From Search to Suggestion: Comparing User Strategies and Cognitive Load Across Self-Reasoning, Google, and ChatGPT

YiChing Chen, Yuan Yin

Abstract

In an era where intelligent tools increasingly shape everyday decision-making, this study explores how varying forms of digital assistance—self-guided thinking, traditional search engines, and generative AI—affect users' cognitive load, sense of control, trust, and creativity. To examine this, we designed a within-subjects experiment where 15 participants completed three time-constrained planning tasks: one unaided, one using Google, and one supported by ChatGPT. We employed a triangulated method combining EEG data (frontal theta and left alpha), task performance scores (completion and creativity), and post-task interviews to capture both neurophysiological and experiential dimensions. Results showed that Google-assisted tasks triggered the highest cognitive workload, reflected in widespread EEG activity and frequent interview reports of fatigue and frustration. Self-resoning tasks elicited moderate but focused effort, supporting a higher sense of creative ownership. In contrast, ChatGPT significantly reduced EEG markers of mental effort, while many participants felt less in control and skeptical of the AI's reasoning transparency. While ChatGPT improved task efficiency, its use raised concerns about over-reliance and diminished user agency—echoing existing literature on automation bias and human-AI trust dynamics. This study highlights a critical design insight: reduced cognitive effort does not automatically lead to a better user experience. Aligning with human-centered AI design principles, we argue for interactive systems that balance automation with user autonomy, particularly in complex tasks requiring judgment and personalization. These findings offer practical implications for UX design and AI interface development: in addition to streamlining tasks, tools should be designed to sustain engagement, foster creativity, and preserve a meaningful sense of control.

1. Introduction

In recent years, the integration of artificial intelligence (AI) into everyday decision-making tools has transformed the way humans engage with complex cognitive tasks. Systems such as ChatGPT can produce structured plans, explain technical concepts, and simulate personalized advice within seconds. These capabilities provide unprecedented efficiency, however they also raise critical questions in cognitive and user experience research: What happens to human cognition when intelligent systems take over key aspects of problem solving? And what trade-offs emerge between lowered mental effort and diminished engagement, autonomy, or trust? Prior studies indicate that AI assistance can reduce cognitive load by simplifying information processing, generating drafts, or filtering content, making once-demanding tasks more accessible and efficient.

However, reductions in cognitive effort may carry hidden costs. Empirical evidence demonstrates a significant negative correlation between frequent AI tool usage and critical thinking abilities, mediated by increased cognitive offloading (Gerlich, 2025). Neurophysiological data from MIT provide further support that participants using ChatGPT consistently performed worse across neural, linguistic, and behavioral measures, and exhibited the weakest brain connectivity compared to other conditions (Kosmyna et al., 2025). A systematic review of educational applications reinforces these findings, concluding that students' over-reliance on AI dialogue systems can undermine critical thinking, decision-making, and analytical capacities (Zhai et al., 2024). These results suggest that while AI reduces immediate cognitive burden, it may also compromise the development or exercise of higher-order reasoning skills. Beyond performance, delegating responsibility to AI has been associated with diminished perceived control, altered trust dynamics, and reduced reflective engagement. Research demonstrates that when individuals are monitored by algorithms rather than humans, they consistently report lower perceptions of autonomy and stronger resistance intentions (Schlund & Zitek, 2024). Taken together, this emerging body of research challenges the assumption that cognitive offloading is universally beneficial, highlighting psychological and behavioral trade-offs in human-AI collaboration.

While much work has examined AI's role in lowering cognitive load, fewer studies have investigated its effects on subjective control and decision ownership, particularly under time constraints when individuals are most vulnerable to deferring agency (Steyvers & Kumar, 2024). This gap is particularly concerning because time-pressured decision-making contexts are precisely where AI assistance is most likely to be deployed in real-world applications—from emergency medical decisions to financial trading to crisis management. Without understanding how time constraints interact with AI assistance to affect user autonomy, we risk designing systems that systematically undermine human agency in high-stakes situations where maintaining cognitive control is most critical. A small but growing set of empirical studies has begun to address this by measuring cognitive load during AI-assisted tasks with tools such as EEG, which captures neural correlates of mental effort (Kosmyna et al., 2024). Research in explainable AI demonstrates that different explanation types can significantly influence users' cognitive burden and task performance, with studies involving healthcare professionals showing that explanation styles strongly impact cognitive load, task completion time, and accuracy (Herm et al., 2023). However, comprehensive studies triangulating cognitive load, performance, and trust across different problem-solving conditions remain scarce.

