuals acquire knowledge and skills in everyday contexts (Qian, Fu, and Zhou 2025). Whether to develop an artistic skill, learn a new language, or acquire practical life skills, individuals often rely on informal, self-directed pursuits mediated by information systems. While these learning pathways were once defined by search engines—studied under the paradigm of “searching as learning” (Ghosh, Rath, and Shah 2018)—and social media platforms (Sengupta 2020), they are increasingly being reshaped by GenAI tools (Terzimehić, Bühler, and Kasneci 2025; Lira et al. 2025). Crucially, informal contexts lack the institutional guardrails and instructional oversight found in formal

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

*Work done while at Microsoft Research Montréal.

settings, making users more likely to offload cognitive efforts that may be essential for building durable expertise.

To understand the mechanisms through which GenAI tools (and their unique affordances) may impact knowledge acquisition, we investigate the shifts in people's information-seeking behaviors during informal learning in naturalistic settings. We conducted a between-subjects, longitudinal field experiment (8 days) in which participants were randomly assigned to perform an informal learning task using either ChatGPT, a widely used GenAI tool (OpenAI 2022), or the conventional Google Search. In addition to measuring participants' knowledge gain, we utilized a daily diary protocol to gather in-situ data about their information-seeking processes, and contrasted those for the two groups (ChatGPT vs. Google) through the lens of an information-seeking stage model (Ellis, Cox, and Hall 1993). Our work seeks to answer the following research questions:

RQ1: How does the use of GenAI tools impact people's informal learning outcomes?

RQ2: How does the use of GenAI tools impact people's information-seeking behaviors for learning?

Our research demonstrates the risks of offloading information-seeking tasks to GenAI tools. With GenAI acting as a “filter” of knowledge spaces and providing fewer affordances for differentiating and assessing information, participants had diminished agency and control in their information-seeking processes, and experienced greater meta-cognitive load as a result of this diminished sense of control. In addition, we highlight two sources of *distortion* in information access. First, distortion may arise from biases within the technology itself: ChatGPT currently exhibits biases in output generation—towards generating solution-oriented artifacts over principled knowledge, while providing little transparency or controllability over these biases. Second, distortion can occur via systematic shifts in people's information-seeking behaviors, in which the socially-oriented interaction paradigm of current GenAI tools, characterized by a “chat” interface and personalization as popular design choices, may inadvertently narrow query scope and reduce exploration of the broader knowledge space. This diminished agency and distortion in information-seeking processes were associated with worse informal learning outcomes, on average, in the ChatGPT group, particularly for higher-order critical learning.

Background and Related Work

Generative AI and Cognitive Offloading

Cognitive offloading refers to the use of physical actions or external tools to reduce the mental demands of a task (Risko and Gilbert 2016). People routinely offload cognitive efforts to digital systems in everyday contexts (Risko and Gilbert 2016). Most notably, with the wide adoption of information retrieval systems such as Google Search, people increasingly offload information storage, remembering where to find information rather than the information itself (Sparrow, Liu, and Wegner 2011). Traditionally, such offloading has been understood as a way to redistribute cognitive effort, allowing individuals to outsource lower-level cognitive processes

or sub-processes. Recent advances in GenAI introduce expanded forms of cognitive offloading. Unlike earlier digital tools that primarily support low-level, narrowly scoped cognitive tasks, GenAI tools increasingly support intermediate cognitive tasks such as synthesis, comprehension, and reasoning, and even automate complex cognitive workflows such as planning, decision making, and creativity. In other words, GenAI tools are increasingly integrated into human thinking processes, and in some cases, may even be replacing human thinking altogether (Lee et al. 2025).

These shifts have spurred great public and academic interest in examining GenAI's implications for human cognition, especially in knowledge work (Budzyń et al. 2025) and educational settings (Walker and Vorvoreanu 2025). A series of studies has examined how the use of popular GenAI tools impacts specific cognitive tasks such as student learning (Yang, Hsu, and Wu 2025), decision making (Spatharioti et al. 2025), creativity (Kumar et al. 2025), and critical thinking (Lee et al. 2025). A few common observations have emerged from this literature. First, the use of GenAI tools involves cognitive offloading and may therefore lead to reductions in overall cognitive effort and engagement, including meta-cognitive activities (Lee et al. 2025). This is further evidenced by a study comparing EEG-based brain activity among participants using ChatGPT, a search engine, and no external tools (Kosmyna et al. 2025).

Second, this bypassing of cognitive effort impairs knowledge and skill acquisition (Bastani et al. 2024; Walker and Vorvoreanu 2025; Qu et al. 2025), contributes to deskilling (Macnamara et al. 2024), and creates psychological harms such as loss of agency and ownership of one's work (Kosmyna et al. 2025). Even if GenAI tools are found to enable performance boosts in some contexts, studies commonly show that knowledge workers and students do worse, compared to those not using AI from the beginning, when they lose access to these tools (Kumar et al. 2025).

While situated in this broad research examining GenAI's impact on human cognition, particularly learning and knowledge acquisition, we identify a gap in the literature where the specific link between the increasing use of GenAI tools for information seeking and learning outcomes is under-investigated. As we review below, information seeking is a foundational component of learning, especially in everyday informal learning contexts where structured instruction and curricula are absent. Although GenAI offers greater opportunities for cognitive offloading than conventional retrieval-based search engines, especially through its information synthesis capabilities, information seeking remains a complex cognitive process whose transformation under GenAI warrants empirical examination.

Information-Seeking Behaviors and Learning

In information science, a large body of work on “searching as learning” (Ghosh, Rath, and Shah 2018) shows that people use search engines not only to retrieve or look up information, but also to engage in learning processes such as sense-making (Dervin et al. 1983) and bridging knowledge gaps (Shah and Bender 2022). In fact, the specific behaviors or actions people adopt during their search support differ-

ent cognitive modes of learning (Rieh et al. 2016; Lee et al. 2015). For example, receptive learning—i.e., remembering and reproducing what is learned—is supported by behaviors like targeted fact retrieval. Critical learning—i.e., thinking differently with respect to one’s own viewpoints—is supported by differentiating the value of information across multiple sources. Lastly, creative learning—i.e., generating new ideas or artifacts based on what is learned—is supported by sense-making. Therefore, scholars have looked at how search engines can be designed to better support search behaviors that lead to improved learning outcomes (Allan et al. 2012; Agosti et al. 2014; Bates 1989).

Many studies have produced models of information-seeking behaviors to elucidate how people retrieve information and to inform the design of better information systems (Allam et al. 2019). Ellis’s model of information-seeking behaviors (Ellis, Cox, and Hall 1993) identifies eight stages that people engage in when seeking information: starting, chaining, browsing, differentiating, monitoring, extracting, verifying, and ending. Wilson’s model highlights that information-seeking behaviors are influenced by personal, social, and environmental factors (Wilson 1999), while Kuhlthau’s model describes the interplay between affect and seeking behaviors (Kuhlthau 1991). These models have been used to study seeking behaviors across various environments (Ge 2010). Choo, Detlor, and Turnbull (1999) used Ellis’s model as an analytical lens to study knowledge workers’ information-seeking behaviors on the web.

Changes in information technologies are often followed by shifts in user behaviors (Mitra et al. 2014), with GenAI representing the latest disruption (Hirvonen et al. 2024). Moran, Rosala, and Brown (2025) found that the “AI overviews” panel appearing at the top of most search results captures users’ attention and discourages them from visiting the actual web pages. Vaithilingam, Zhang, and Glassman (2022) observed that participants prefer Copilot over search engines to initiate information seeking for programming tasks, but rely on human-vetted sources for information verification. Yet, we lack a systematic understanding of *how* this new generation of GenAI tools reshapes information-seeking behaviors that are essential for learning.