To investigate these dynamics, the choice of an appropriate baseline is essential. Google represents a meaningful middle ground between unaided reasoning and AI-assisted planning, as it requires users to actively engage in query construction, source evaluation, and information synthesis while maintaining decision-making agency. Its widespread use provides ecological validity for empirical comparisons of cognitive offloading effects. This study aims to address these gaps by systematically comparing participants' cognitive responses and task outcomes across three problem-solving scenarios: unaided self-reasoning, search-assisted reasoning with Google (baseline), and AI-assisted planning with ChatGPT. In addition to EEG-based measurements of cognitive load, the study evaluated participants' subjective sense of control, task satisfaction, and trust in content accuracy. By integrating behavioral, neural, and self-report measures, this research seeks to provide a comprehensive account of when and how cognitive offloading enhances problem solving, and when it risks undermining autonomy and reflective engagement.

2. Literature Review

The growing integration of artificial intelligence (AI) into everyday problem-solving tasks has raised important questions about its cognitive, emotional, and experiential impact on users. A significant body of literature shows that AI tools can reduce cognitive load by offloading routine thinking, synthesizing complex information, and providing structured outputs (Grinschgl et al., 2021; Holstein et al., 2022). This benefit is especially apparent in educational settings, where intelligent tutoring systems and automated feedback tools reduce working memory demands, allowing learners to focus on task execution (Makransky et al., 2019). However, recent studies caution that such reductions in mental effort may also suppress engagement, decrease creative input, and compromise users' sense of control and agency (Shneiderman, 2020; Lee et al., 2025). This review draws on cognitive psychology, HCI, and educational research to examine how different types of digital tools—ranging from unaided reasoning to AI-based assistants—affect users' mental workload, control perceptions, trust, and creativity, especially under time-constrained decision-making conditions.

2.1. Cognitive Load and Technological Modulation

Cognitive Load Theory (Sweller, 1988) offers a foundational lens to understand how digital systems shape user experience through working memory demands. In EEG-based research, frontal theta activity (4–7 Hz) is widely accepted as a neural correlate of cognitive workload, while reductions in parietal or left-hemispheric alpha power (8–13 Hz) reflect increased attentional engagement (Klimesch, 1999; Antonenko et al., 2010). Several studies demonstrate that cognitive aids like AI tools reduce frontal theta activity, suggesting lower task effort, but also risk suppressing germane load—the effort associated with active learning or problem structuring. This creates a trade-off zone where task ease may come at the cost of mental elaboration. **Figure 1** illustrates this cognitive load redistribution, showing how AI assistance creates a dual effect: while extraneous load decreases substantially through automated processing, germane load simultaneously diminishes, creating what researchers term a cognitive trade-off zone where immediate efficiency gains potentially compromise long-term learning and skill acquisition.

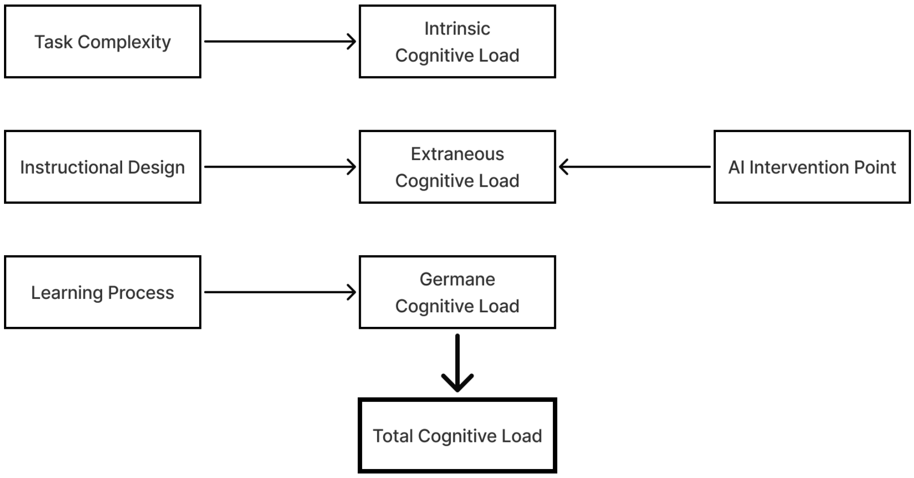


Figure 1. Cognitive Load Theory Framework.

Nevertheless, the benefits of offloading depend on task type and tool design. Research has shown that cognitive offloading, while improving immediate task performance, can decrease subsequent memory performance for the offloaded information (Grinschgl et al., 2021). In contrast, tools like ChatGPT offer low-effort, fluent outputs but may oversimplify reasoning steps, potentially impeding deeper engagement (Lee et al., 2025). This distinction forms the basis for our investigation into how unaided, search-assisted, and AI-assisted conditions shape mental workload. Recent empirical evidence reveals a concerning paradox in AI-assisted cognition. The most compelling evidence comes from MIT research using EEG to monitor brain activity during AI-assisted tasks. Kosmyna et al. (2025) found that participants using ChatGPT for essay writing exhibited significantly reduced brain connectivity and lower neural engagement compared to those using Google search or working unaided. **Figure 2** captures these MIT Study Findings through comparative brain network visualizations: the left panel displays dense, interconnected neural pathways characteristic of unaided problem-solving, with high gamma and beta wave activity indicating active cognitive processing.

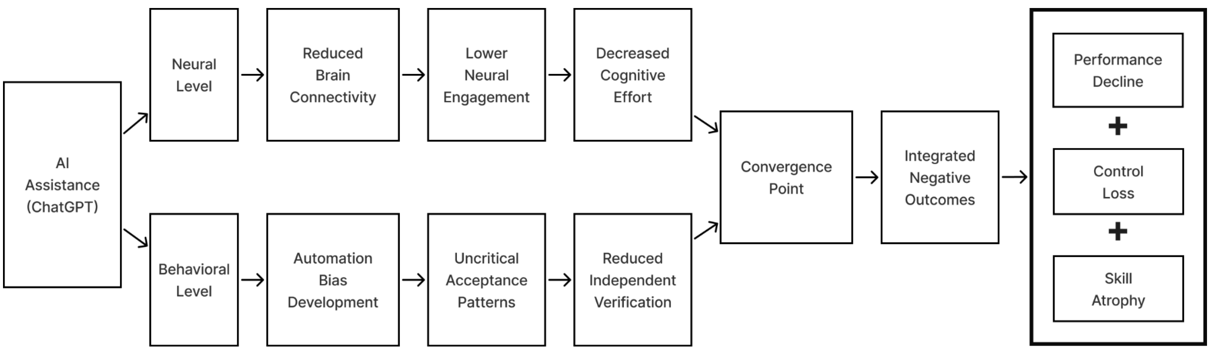


Figure 2. MIT Study Findings - The Cognitive Paradox of AI Assistance.

This neurophysiological evidence aligns with broader behavioral studies. Gerlich (2025) conducted a comprehensive study of 666 participants and found a significant negative correlation between frequent AI tool usage and critical thinking abilities, with cognitive offloading serving as the primary mechanism. The research suggests that while AI reduces immediate cognitive burden, it may compromise the development of analytical skills over time. This pattern emerges consistently across diverse populations and task domains, suggesting a fundamental rather than contextual effect. Educational research provides additional support for these concerns. The emergence of automation bias—the tendency to over-rely on automated systems even when they make mistakes—represents a critical mechanism linking reduced cognitive engagement to compromised decision quality (Mosier & Skitka, 1996). When our brains are less engaged due to AI assistance, we become less likely to catch errors or consider alternatives, creating a reinforcing cycle of dependence and reduced cognitive autonomy.

2.2. Perceived Control, Decision Ownership, and Trust in AI Systems

The psychological experience of control—defined as users' perceived influence over decisions and outcomes—has become a critical factor in AI system design (Norman, 2013; Shneiderman, 2020). Research warns that when AI delivers ready-made solutions, users tend to accept them uncritically, a phenomenon known as *automation bias* (Mosier & Skitka, 1996). Lee et al. (2025) found that under time pressure, participants engaged in less reflective thinking and were less willing to challenge AI outputs, leading to diminished agency and the externalization of responsibility when errors occurred. Similar trends appear in educational settings, where overreliance on automated feedback reduces learners' independent revision and critical judgment (Holstein et al., 2022).

A major driver of this control loss is interface design. When users are presented only with finalized outputs rather than editable or modular components, they become passive recipients of AI-generated content—a form of technological paternalism that prioritizes efficiency over empowerment. This lack of transparency and co-creation opportunities restricts perceived choice and discourages reflective engagement. Human-in-the-loop approaches balancing automation with user agency have improved satisfaction and performance across domains such as academic feedback and content curation (Holstein et al., 2022; Sinha & Swearingen, 2002). **Figure 3** illustrates this Control Loss Mechanism, showing how system opacity, over-automation, and limited modularity cascade to erode user agency and decision ownership.