To close this gap, we conducted a longitudinal field experiment with a diary study protocol comparing information-seeking behaviors using ChatGPT (a popular GenAI tool) and a conventional search engine. While researchers have begun to examine the differences between the two paradigms (Spatharioti et al. 2025; Sharma, Liao, and Xiao 2024), prior work focuses on controlled lab settings, limiting their external validity, especially for learning-related behaviors and outcomes. The diary protocol allows us to collect in-situ data about how and why participants engage in all stages of information seeking, which we analyze through the lens of Ellis’s stage model (Ellis, Cox, and Hall 1993).

Methods

Study Design

We conducted a between-subjects, IRB-approved, longitudinal field experiment with a daily diary study protocol. Partic-

ipants were recruited from Prolific (Section “Participants”), a crowdsourcing platform widely used for human-subjects experiments. Over a period of 8 days (excluding weekends), participants were asked to use the assigned technology to seek information to learn about “nutrition and meal planning” (i.e., to perform an informal learning task)—a popular topic for informal learning (Terzimehić, Bühler, and Kasneci 2025). They were also encouraged to use (but were not restricted to) the technology for their other usual day-to-day information needs. They were randomly placed into one of two conditions, representing the assigned technology to use:

- ChatGPT: a popular GenAI-powered application, with information support (“ask anything”) as one of its primary functions.
- Google: the long-standing popular search engine, which currently controls over 90% of the global search engine market (Global Stats 2025). Google Search recently introduced “AI Overviews” (Google 2025), a feature that uses GenAI to provide a snapshot of key information related to users’ queries. To minimize infusion of direct GenAI outputs, we asked participants to append “-ai” (verified via screenshot review) to every search query to temporarily remove the panel (Rasool 2024).

By comparing ChatGPT with Google Search without AI overviews, we captured two distinct ends of the information-seeking spectrum, i.e., GenAI-driven vs. conventional approaches. The field experiment included three components:

- Onboarding; Pre-learning knowledge test: Participants consented to the study and completed (1) a demographics survey, (2) a pre-learning knowledge test to capture their baseline knowledge about the learning topic, and (3) example tasks to prepare for the daily diary study (see below). Refer to Appendix for the onboarding survey.
- Diary study: Following onboarding, each day for eight days, participants were asked to complete two tasks as part of the diary study: (1) *Informal learning task*: Every morning, we provided a guiding prompt, asking participants to learn about a subtopic related to the informal learning topic (see below). They were required to use the assigned technology for at least 10 minutes and could use it however they preferred. While completing the learning task, they were asked to upload screenshots for any number of interactions with the technology and comment on them; (2) *Diary entry task*: After completing the learning task, participants were required to submit a diary entry reflecting on their information-seeking activities for the day. They were asked to reflect on (a) how they sought information to complete the learning task and (b) at least one additional instance of information seeking.¹
- Post-learning knowledge test: Post experiment, participants completed a knowledge test designed to assess their learning outcomes.

Informal Learning Task. We selected an informal learning task centered on acquiring a practical life skill (Qian, Fu,

¹For the scope of this paper, we only focused on participants’ responses to the first diary entry prompt.

and Zhou 2025)—specifically, to learn about “nutrition and meal planning.” This topic fits the criteria of informal learning as it is self-directed, potentially personally relevant, requires active information seeking, and allows people to construct knowledge in informal contexts. Selecting this topic also enabled us to create a knowledge test and measure participants’ learning outcomes, suitable for answering RQ1.

To facilitate learning over the eight days, we created broad daily guiding prompts, outlined in the Appendix (Section “Guiding Prompts”), in reference to beginner-friendly, highly-rated, and heavily registered relevant courses (Felix Harder 2025a,b). For example, over the course of the eight days, participants were progressively asked to learn about macronutrients, how to plan meals tailored to a specific fitness goal, and strategies for eating healthy when traveling.

Diary Study Protocol. Diary studies provide an opportunity to collect data on participants’ situated practices in naturalistic settings over an extended period of time, while minimizing recall bias, making them particularly suited for RQ2 (Bolger, Davis, and Rafaeli 2003). After completing the learning task and before the end of the day, participants were asked to spend at least 20 minutes writing diary entries reflecting on their information-seeking activities. To design our diary entry protocol (Appendix), we drew on prior work (Hiniker et al. 2016). The protocol asked participants to answer two open-ended questions: to (1) elaborate on how they sought information to learn about the day’s subtopic using the assigned technology; and (2) describe how they addressed at least one additional information need. Under each question, we added prompts to guide participants to describe how they used the technology, what they learned and how they arrived at it, any challenges they encountered, the steps taken to address them, and their overall experience seeking information. We took several measures to ensure good quality diary data, including daily review of submissions, participant feedback, and safeguards against AI-generated entries (Appendix Section “Diary Study: Quality Control”).

Quantitative Measurement: Learning Outcome Score (RQ1)

Knowledge Test. We operationalized Bloom’s taxonomy (Anderson and Krathwohl 2001) to design a knowledge test (Appendix Section “Knowledge Test”) that measured participants’ different levels of learning on the assigned topic following the eight-day study (see “Background and Related Work” for the learning levels). Bloom’s taxonomy is widely used in education to characterize different levels of learning objectives by cognitive complexity (from low to high): remembering, understanding, applying, evaluating and creating. To measure participants’ *receptive learning outcomes*, we gathered (with minor modification) multiple-choice questions from relevant courses on Udemy that focused on remembering, recalling, and understanding key concepts. To measure *critical learning outcomes*, we designed a mix of multiple-choice and open-ended questions, asking participants to apply, analyze, or evaluate certain concepts based on what they had learned. For instance, participants were asked to critically evaluate two meal plans ac-

cording to their nutritional quality. Relevant course materials, e.g., meal plans and end of course quizzes, from Udemy were used to create these questions. Finally, to measure *creative learning outcomes*, we asked participants to create an artifact: in our case, a three-day meal plan.

We iteratively refined this knowledge test using feedback and responses from 11 pilot participants. Adjustments included removing questions that were answered correctly by a majority of the participants using just their baseline knowledge, and making minor tweaks to reduce ambiguity in the open-ended questions. We also ensured that obtaining the information needed to answer the questions required effort in both conditions (e.g., it could not be found in the first sentences or titles displayed when querying ChatGPT or Google Search with a relevant subtopic).

Scoring the Knowledge Test. The knowledge test consisted of nine multiple-choice and four open-ended questions. Since the multiple-choice questions were directly taken from courses on Udemy, we were able to score participants’ responses against the course-provided ground truth using a binary correct (1), incorrect (0) scale.

Through multiple discussions and in reference to existing guidelines on nutrition (Mayo Clinic 2024; USDA 2020), we iteratively developed rubrics to score participants’ responses to the four open-ended questions, by first independently rating responses from seven participants in each group, and then discussing and resolving disagreements. Responses to two of the questions (Q11, Q12) for evaluating artifacts (nutrition labels and meal plans) were scored on a binary scale. Responses to the other two questions (Q13, Q14) for improving or creating a meal plan were scored on a 1–3 scale, with higher scores for better responses according to the designed rubric (Appendix Section “Rubric for Open-ended Questions”). Participants’ scores on all knowledge test questions were summed to obtain their *learning outcome score*.