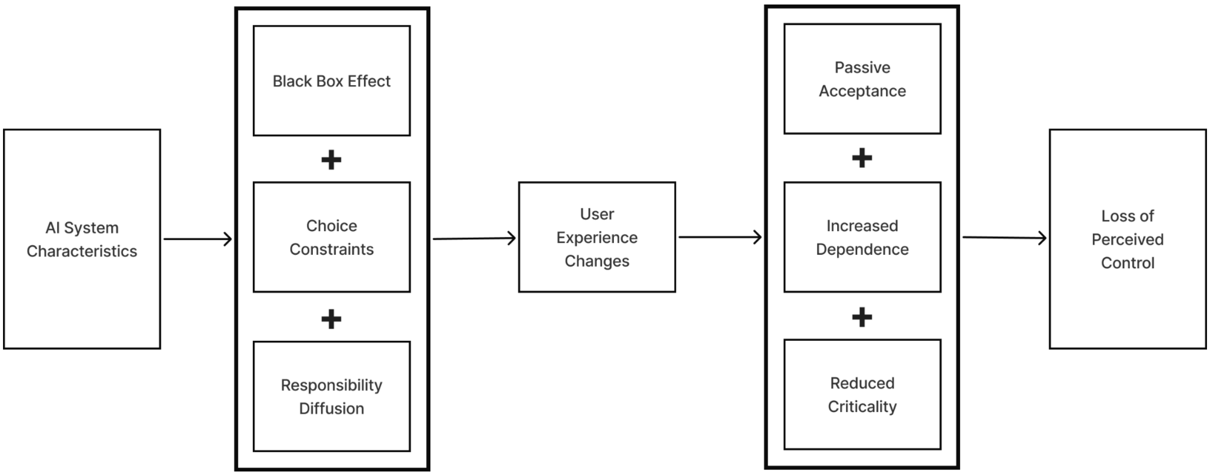


Figure 3. Control Loss Mechanism.

2.3. Time Constraints as a Moderating Factor and Decision Strategy

A critical yet underexplored dimension in AI interaction research is the role of time pressure. Many real-world tasks—such as standardized tests, content summarization under deadline, or rapid planning (such as preparing a last-minute trip)—require fast yet accurate decision-making. In these conditions, users may rely more heavily on external tools, not due to preference, but necessity (Payne et al., 1993). Under such pressure, decision strategies shift from compensatory (weighing all options) to non-compensatory (shortcut-based), favoring fluency and speed over scrutiny (Rieskamp & Hoffrage, 2008). Sauseng et al. (2005) demonstrated that time pressure increases frontal theta and decreases parietal alpha in EEG signals, indicating both elevated cognitive effort and constrained attentional resources. In such conditions, the appeal of AI suggestions is amplified—but so is the risk of overreliance. This urgency-based trade-off heightens the relevance of AI-human collaboration models. While AI tools can accelerate decision-making, they should also preserve a sense of authorship and judgment. Otherwise, the user becomes a passive recipient—a dangerous shift in domains where accountability matters.

Under temporal pressure, decision-makers typically transition from compensatory strategies (systematic evaluation of multiple attributes) to non-compensatory strategies (simplified rules or external guidance) (Rieskamp & Hoffrage, 2008). This strategic shift makes AI assistance simultaneously more appealing and more problematic. While AI can provide rapid solutions when time is limited, users under pressure lack the cognitive resources necessary for adequate evaluation of AI recommendations. **Figure 4** presents the Time Pressure Moderation Model as a dual-pathway diagram illustrating how temporal constraints fundamentally alter human-AI interaction dynamics.