Qualitative Analysis of Diary Entries (RQ2)

To answer RQ2, we analyzed participants’ responses to the first diary entry prompt, where they described how they sought information to complete the daily informal learning task. Prior work has proposed a range of models to characterize people’s information-seeking behaviors (Ellis, Cox, and Hall 1993; Kuhlthau 1991; Wilson 1999). Among these, Ellis’s model offers a particularly suitable analytical lens for our study, as while other models emphasize contextual or affective factors governing information seeking, Ellis’s model focuses on observable behaviors, i.e., how information seekers interact in an information environment.

Ellis’s model proposes eight behaviors people may engage in during their information-seeking processes: starting (*initiating information seeking*), chaining (*following connections to pursue a line of inquiry*), browsing (*casually seeking information in potential areas of interest*), differentiating (*using known differences to filter information*), monitoring (*regularly following particular sources*), extracting (*selectively identifying relevant information*), verifying (*checking the correctness of information*), and ending (*wrapping up*). It has been adopted to study behaviors of dif-

ferent user populations within varied information environments, both digital and offline (Choo, Detlor, and Turnbull 1999; Ganie et al. 2019; Chen and Feng 2024). Using Ellis's model as a guiding framework, the first author conducted a preliminary thematic analysis of a subset of diary entries. Although the analysis began with a deductive coding-based approach, several sub-codes were generated inductively to capture participants' actions at a more granular level. These sub-codes further emerged as a consequence of the distinct ways in which participants engaged with ChatGPT or Google Search. Following this, the first author met with the other co-authors to discuss, refine, and finalize the coding framework. The final framework is summarized in Table 1.

After finalizing the codes and their definitions, the first author conducted a thematic analysis of participants' entries from days 2 to 8. Entries from day 1 were intentionally excluded, as participants were still familiarizing themselves with the writing process. In total, 98 entries from the ChatGPT condition and 140 entries from the Google condition were analyzed (with 7 entries missing from the Google condition). The difference in the number of entries across groups is due to the differing number of participants who completed the study in each group (see below).

Participants

Recruitment. In total, we invited 144 participants on Prolific to complete a preliminary screening survey. Among those that fit our participation criteria (below), 80 were invited to participate in the experiment, recruited in two batches spaced one week apart (N1=50; N2=30). The second batch was recruited to account for participant attrition in the first batch (see below).

The screening survey identified participants who (1) were moderate users of GenAI, (2) had moderate knowledge about and interest in "nutrition and meal planning" (the informal learning task), and (3) were likely to engage in informal learning. These criteria were intended to filter potential outlier behaviors, avoid floor and ceiling effects when measuring learning outcomes, ensure that participants engaged in the informal learning task voluntarily, and support participant retention. We also disclosed the topical focus of the experiment and screened out participants who might not feel comfortable seeking nutrition-related information. Participants were compensated up to \$105; \$8 per day over a 10-day period for completing 30-minute daily tasks, and \$25 for a 45-minute semi-structured interview. Compensation was provided on a daily basis.

Of the 40 participants in each condition, 14 in the ChatGPT condition and 21 in the Google condition completed the entire study. The high attrition rate can be attributed to the nature of a diary study (Ohly et al. 2010). While not statistically significantly different, the higher drop-out rate in the ChatGPT condition is consistent with our results that participants in this group had more negative experiences and higher cognitive load. Participants who voluntarily dropped out expressed difficulty with writing diary reflections, uploading screenshots, or sustaining participation over the full study period. To maintain data quality, we also removed participants who consistently submitted low-quality diary entries

despite receiving repeated feedback. We excluded data from participants who dropped out or were removed early.

Background. During recruitment, we aimed to achieve a balanced demographic distribution across the two groups. Among participants in the ChatGPT and Google conditions, respectively, 42.86% and 19.05% identified as women, 57.14% and 76.19% identified as men, and 4.76% in the Google condition identified as non-binary. Age distributions were: 26–35 (21.43% and 23.81%), 36–45 (50.00% and 42.86%), 46–55 (21.43% and 23.81%), and above 55 (7.14% and 9.52%) for the ChatGPT and Google conditions, respectively. University or college was the most commonly reported level of education (42.86% and 71.43%), followed by high school (28.56% and 23.81%), and graduate degree (14.29% and 4.76%).

During onboarding, participants completed a pre-learning knowledge test. This consisted of three multiple-choice questions (scores ranged from 0–3) from a subset of the final knowledge test (Q4, Q5, and Q8). We included only a subset of questions to keep the onboarding process at a reasonable length and to minimize potential recall effects during the final post-learning assessment. A Welch's *t*-test revealed no significant differences in the pre-learning topical knowledge between the two groups (*t*-statistic: 0.647, *p*-value (*p*): 0.261). On average, participants in the ChatGPT group scored 1.64 (standard deviation (SD): 0.72), and participants in the Google condition scored 1.48 (SD: 0.73).

Results

Impact of GenAI on Learning (RQ1)

Learning Outcome Score. Given that the two groups started with comparable baseline knowledge on the assigned topic, we performed a Welch's *t*-test to compare participants' learning outcome scores as measured by the post-learning knowledge test. On average, participants in the ChatGPT and Google conditions scored 11.62 (SD: 2.32) and 12.86 (SD: 2.21), respectively. The test did not reach statistical significance but indicates a trend that participants in the ChatGPT condition had worse learning outcomes than those in the Google condition (*t*-statistic: -1.523 , *p*: 0.069). Regression analyses, controlling for baseline knowledge, showed a similar trend.

Question-level Analysis Next, since the questions were designed to cover different aspects (i.e., subtopics and levels of learning), we performed an exploratory, question-level analysis to examine on which aspects participants in the two groups showed different learning effects. Specifically, for each multi-choice question, we performed a Fisher's exact test, recommended for comparing binary outcomes within small samples, on participants' answer correctness between the ChatGPT and Google conditions. This analysis revealed that participants in the Google group were significantly more likely to correctly answer two of the multiple-choice questions designed to assess critical learning, i.e., application or evaluation of concepts learned.

First, one of these questions (Q8) required participants to apply their knowledge and interpret a given nutrition

label. In the post-learning knowledge test, a significantly smaller proportion of participants in the ChatGPT condition answered Q8 correctly, compared to those in the Google group (57% vs. 95%; odds ratio: 0.07; Fisher's p : 0.01). This difference was not significant when the participants attempted the same question in the pre-learning knowledge test (64% vs. 62%; odds ratio: 0.56; Fisher's p : 0.47). Based on qualitative analysis of participants' diary entries on the day they were assigned a subtopic relevant to Q8 (day 6), the ChatGPT group's worse performance could be attributed to their more limited exposure to realistic and diverse examples of nutrition labels. For example, one participant noted (P6; ChatGPT), "*while learning with ChatGPT, one aspect I found unsatisfying was the lack of visual examples, which made it harder to apply the advice directly,*" while another reflected (P7; ChatGPT), "*I wish [ChatGPT] could give me an example of a real nutrition label to gain more practical knowledge.*" This limitation forced some participants to engage in additional information-seeking activities from other channels: "*I paired the guidance with labels from products in my pantry*" and "*I will now try [to get examples of some real nutrition labels] using Google.*"

Next, the other question (Q7) required participants to compare and evaluate animal and plant sources of protein. Again, a significantly smaller proportion of participants in the ChatGPT condition answered correctly compared to those in the Google condition (21% vs. 57%; odds ratio: 0.21; Fisher's p : 0.04). This difference could stem from content bias within ChatGPT: in the diary entries, participants reflected that ChatGPT by default provided meat-based sources of protein. For example, one participant (P3; ChatGPT) noted "*[ChatGPT] automatically assumes that everyone eats meat.*" Taken together, these limited and biased outputs provided by ChatGPT, combined with participants' reduced exploration and agency in learning, as discussed in the qualitative results below, may have contributed to their worse learning outcomes, particularly those for higher-level critical learning.

Impact of GenAI on Information-Seeking Behaviors (RQ2)

As described earlier, we conducted a thematic analysis of participants' diary entries using the coding framework summarized in Table 1. n_C/n_G represents the number of entries from participants in the ChatGPT/Google condition that mentioned the information-seeking behavior or code. These numbers should not be interpreted as quantitative measurements, but serve to guide our attention towards how information-seeking behaviors differed in the two conditions. We present the main contrasts below.

Reduced Querying Scope in the ChatGPT Condition.

The frequencies of the "Starting" codes (see Table 1) show that participants in the ChatGPT condition were more likely to begin their information-seeking with a narrow or specified query, a behavior observed in 32.65% of their entries ($n_C=32/98$) compared to only 3.57% of entries from Google participants ($n_G=5/140$). This suggests that interacting with ChatGPT may have constrained participants' information-

seeking trajectories and reduced their exploration of the knowledge space from the very beginning.

The narrower start happened in two ways: First, instead of seeking a general overview of the informal learning topic, participants tailored their conversational queries to get specific, personally relevant information (see excerpt below):

"I used ChatGPT to learn about macronutrients, specifically, by asking how much of each a person should consume per day by providing a rough estimate of my height, weight and activity levels." – ChatGPT participant (P3)

This behavior could be attributed to the conversational interface invoking a social interaction-oriented schema (Lee, Kiesler, and Forlizzi 2010): "*[feels] as if I am having a conversation*" and "*led to an interesting conversation.*" In some cases, participants even referred to it as a nutritional coach: "*It felt as though I was getting a nutritional coach for free.*" In contrast, participants in the Google condition started their information-seeking process with a broader query: one participant noted, "*I searched how much of each macronutrient is recommended for a balanced meal.*"

Second, ChatGPT's interface affordance—a chat-based UI enabling easy revisiting of prior interactions—made participants more inclined to continue their previous interactions rather than initiating a completely new query:

"To learn about today's topic, I simply opened my prior interaction with ChatGPT and asked it to summarize all the lessons learned so far." – ChatGPT participant (P3)

Moreover, even when participants began with broad queries, ChatGPT's context retention and personalization limited exploration:

"I asked ChatGPT how much of each macronutrient one should have. Since my information was saved from yesterday, it gave me a personalized guide." – ChatGPT participant (P11)

In short, certain affordances of ChatGPT, including conversational interactions and easy access to and personalization based on past interactions, can constrain what people choose to query (Shah and Bender 2022), reducing knowledge exploration. This echoes prior findings that conversational interactions can bias people towards issuing personal, opinionated, and highly specified inputs, potentially reinforcing "information bubbles" (Sharma, Liao, and Xiao 2024).

Reduced Information Differentiation and Verification in the ChatGPT Condition.

Participants in the ChatGPT condition reported fewer instances of differentiating information, i.e., assessing (and filtering) information based on specific criteria (e.g., personal relevance, credibility, type of information need) and their indicators ("Differentiating" in Table 1, 43/125 instances in ChatGPT/Google conditions).

First of all, they applied a narrower set of criteria to filter information, relying almost entirely on personal relevance:

"I no longer eat meat, so [...] I focused on the vegetarian foods that were recommended by ChatGPT." – ChatGPT participant (P12)

Behavior (or Code)	Description	n_C	n_G	$\%_C$	$\%_G$
Starting (broad)	Beginning with exploratory, wide-ranging queries to serve as starting points for the information-seeking process.	66	135	67.35	96.43
Starting (narrow)	Beginning with specified, focused queries (e.g., tailored to personal preferences or well-defined needs) to serve as starting points for the information-seeking process.	32	5	32.65	3.57
Chaining (self-initiated)	Following a line of inquiry by using follow-up queries, links, or citations that build on previously retrieved information.	36	39	36.73	27.86
Chaining (technology-mediated)	Following a line of inquiry by using queries, links, or citations suggested by the technology (e.g., via ChatGPT-suggested follow-up prompts, a previous interaction, or ‘people also search for’ panel in Google).	4	7	4.08	5.00
Browsing	Casually seeking information in potential areas of interest, without a specific information need.	9	62	9.18	44.29
Differentiating (based on personal relevance)	Using indicators of personal relevance to filter information.	43	68	43.88	48.57
Differentiating (based on credibility)	Using indicators of credibility to filter information.	0	43	0	30.71
Differentiating (based on type of information need)	Using indicators of the type of information needed (e.g., level of detail, format, topic) to filter information.	0	14	0	10.00
Monitoring	Regularly following particular sources of information (e.g., a trusted source, or a source encountered earlier).	0	26	0	18.57
Extracting (information)	Selectively identifying relevant fundamental or principle-oriented information.	65	136	66.33	97.14
Extracting (artifacts)	Selectively identifying artifact-oriented information, such as meal plans, grocery lists, or nutrient-intake trackers.	46	9	46.94	6.43
Expressing an intent to verify	Indicating an intention to verify information, without actually performing the verification action.	3	0	3.06	0
Verifying (via another source)	Checking the correctness of information by consulting another source, such as a different information channel or another webpage.	3	27	3.06	19.29
Verifying (via prior knowledge)	Checking the correctness of information via prior knowledge.	4	4	4.08	2.86
Ending (Satisfied)	Concluding the information-seeking process with overall satisfaction (information need was met), even if minor frustrations or interruptions were expressed.	86	125	87.76	89.29
Ending (Unsatisfied)	Concluding the information-seeking process with clear dissatisfaction, as the information need remained unmet.	12	15	12.24	10.71

Table 1: Summary of the codebook used to study participants’ information-seeking behaviors as described in their diary entries. n_C/n_G represents the number of entries from ChatGPT/Google participants. $\%_C/\%_G$ represents the percentage of total entries from each group that described a given behavior. We analyzed 98/140 diary entries from the ChatGPT/Google groups.

On the other hand, participants in the Google condition considered additional indicators, such as source credibility and the type of information need to filter:

“I looked for sources that I know to be reputable [...] I clicked on an NIH result.” – Google participant (P7)

These differences in information differentiation can be attributed to offloading the information selection process to ChatGPT’s generative capabilities, and ChatGPT providing fewer affordances for people to make their own selection. While intended to provide fast and convenient information access, this may deprive information seekers of control over their own information-seeking processes. For example, some participants in the ChatGPT condition expressed a reduced sense of control, and noted that they were unable to learn in the ways they would normally prefer:

“I prefer to seek out my own sources. Using [ChatGPT] is just not what I would prefer. [When learning]

I like to read as much or as little detail as I would like.” – ChatGPT participant (P2)

In contrast, participants in the Google condition noted that the ability to differentiate information sources based on their information needs gave them greater control over what to consume. For instance, one participant highlighted the ability to switch sources depending on whether they sought general advice or more nuanced, scientific information:

“I used a Reddit thread and a blog, which worked well for tips and advice, but earlier in the week, when I was searching for information on macronutrients and scientific topics, these sources were less helpful. I was able to adjust my sources accordingly as my information needs changed.” – Google participant (P7)

Next, in the verifying stage (after information extraction), there were only seven instances where ChatGPT participants checked the accuracy of the information. They either did so

by consulting another source (three instances, two by the same participant), such as referring to official health-related web pages (e.g., “*cross-referenced some of the suggestions with official health websites*”) or by relying on their own prior knowledge (four instances; e.g., “*I have also heard that mentioned elsewhere*”). Interestingly, a few reflections mentioned an *intent to verify* the information obtained (three instances); however, they did not describe engaging in the actual verification behavior. Such an intent was expressed as skepticism about how ChatGPT worked (e.g., “*I am not sure how accurate that is since the data is probably from 2023*”) and whether the information provided by it could be trusted (e.g., “*[seeking health-related information] is not something I [would] trust a chatbot with right now*”).