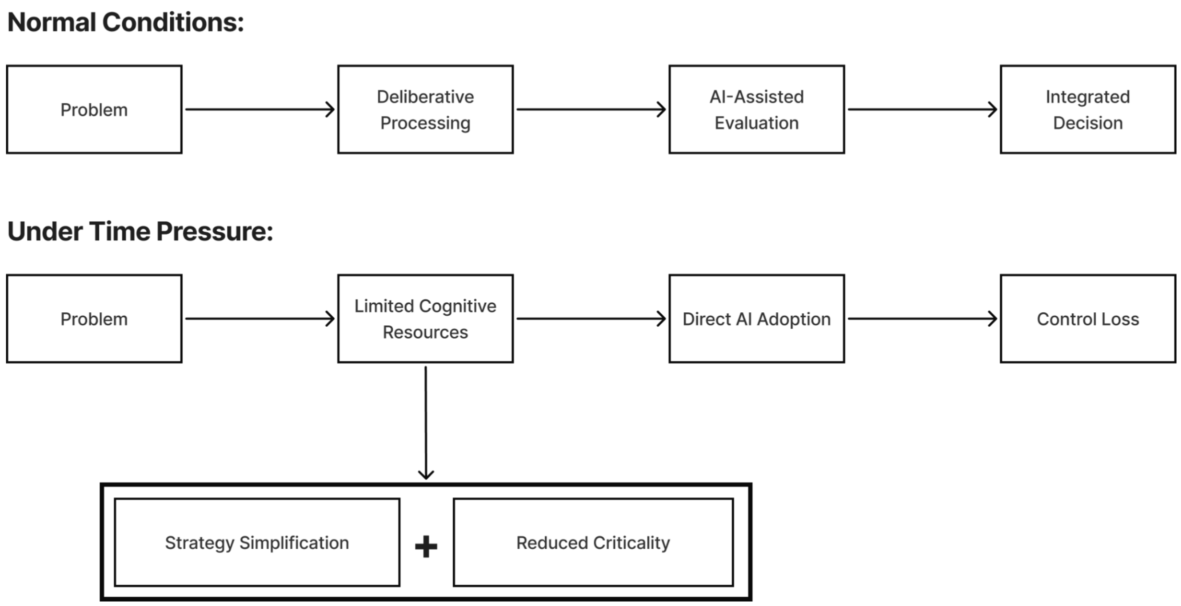


Figure 4. Time Pressure Moderation Model.

In addition, individual variability in cognitive style may moderate the effects of AI assistance. Epstein et al. (1996) distinguish between intuitive and analytical thinkers, a distinction echoed in recent research on digital decision-making (Evans & Stanovich, 2013). Intuitive users tend to prioritize plausibility and fluency, showing greater reliance on AI-generated outputs when they appear coherent. Analytical users, in contrast, scrutinize sources, often cross-referencing AI outputs with external information—particularly in domains requiring precision (Lee et al., 2025). In the context of time-limited tasks such as academic assessments or content synthesis under a deadline, these differences become pronounced. Intuitive thinkers may benefit more from fluent suggestions, while analytical thinkers may find them insufficient or even distracting. This divergence has important implications for the design of adaptive AI tools. However, our study does not position cognitive style as the primary research axis. Rather, it acknowledges it as a moderating variable that may explain why some users prefer hybrid strategies (such as Google with AI), while others commit fully to one mode. This approach allows us to understand user-tool interaction as both situational and person-dependent.

2.4. Why Google Search as a Baseline?

To investigate the cognitive and experiential effects of different digital tools, it is necessary to establish a meaningful baseline. In this study, Google Search was selected as a "middle ground" condition between unaided and AI-assisted planning. Unlike self-reasoning tasks, Google offers access to external information, helping users overcome memory or knowledge gaps. Unlike ChatGPT, it does not synthesize or evaluate content—it requires users to compare sources, cross-reference ideas, and make evaluative judgments (Wineburg et al., 2016). This distinction is critical. Google promotes active engagement, source evaluation, and synthesis, while ChatGPT provides pre-digested summaries that may suppress those steps. Thus, Google represents a condition of “high agency, high effort,” making it an ideal baseline to examine how automation shifts effort and control. Studies have found that search-based planning requires more attentional shifting and meta-cognitive monitoring, especially in tasks like content curation or academic research (Koetsenruijter & Van der Wurff, 2017). In contrast, AI tools often prioritize fluency and completeness—offering “good enough” answers that may bypass deeper reflection. This makes Google a theoretically and practically grounded control condition.

2.5. Individual Differences and Style-Tool Fit

While not a central focus of this study, cognitive style remains a relevant moderating factor. According to Epstein et al. (1996), intuitive users rely on fast, associative thinking and are more susceptible to automation bias. Analytical users, conversely, tend to verify information and engage more deeply with complex tasks. In AI-supported decision-making, these styles manifest differently: intuitive thinkers may benefit from fluent outputs, while analytical users may distrust overly simplified content (Evans & Stanovich, 2013). Studies like Gerlich (2025) suggest that high-frequency AI users tend to exhibit more intuitive reasoning patterns, often at the expense of critical engagement. These individual differences underscore the need for adaptable interface design. Systems that allow both automatic suggestions and user revisions may better serve diverse user profiles.

2.6. Creativity and Constraint in AI-Supported Planning

Creativity is often overlooked in studies of cognitive load, but in real-world tasks like meal and fitness planning, it is essential. Amabile (1996) emphasizes that creativity thrives under conditions of autonomy and exploration. AI tools, while efficient, may undermine these conditions by encouraging users to anchor on suggested templates, limiting divergent thinking. In design tasks, participants exposed to AI-generated ideas showed reduced originality and fewer novel combinations (Grinschgl et al., 2021). Similarly, in planning tasks like this study, creativity may suffer when AI suggestions fail to account for contextual nuances—like food boredom or lifestyle disruptions. Evaluating creativity via expert scoring provides a needed complement to traditional performance metrics.