Meanwhile, as reflected in the frequency of the “Verify-ing” code (Table 1), participants in the Google condition verified information more frequently (31 instances). Participants often did so by consulting multiple web pages:

“I relied on an article from a website I considered reliable. It all looked legitimate, but I fact-checked to be sure. [...] I looked at several other websites quickly.”
– Google participant (P6)

Overall, these findings suggest a shift toward a more passive form of learning (Bonwell and Eison 1991) in the ChatGPT condition, as ChatGPT provides synthesized information in its output at the cost of taking away information seekers’ agency in filtering, assessing, and selecting information. This not only results in diminished sense of control over one’s learning process, but can also be especially detrimental to critical and creative learning outcomes (Rieh et al. 2016).

ChatGPT Reduced Extraction of Fundamental or Principle-Oriented Information. As shown by the frequencies of the “Extracting” codes in Table 1, participants in the ChatGPT condition were less likely to extract fundamental or principle-oriented information (i.e. the underlying principles for how things should be done and why, such as information about the recommended amounts of macronutrients) as takeaways from their information-seeking process. 66.33% ($n_C=65/98$) entries from ChatGPT participants described this behavior, compared to 97.14% ($n_G=136/140$) of entries from participants in the Google group. They were more likely to extract solution-oriented, artifact-based information, such as meal plans, grocery lists, or nutrient-intake PDF trackers, compared to the Google group (“Extracting (artifacts)” in Table 1); this behavior was observed in 46.94% ($n_C=46/98$) of entries from ChatGPT participants, compared to only 6.43% ($n_G=9/140$) of entries from the Google group.

Even when a participant queried to obtain principle-oriented information, such as recommended macronutrient amounts for a balanced meal, ChatGPT provided a sample meal plan, followed by a suggested grocery list:

“I asked ChatGPT how much of each macronutrient one should have for a balanced meal. It gave me a balanced meal plan. [...] It asked if I’d want a grocery list, which it provided.” – ChatGPT participant (P11)

By contrast, a similar query by a participant in the Google group led them to learn about the principles of recommended amounts of macronutrients:

“I searched how much of macronutrients are needed for a balanced meal. A dietitian blog explained: carbs (45% to 65% of total calories), protein [...] and fats [...], for an average diet.” – Google participant (P6)

Receiving solution- or artifact-oriented information may also lead participants to expect to rely on ChatGPT for situations where they need to apply the knowledge, reducing their motivation to develop the expertise in themselves (Cronin-Golomb and Bauer 2023), as shown by the quote from the same participant (P11):

“I plan to use ChatGPT, even after this study, to create meal plans for me. It has always been overwhelming for me to do so.” – ChatGPT participant (P11)

In short, ChatGPT currently exhibits a bias towards generating solution-oriented artifacts, rather than principle-based information to allow someone to build fundamental knowledge about how things should be done and why. This bias could potentially be a by-product of training ChatGPT as a product targeting knowledge workers who may rely heavily on artifact generation (e.g., programming assistance, writing support), making it a mismatch for learning purposes.

Unsatisfied Ending from Meta-cognitive Load in the ChatGPT Condition. Participants in the ChatGPT condition were slightly more likely to end their information-seeking process expressing dissatisfaction (“Ending (Unsatisfied)” in Table 1). The code was observed in 12.24% of entries from participants in the ChatGPT group ($n_C = 12/98$), compared to 10.71% of entries from those in the Google group ($n_G = 15/140$).

In the ChatGPT condition, many unsatisfactory endings were rooted in meta-cognitive load and frictions (Tankelevitch et al. 2024)—i.e., the extra mental effort required to monitor and attempt to control one’s information-seeking process and the frustrations from failures in doing so:

“I asked: shouldn’t a person consult a doctor before making any major changes. I was frustrated when it said it’s good to consult a doctor but not necessary. [...] I don’t agree with this. I got more direct and expressed my concerns [...], asking ChatGPT if it was okay for it to provide that advice. Again, it hedged to say that it tries to provide information that does not harm anyone. [...] Using ChatGPT was frustrating and I just left the chat.” – ChatGPT participant (P2)

Due to a lack of affordances for controlling information selection from ChatGPT’s outputs, the main path through which one can control their information seeking is prompting and re-prompting. However, prompting remains an unpredictable task where users do not always get the intended output (the right content in the right format), referred to as the “instruction gap” (Subramonyam et al. 2024), as illustrated in the excerpt below:

“ChatGPT chose to provide two separate answers this time. It provided the information in a different layout,

which I found confusing. It made me second-guess my query.” – ChatGPT participant (P8)

More often than not, this burden on users led them to question their own capability for using ChatGPT:

“Users have to learn how to use ChatGPT. [...] I tried my best to then ask a basic query: how to go about losing weight, but the assistant again gave me responses that included information on both losing weight AND gaining muscle.” – ChatGPT participant (P4)

In some instances, the inability to obtain information in a desirable format led participants to abandon their information-seeking altogether:

“[ChatGPT] listed the macronutrient quantity in grams. I was confused, as I have no idea how to measure food in grams. [...] This was of no use to me.” – ChatGPT participant (P7)

In the ChatGPT condition, even when information needs were met in the end (among those who expressed satisfaction), participants experienced heightened meta-cognitive load because suggestions or follow-up prompts from ChatGPT were distracting them from their learning process and they had to make a conscious effort to stay on track:

“ChatGPT’s follow-up prompts were interruptive. They steered the conversation towards diet-related topics, even when I was asking more science-focused questions.” – ChatGPT participant (P6)

“ChatGPT was ‘eager’ to give me a sample meal plan, even though I was asking about something else altogether. I declined.” – ChatGPT participant (P11)

These results further show that participants in the ChatGPT condition experienced a diminished sense of control in their information-seeking process. This placed high meta-cognitive demands and frictions on them to monitor, attempt to regain control over, or regulate their interactions to stay “on track” with their learning goals. Currently, ChatGPT provides few affordances for users to meet these demands, leading to dissatisfaction and abandonment of the system.

Discussion

Through a between-subjects, longitudinal field experiment using a diary study protocol, we found evidence that using ChatGPT for informal learning may lead to worse learning outcomes than using conventional search engines, especially for higher-order critical learning (Lee et al. 2015). Participants’ diary entries revealed that the worse learning outcomes in the ChatGPT group can be attributed to their reduced exploration of the knowledge space and lack of agency in information selection, combined with limited and biased information provided by ChatGPT, leading to an overall more unsatisfactory learning experience from a diminished sense of control. Reflecting on these observations, below we discuss two implications of cognitive offloading to AI for information seeking, learning, and broader contexts.