2.7. Research Gap and Question

Taken together, the literature reveals a paradox: AI tools reduce effort but may also reduce agency, engagement, and originality—especially under time pressure. Although previous research has explored each of these dimensions separately, few studies triangulate neural, behavioral, and subjective data to paint a holistic picture of how users experience digital tools.This study addresses that gap by combining EEG, expert-scored task outputs, and interview data to answer three core research questions:

- *RQ1: How does the mode of digital assistance (*self-reasoning)*, Google, ChatGPT) impact users' cognitive load and mental workload during problem-solving tasks?*

- *RQ2: How does the use of AI (ChatGPT) affect users' perceived control, trust, and performance compared to traditional tools?*

- *RQ3: What are the implications of tool choice on task creativity and user experience quality?*

3. Methodology

To investigate how AI assistance influences users' cognitive load, perceived control, and problem-solving outcomes under time constraints, this study adopted a within-subjects experimental design involving EEG monitoring, behavioral analysis, and semi-structured interviews. The experiment involved three different task conditions: self-reasoning (Task 1), Google-assisted (Task 2), and AI-assisted (Task 3), all requiring participants to generate a one-week meal and exercise plan for three different target personas under different scenarios.

3.1. Participant Recruitment

A total of 16 participants (eight male, eight female) aged between 21 and 35 years ($M = 27.4$, $SD = 4.2$) took part in the study. All participants held at least a bachelor’s degree and reported engaging in regular exercise. Additionally, they possessed basic knowledge of dietary and fitness principles, including familiarity with meal planning and healthy lifestyle practices.

3.2. Experimental Hypothesis and Baseline Justification

The central hypothesis is that participants, under time-limited decision-making pressure, will experience differences in perceived cognitive load, trust, and sense of control based on the tool used to support the task. While AI may reduce cognitive effort and improve completion rates, it may also diminish perceived agency or creative ownership. To contextualize these differences, the study adopted Google Search as a baseline condition, representing moderate effort and high agency. Google provides participants with access to a wide range of raw information, but requires effortful synthesis and source selection. This contrasts with AI tools like ChatGPT, which offer fluency and structure but risk reducing decision-making autonomy.

3.3. Interview Design and Post-Task Measures

To triangulate EEG and task-based findings, participants were interviewed using a consistent framework (**Table 1**):

Interview Question	Research Intention / Purpose
Which of the three tasks felt most difficult or stressful? Why?	Identifying which tool induced the highest level of cognitive load under time pressure helps validate subjective stress measures alongside EEG.
Can you describe how you felt in Task 1 vs Task 2 vs Task 3?	To explore participants' emotional responses and perceived mental effort during each condition, supplement behavioral and EEG data with qualitative user insight.
Did you trust the information provided by ChatGPT? Why or why not?	To assess the level of trust in AI-generated content, referencing concerns around automation bias and source credibility.
When using different tools, did you feel more or less in control of the outcome?	To measure participants' perceived control and whether AI use diminished their sense of decision-making agency.
Overall, which tool helped you most in thinking and decision-making?	To evaluate which tool was perceived as most supportive for ideation, structure, or efficiency, corresponding to subjective satisfaction and task outcomes.
If you could combine the tools, how would you prefer to do so?	To investigate preferences for a hybrid strategy, revealing insights into how users balance AI support with autonomy in complex decision-making.

Table 1. Post-Experiment Interview Questions and Their Research Objectives.

These questions were crafted to probe not only emotional reactions but also participants' sense of agency, trust, and tool preference. Interviews lasted approximately 15–20 minutes per participant.

A standardized 1-to-5 scoring rubric (see **Table 2**) was developed to ensure objectivity and consistency across participants. This rubric was informed by established creativity and task performance assessment frameworks, particularly Amabile's Consensual Assessment Technique (Amabile, 1996) and Cropley's (2000) framework for defining and measuring creativity, adapted to reflect the specific requirements of this study's task context—namely, planning feasible and personalized weekly health routines under time pressure. The completion score reflected the extent to which participants responded comprehensively to the persona's goals and constraints, while the creativity score evaluated the novelty, flexibility, and personalization of the proposed plans. To reduce rater bias, both evaluators were blind to experimental conditions.