Fundamental Tensions between Cognitive Offloading and Learning

Compared to conventional search engines, which arguably allow users to offload the cognitive tasks of information retrieval and storage, GenAI tools’ synthesis capabilities can offload much of the extended processes of collecting, selecting and assessing information to AI, resulting in diminished human agency and control over one’s own information-seeking process. While potentially bringing value via, e.g., improving the efficiency and potentially the quantity of information access, this shift of agency may present fundamental tensions for learning. As our qualitative results suggest, learners may lose the means and motivation to thoroughly explore the knowledge space, miss opportunities to engage in information filtering and assessment, which are key to critical learning and developing durable knowledge, and become passive receivers of information, forming only surface-level understanding with an “illusion of knowledge” (Walker and Vorvoreanu 2025). Since agency and control play critical roles in learning, we also observe participants experiencing heightened meta-cognitive load and frustration, and overall unsatisfactory user experience with GenAI tools for learning purposes.

Our results echo, and provide explanations for, recent literature in the education field which finds that the “default” designs of GenAI tools, which are oriented towards cognitive offloading for productivity gain, combined with students’ unconstrained use without guardrails, are often detrimental to learning (Bastani et al. 2024; Qu et al. 2025). Notably, research efforts have emerged to develop dedicated, structured GenAI tools for education, which often focus on cognitive *scaffolding* instead of offloading, such as providing feedback, prompting reflection, and engaging in Socratic questioning to facilitate structured exploration of the knowledge space (Tabarsi et al. 2025). However, the development and adoption barriers for these structured tools remain steep, especially as unstructured, “general-purpose” GenAI tools such as ChatGPT are easily accessible.

Outside formal education settings, people are even more likely to continue using these general-purpose GenAI tools for knowledge and skill acquisition, in both planned and unplanned fashions. To preserve and even enhance these opportunities for informal learning we suggest two areas to develop new sets of affordances in GenAI tools in addition to the scaffolding approaches explored in educational technology research. One area is to provide affordances for users to understand and control which processes or subprocesses—e.g., which of Ellis’ model’s information-seeking stages—are offloaded to AI. Such modular support would allow GenAI tools to be flexibly configured to meet different goals (e.g., learning versus productivity) and the demands of diverse cognitive tasks. Second, the (sometimes unbounded) opportunities to offload cognitive tasks to AI call for additional meta-cognition support (Tankelevitch et al. 2024): improving users’ awareness of their own goals, processes, capabilities, and—in learning contexts in particular—their self-regulation capabilities (Ertmer and Newby 1996) to take charge of their own learning goals while using AI; monitor-

ing their progress; and managing their motivation, emotions, and behaviors, including their use of AI.

GenAI's Risks in Cognitive Distortion

GenAI tools are often framed as opportunities for intelligence augmentation (Engelbart 2023), offloading some cognitive tasks to extend human capabilities. This framing often invokes the analogy of the introduction of calculators, which undoubtedly have extended people's mathematical capacities in everyday life. However, unlike calculators with direct control of input and deterministic output, our results highlight two forms of distortion risk—systematic change in the outcome of a cognitive process as compared to without AI—associated with cognitive offloading to GenAI: (1) distortions introduced through the generated outputs themselves, and (2) distortions introduced through systematic shifts in people's interaction behaviors. Together, these distortions can reshape downstream cognitive tasks, learning processes, and broader patterns of information engagement.

First, biases are inevitable with GenAI outputs. The generative synthesis capability acts as “information filters” of a much broader information space, and these filters are not value-free, but laden with values and judgments of technologies' developers. For example, we observed that ChatGPT is currently biased towards generating solution-oriented artifacts rather than principle-based information. This is likely influenced by the technology developers' biases towards supporting artifact-oriented knowledge work tasks (e.g., coding or writing). As user logs of GenAI tools report heavy use towards knowledge work-related tasks (Suri et al. 2024; Handa et al. 2025), these usage patterns or user bases likely feed back into these technologies' training and design decisions. Unfortunately, such a bias runs counter to the fundamental purpose of learning: to gain principled knowledge about *how* and *why* (Lee et al. 2015). Problematically, these biases are often not understood, nor made transparent or controllable, leaving the distortion effects of cognitive loading potentially long-lasting but unexamined.

Second, our results suggest that this distortion risk can also result from systematic shifts in people's interaction behaviors, i.e. what and how one offloads to AI. These behavioral shifts can inadvertently result from seemingly popular design choices of GenAI tools. For example, echoing prior work (Sharma, Liao, and Xiao 2024), we found that the social, chat-based paradigm had unintended consequences on narrowing participants' information querying. Certain technical affordances that have been extolled for the new generation of “companion”-like GenAI tools, such as context preservation, personalization, and proactive interactions, may risk further homogenizing people's interactions with these tools. We urge future research to more carefully examine how various design choices interfacing with people's offloading of different cognitive tasks can create systematic shifts from how these tasks are performed without AI.

Limitations and Future Work

We note several limitations of our study. First, owing to our study design, the informal learning task was not completely self-directed or unstructured. Nevertheless, while we

provided broad topical prompts to help initiate participants' learning, we emphasized that these prompts were only meant as guidance (see Appendix), and participants were free to explore and engage with the task in ways that felt natural and meaningful to them. Second, the extent to which people engage in cognitive offloading likely varies across individuals (O'Brien, Dickinson, and Askin 2017), learning strategies, and patterns of GenAI use. We focus on aggregate behavioral differences rather than modeling individual differences. Therefore, our findings should not be interpreted as characterizing all learners or all forms of learning. Future work should investigate the interplay between individual characteristics and information-seeking behaviors. Further, behaviors may also be influenced by the information domain. For example, in high-stakes contexts such as healthcare, information-seeking behaviors such as differentiation based on credibility or verification may be more pronounced (Bhavnani 2002). We used a diary study protocol to capture participants' in-situ seeking behaviors. We acknowledge that these accounts may represent only some of these behaviors, as participants may not have documented every instance or fully articulated the nuances of their information-seeking process. Next, through this work, we do not make any causal claims linking information-seeking behaviors to learning outcomes. We provide a speculative discussion describing how certain behaviors may have led to particular outcomes. Future work should investigate this relationship empirically. Lastly, participant attrition led to a small sample size. Nonetheless, our final sample remained within the recommended range for diary studies (Flaherty 2024).

Conclusion

People seek information not only to retrieve facts, but to acquire knowledge and develop expertise. GenAI-powered applications are increasingly mediating access to information, and, for many, replacing conventional search engines. This paradigm change involves offloading of the information-seeking process to AI. To understand how GenAI tools may reshape information seeking and impact learning, we conducted a between-subjects, longitudinal field experiment comparing participants using ChatGPT versus Google Search to pursue informal learning. We found that, by offloading much of the information selection, the ChatGPT group had diminished agency in their information seeking and took a more passive role in learning. This presents fundamental tensions with learning, which also manifested as heightened meta-cognitive demands and unsatisfactory user experience for the ChatGPT group in attempting to regain control in their learning processes. The diminished agency led to, on average, worse learning outcomes in the ChatGPT group, accompanied by two sources of distortion that systematically shifted their information access. First, ChatGPT exhibits a bias toward generating solution-oriented artifacts over principled knowledge about ‘how’ and ‘why.’ Second, the socially-oriented interaction paradigm of current GenAI tools may have unintended consequences in shifting people's information-seeking behaviors. Popular design choices such as the “chat” interface and personalization may inadvertently reduce exploration of the knowledge space.

Ethical Statement

The study design was approved by our organization's internal IRB and we obtained participant consent. Following guidelines provided by Prolific (Prolific 2025b), participants were compensated at an hourly wage of 15 USD. We paid and approved the work of everyone who completed the study, regardless of whether their responses passed our quality checks. We did not collect any personally identifiable information, except for participants' Prolific IDs for compensation and communication purposes. All communication between the research team and participants was conducted anonymously via Prolific's internal messaging system (Prolific 2025a). Participants were free to discontinue at any time without penalty. Participants were compensated on a daily basis for their contributions. All participant data were anonymized prior to analysis.

Acknowledgments

We thank our participants for their time and effort. We also thank the FATE groups at Microsoft Research Montréal and NYC, Daniel Russell, and Jiawei Zhou for thoughtful discussions and feedback.

References

- Agosti, M.; Fuhr, N.; Toms, E.; and Vakkari, P. 2014. Evaluation methodologies in information retrieval dagstuhl seminar 13441. In *Acm sigir forum*, volume 48, 36–41. ACM New York, NY, USA.
- Allam, H.; Bliemel, M.; Nassiri, N.; Toze, S.; Peet, L. M.; and Banerjee, R. 2019. A review of models of information seeking behavior. 2019 *Sixth HCT Information Technology Trends (ITT)*, 147–153.
- Allan, J.; Croft, B.; Moffat, A.; and Sanderson, M. 2012. Frontiers, challenges, and opportunities for information retrieval: Report from SWIRL 2012 the second strategic workshop on information retrieval in Lorne. In *Acm sigir forum*, volume 46, 2–32. ACM New York, NY, USA.
- Anderson, L. W.; and Krathwohl, D. R. 2001. *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives: complete edition*. Addison Wesley Longman, Inc.
- Bastani, H.; Bastani, O.; Sungu, A.; Ge, H.; Kabakci, Ö.; and Mariman, R. 2024. Generative AI can harm learning. *The Wharton School Research Paper*.
- Bates, M. J. 1989. The design of browsing and berrypicking techniques for the online search interface. *Online review*, 13(5): 407–424.
- Bhavnani, S. K. 2002. Domain-specific search strategies for the effective retrieval of healthcare and shopping information. In *CHI '02 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '02, 610–611. New York, NY, USA: Association for Computing Machinery. ISBN 1581134541.
- Bolger, N.; Davis, A.; and Rafaeli, E. 2003. Diary methods: Capturing life as it is lived. *Annual review of psychology*, 54(1): 579–616.
- Bonwell, C. C.; and Eison, J. A. 1991. *Active learning: Creating excitement in the classroom. 1991 ASHE-ERIC higher education reports*. ERIC.
- Budzyń, K.; Romańczyk, M.; Kitala, D.; Kołodziej, P.; Bugajski, M.; Adami, H. O.; Blom, J.; Buszkiewicz, M.; Halvorsen, N.; Hassan, C.; et al. 2025. Endoscopist deskilling risk after exposure to artificial intelligence in colonoscopy: a multicentre, observational study. *The Lancet Gastroenterology & Hepatology*, 10(10): 896–903.
- Chen, X.; and Feng, S. 2024. Analyzing Students' Information Behavior in Generative AI-Supported Small Group Discussions. In *Proceedings of the Eleventh ACM Conference on Learning @ Scale, L@S '24*, 325–329. New York, NY, USA: Association for Computing Machinery. ISBN 9798400706332.
- Choo, C. W.; Detlor, B.; and Turnbull, D. 1999. Information Seeking on the Web—An Integrated Model of Browsing and Searching.
- Cronin-Golomb, L. M.; and Bauer, P. J. 2023. Self-motivated and directed learning across the lifespan. *Acta Psychologica*, 232: 103816.
- Dervin, B.; et al. 1983. An overview of sense-making research: Concepts, methods, and results to date.
- Ellis, D.; Cox, D.; and Hall, K. 1993. A comparison of the information seeking patterns of researchers in the physical and social sciences. *Journal of documentation*, 49(4): 356–369.
- Engelbart, D. C. 2023. Augmenting human intellect: A conceptual framework. In *Augmented education in the global age*, 13–29. Routledge.
- Ertmer, P. A.; and Newby, T. J. 1996. The expert learner: Strategic, self-regulated, and reflective. *Instructional science*, 24(1): 1–24.
- Felix Harder. 2025a. Nutrition Certification Diploma: Diet Planning & Meal Prep. <https://www.udemy.com/course/nutrition-certification-diploma-diet-plan-meal-prep/>. [Accessed 09-11-2025].
- Felix Harder. 2025b. Nutrition Masterclass: Build Your Perfect Diet & Meal Plan. <https://www.udemy.com/course/nutrition-masterclass-build-your-perfect-diet-meal-plan/>. [Accessed 09-11-2025].
- Flaherty, K. 2024. Diary Studies: Understanding Long-Term User Behavior and Experiences. <https://www.nngroup.com/articles/diary-studies/>. [Accessed 10-09-2025].
- Ganie, S. A.; et al. 2019. Information seeking behavior of faculty members and students in colleges of Kashmir region: ellis model approach.
- Ge, X. 2010. Information-seeking behavior in the digital age: A multidisciplinary study of academic researchers. *College & Research Libraries*, 71(5): 435–455.
- Ghosh, S.; Rath, M.; and Shah, C. 2018. Searching as Learning: Exploring Search Behavior and Learning Outcomes in Learning-related Tasks. In *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval, CHIIR '18*, 22–31. New York, NY, USA: Association for Computing Machinery. ISBN 9781450349253.