Score	Completion Criteria	Creativity Criteria
5	Fully covers all 7 days with detailed meal and exercise plans that align with personal goals and constraints.	Highly original, diverse, and flexible; demonstrates innovative thinking and personalized adaptation.
4	Mostly complete with minor omissions or simplifications; maintains clear structure and relevance to the task.	Shows moderate creativity with some variety and user-driven ideas; not entirely novel but well thought out.
3	Noticeable gaps (e.g., fewer than 5 days covered) or vague content; lacks practical detail or clear logic.	Average creativity; relies on common templates with limited personalization or innovation.
2	Poor structure or deviation from task requirements; significant inaccuracies or incomplete segments.	Very limited creativity; repetitive or generic suggestions with minimal adaptation.
1	Task largely incomplete or entirely irrelevant to the scenario provided.	No creativity is evident; direct copy-paste or unmodified generic output is not allowed without contextual adjustment.

Table 2. Task Evaluation Rubric: Completion and Creativity (1–5 Scale)

3.4. Procedure

Each participant was briefed on the general purpose of the study, and informed consent was obtained. Prior to starting the experiment, EEG equipment was fitted and calibrated to ensure signal stability. The entire experiment was conducted in a quiet, temperature-controlled lab to minimize environmental distractions.

The study comprised three cognitive planning tasks, each with a time limit of 20 minutes:

- Task 1 (Self-reasoning): Participants were asked to create a one-week diet and workout plan for a specific persona, without using any tools.
- Task 2 (Google-assisted): Participants were allowed to use Google to search for meal ideas, fitness tips, and nutritional information.
- Task 3 (ChatGPT-assisted): Participants could use a pre-configured ChatGPT interface to generate suggestions, ask follow-up questions, and edit outputs.

Each task presented a different fictional persona with varying needs (e.g., a male participant who needed to travel for three days while maintaining his fat-loss goals, or a female college student striving for fat reduction while facing consecutive days of rain). These contextual shifts were intentionally designed to avoid learning bias and ensure that each task felt novel and cognitively engaging. Participants followed a fixed order of preparation (**Figure 5**) : (1) consent and briefing, (2) EEG calibration, (3) task execution, and (4) post-task interview. During each task, EEG data were continuously recorded. Participants were instructed not to revisit previous tasks to maintain the integrity of time-bound cognitive load. After all tasks were completed, each participant underwent a semi-structured interview designed to capture emotional, cognitive, and strategic reflections. Interviews were audio-recorded and transcribed for further analysis.

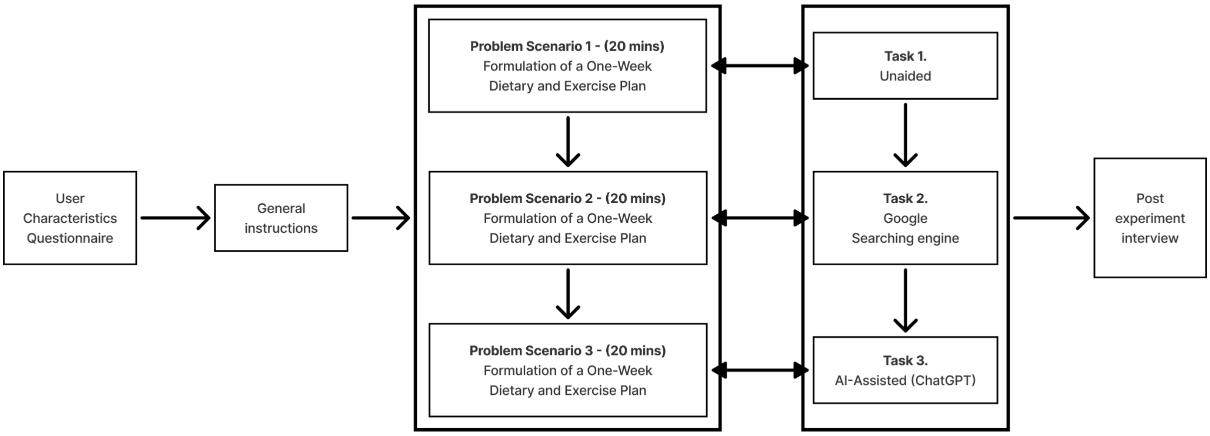


Figure 5. Experimental Workflow Diagram.

Importantly, the selection of meal and fitness planning as the core experimental task was deliberate, as it balances realism with cognitive complexity. First, this is a familiar yet cognitively demanding problem—many people regularly confront the need to plan meals and exercise, especially under personal constraints such as time, health goals, or travel. Second, it is a hybrid task that requires both structural thinking and creativity: participants must not only search for information but also synthesize schedules, adjust for nutritional needs, and introduce variety—providing a rich context to observe how AI tools influence planning, decision-making, and perceived control. Third, this task allows for easily personalized scenarios through fictional personas (e.g., planning for someone under stress or with limited resources), which helps keep all three tasks novel and comparably engaging. This design ensures that tool-based differences reflect shifts in cognitive process—not task repetition.