- Global Stats. 2025. Search Engine Market Share Worldwide. <https://gs.statcounter.com/search-engine-market-share/>. [Accessed 09-11-2025].
- Google. 2025. Google Search. <https://search.google/ways-to-search/ai-overviews/>. [Accessed 09-09-2025].
- Handa, K.; Bent, D.; Tamkin, A.; McCain, M.; Durmus, E.; Stern, M.; Schiraldi, M.; Huang, S.; Ritchie, S.; Syverud, S.; Jagadish, K.; Vo, M.; Bell, M.; and Ganguli, D. 2025. Anthropic Education Report: How University Students Use Claude.
- Hiniker, A.; Suh, H.; Cao, S.; and Kientz, J. A. 2016. Screen Time Tantrums: How Families Manage Screen Media Experiences for Toddlers and Preschoolers. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, 648–660. New York, NY, USA: Association for Computing Machinery. ISBN 9781450333627.
- Hirvonen, N.; Jylhä, V.; Lao, Y.; and Larsson, S. 2024. Artificial intelligence in the information ecosystem: Affordances for everyday information seeking. *Journal of the Association for Information Science and Technology*, 75(10): 1152–1165.
- Kosmyna, N.; Hauptmann, E.; Yuan, Y. T.; Situ, J.; Liao, X.-H.; Beresnitzky, A. V.; Braunstein, I.; and Maes, P. 2025. Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task. *arXiv preprint arXiv:2506.08872*, 4.
- Kuhlthau, C. C. 1991. Inside the search process: Information seeking from the user's perspective. *Journal of the American society for information science*, 42(5): 361–371.
- Kumar, H.; Vincentius, J.; Jordan, E.; and Anderson, A. 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*, 1–18.
- Lee, H.-J.; Lee, J.; Makara, K. A.; Fishman, B. J.; and Hong, Y.-I. 2015. Does higher education foster critical and creative learners? An exploration of two universities in South Korea and the USA. *Higher Education Research & Development*, 34(1): 131–146.
- Lee, H.-P.; Sarkar, A.; Tankelevitch, L.; Drosos, I.; Rintel, S.; Banks, R.; and Wilson, N. 2025. The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the 2025 CHI conference on human factors in computing systems*, 1–22.
- Lee, M. K.; Kiesler, S.; and Forlizzi, J. 2010. Receptionist or information kiosk: how do people talk with a robot? In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, 31–40.
- Lira, B.; Rogers, T.; Goldstein, D. G.; Ungar, L.; and Duckworth, A. L. 2025. Learning not cheating: AI assistance can enhance rather than hinder skill development. *arXiv preprint arXiv:2502.02880*.
- Macnamara, B. N.; Berber, I.; Çavuşoğlu, M. C.; Krupinski, E. A.; Nallapareddy, N.; Nelson, N. E.; Smith, P. J.; Wilson-Delfosse, A. L.; and Ray, S. 2024. Does using artificial intelligence assistance accelerate skill decay and hinder skill development without performers' awareness? *Cognitive Research: Principles and Implications*, 9(1): 46.
- Mayo Clinic. 2024. Healthy meals start with planning. <https://www.mayoclinic.org/healthy-lifestyle/nutrition-and-healthy-eating/in-depth/healthy-meals/art-20546806>. [Accessed 09-11-2025].
- Mitra, B.; Shokouhi, M.; Radlinski, F.; and Hofmann, K. 2014. On user interactions with query auto-completion. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval*, 1055–1058.
- Moran, K.; Rosala, M.; and Brown, J. 2025. How AI Is Changing Search Behaviors. <https://www.nngroup.com/articles/ai-changing-search-behaviors/>. [Accessed 08-09-2025].
- Nugroho, A.; Andriyanti, E.; Widodo, P.; and Mutiaraningrum, I. 2025. Students' appraisals post-ChatGPT use: Students' narrative after using ChatGPT for writing. *Innovations in Education and Teaching International*, 62(2): 499–511.
- O'Brien, H. L.; Dickinson, R.; and Askin, N. 2017. A scoping review of individual differences in information seeking behavior and retrieval research between 2000 and 2015. *Library & Information Science Research*, 39(3): 244–254.
- Ohly, S.; Sonnentag, S.; Niessen, C.; and Zapf, D. 2010. Diary studies in organizational research. *Journal of Personnel Psychology*.
- OpenAI. 2022. Introducing ChatGPT. <https://openai.com/index/chatgpt/>. [Accessed 30-08-2025].
- Prolific. 2025a. How can I contact participants? <https://researcher-help.prolific.com/en/article/8a2092>. [Accessed 07-09-2025].
- Prolific. 2025b. Prolific's payment principles. <https://researcher-help.prolific.com/en/article/2273bd>. [Accessed 07-09-2025].
- Qian, Z.; Fu, J.; and Zhou, Y. 2025. Seeking a Sense of Meaning and Companionship in Life: Informal Learning on Douyin Among Chinese Older Adults. *Proc. ACM Hum.-Comput. Interact.*, 9(2).
- Qu, X.; Sherwood, J.; Liu, P.; and Aleisa, N. 2025. Generative AI Tools in Higher Education: A Meta-Analysis of Cognitive Impact. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, CHI EA '25. New York, NY, USA: Association for Computing Machinery. ISBN 9798400713958.
- Rasool, F. 2024. How to Turn Off AI Overview Results in Your Google Searches. <https://www.howtogeek.com/turn-off-ai-overview-results-in-your-google-searches/>. [Accessed 09-11-2025].
- Rieh, S. Y.; Collins-Thompson, K.; Hansen, P.; and Lee, H.-J. 2016. Towards searching as a learning process: A review of current perspectives and future directions. *Journal of Information Science*, 42(1): 19–34.
- Risko, E. F.; and Gilbert, S. J. 2016. Cognitive offloading. *Trends in cognitive sciences*, 20(9): 676–688.

- Sengupta, S. 2020. 'Learning to code in a virtual world' A Preliminary Comparative Analysis of Discourse and Learning in Two Online Programming Communities. In *Companion Publication of the 2020 Conference on Computer Supported Cooperative Work and Social Computing*, 389–394.
- Shah, C.; and Bender, E. M. 2022. Situating Search. In *Proceedings of the 2022 Conference on Human Information Interaction and Retrieval*, CHIIR '22, 221–232. New York, NY, USA: Association for Computing Machinery. ISBN 9781450391863.
- Sharma, N.; Liao, Q. V.; and Xiao, Z. 2024. Generative Echo Chamber? Effect of LLM-Powered Search Systems on Diverse Information Seeking. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24. New York, NY, USA: Association for Computing Machinery. ISBN 9798400703300.
- Shen, J. H.; and Tamkin, A. 2026. How AI Impacts Skill Formation. arXiv:2601.20245.
- Sparrow, B.; Liu, J.; and Wegner, D. M. 2011. Google effects on memory: Cognitive consequences of having information at our fingertips. *science*, 333(6043): 776–778.
- Spatharioti, S. E.; Rothschild, D.; Goldstein, D. G.; and Hofman, J. M. 2025. Effects of LLM-based Search on Decision Making: Speed, Accuracy, and Overreliance. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–15.
- Subramonyam, H.; Pea, R.; Pondoc, C.; Agrawala, M.; and Seifert, C. 2024. Bridging the gulf of envisioning: Cognitive challenges in prompt based interactions with llms. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 1–19.
- Suri, S.; Counts, S.; Wang, L.; Chen, C.; Wan, M.; Safavi, T.; Neville, J.; Shah, C.; White, R. W.; Andersen, R.; Buscher, G.; Manivannan, S.; Rangan, N.; and Yang, L. 2024. The Use of Generative Search Engines for Knowledge Work and Complex Tasks. arXiv:2404.04268.
- Tabarsi, B.; Yasir, T.; Reichert, H.; Tian, X.; Gadireddy, S.; Dimarco, C.; Briceno, D.; and Barnes, T. 2025. Herald of Advancement, Disruption, or Both: A Systematic Literature Review on Student-Facing LLM Tools in Undergraduate Computing Education.
- Tankelevitch, L.; Kewenig, V.; Simkute, A.; Scott, A. E.; Sarkar, A.; Sellen, A.; and Rintel, S. 2024. The metacognitive demands and opportunities of generative AI. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 1–24.
- Terzimehić, N.; Bühler, B.; and Kasneci, E. 2025. Conversational AI as a Catalyst for Informal Learning: An Empirical Large-Scale Study on LLM Use in Everyday Learning. *arXiv preprint arXiv:2506.11789*.
- USDA. 2020. Dietary Guidelines for Americans. https://www.dietaryguidelines.gov/sites/default/files/2021-03/Dietary_Guidelines_for_Americans-2020-2025.pdf. [Accessed 25-11-2025].
- Vaithilingam, P.; Zhang, T.; and Glassman, E. L. 2022. Expectation vs. Experience: Evaluating the Usability of Code Generation Tools Powered by Large Language Models. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI EA '22. New York, NY, USA: Association for Computing Machinery. ISBN 9781450391566.
- Walker, K.; and Vorvoreanu, M. 2025. Learning outcomes with GenAI in the classroom: A review of empirical evidence. Technical Report MSR-TR-2025-42, Microsoft.
- Wilson, T. D. 1999. Models in information behaviour research. *Journal of documentation*, 55(3): 249–270.
- Wu, S.; Liu, Y.; Ruan, M.; Chen, S.; and Xie, X.-Y. 2025. Human-generative AI collaboration enhances task performance but undermines human's intrinsic motivation. *Scientific Reports*, 15(1): 15105.
- Yang, T.-C.; Hsu, Y.-C.; and Wu, J.-Y. 2025. The effectiveness of ChatGPT in assisting high school students in programming learning: evidence from a quasi-experimental research. *Interactive Learning Environments*, 33(6): 3726–3743.