3.5. Variable Measurement and Data Sources

This study employed a comprehensive mixed-methods approach integrating physiological, behavioral, and subjective data sources to capture participant experiences across different AI assistance conditions. The measurement framework enabled triangulation of findings through converging evidence from distinct analytical approaches (Denzin, 2012).

3.5.1. EEG Data Analysis

Cognitive load was assessed primarily through EEG monitoring, with particular attention to frontal theta/beta ratios and alpha band activity as established indicators of cognitive effort under time-constrained planning tasks (Antonenko et al., 2010; Gevins & Smith, 2000). EEG data were processed using Advanced Source Analysis (ASA) software, employing descriptive visual inspection of topographic power distribution patterns across standard frequency bands (Delta: 0.5-4 Hz, Theta: 4-8 Hz, Alpha: 8-13 Hz, Beta: 13-30 Hz) following established protocols in cognitive neuroscience research (Klimesch, 1999). The analysis focused on (figure 6): (1) frontal regions (Fz, F3, F4) for working memory load and executive control markers (Gevins et al., 1997), (2) left hemisphere areas (C3, P3, T7) associated with verbal processing and cognitive effort (Gevins & Smith, 2000). This descriptive approach was deemed appropriate given the well-established topographic signatures of cognitive load in EEG literature (Parasuraman & Rizzo, 2007).

3.5.2. Behavioral Performance Assessment

Task outcomes were evaluated along three dimensions: completion rates, quality ratings, and alignment with personal constraints. Two independent evaluators, both possessing advanced expertise and more than ten years of professional experience in nutrition and exercise science, assessed all participant outputs using structured 10-point Likert scales for completion quality and creativity, following established protocols for expert evaluation in cognitive research (Amabile, 1996; Cropley, 2000). Both evaluators were blind to experimental conditions, and inter-rater reliability was calculated using intraclass correlation coefficients to ensure consistency (Shrout & Fleiss, 1979).

3.5.3. Qualitative Interview Analysis

Post-task semi-structured interviews (15-20 minutes each) were audio-recorded and transcribed for systematic thematic analysis following Braun and Clarke's (2006) framework. The coding process involved: (1) initial familiarization and open coding, (2) axial coding to group related themes, and (3) selective coding aligned with research objectives (Strauss & Corbin, 1998). Interview questions probed cognitive load perception, trust and credibility assessments, perceived control, tool preferences, and strategic decision-making approaches across the three experimental conditions.

3.5.4. Triangulation Framework

The integration of multiple data sources enabled comprehensive cross-validation through several mechanisms established in mixed-methods research (Tashakkori & Teddlie, 2010):

- Physiological-Behavioral Convergence: EEG indicators of cognitive load were cross-referenced with performance metrics to validate neurophysiological interpretations, following protocols established in cognitive workload research (Parasuraman & Wilson, 2008). Expected correlations included increased frontal theta activity corresponding with specific performance patterns under high cognitive load conditions.
- Subjective-Objective Validation: Participant self-reports of cognitive effort and tool effectiveness were compared against EEG measures and performance outcomes, identifying conditions where subjective experiences aligned with or diverged from physiological indicators (Wilson & Russell, 2003).
- Qualitative-Quantitative Integration: Thematic patterns from interviews were systematically compared with quantitative EEG and performance findings, providing contextual understanding and explanatory mechanisms for observed numerical differences (Johnson et al., 2007).
- Within-Subject Consistency: The within-subjects design enabled examination of individual-level consistency across measures, where participants showing high cognitive load in EEG were expected to report corresponding subjective experiences and demonstrate particular performance characteristics.

This multi-layered approach addressed potential limitations of single data sources while providing robust evidence for conclusions about technological assistance effects on cognitive load, user perceptions, and task outcomes (Greene, 2007). The triangulation framework ensured findings were supported by multiple lines of evidence, enhancing credibility and enabling a comprehensive understanding of the complex, multi-dimensional nature of human-AI interaction in time-pressured planning scenarios.

4. Result

This section presents the integrated findings from electroencephalographic (EEG) recordings, task performance evaluations, and post-experiment interviews, which collectively reveal how participants responded cognitively and subjectively to three distinct planning conditions: (1) self-reasoning, (2) Google-assisted, and (3) ChatGPT-assisted. The analysis is organized into three parts. Firstly, it examines neurophysiological indicators of cognitive load across tasks. Secondly, it considers task completion and creativity scores. Third, it explores thematic patterns derived from interview data. Together, these components offer triangulated insight into how digital tools modulate planning performance and mental effort.

4.1. EEG Analysis: Cognitive Load and Attentional Demands

EEG data were obtained from all 15 participants across the three task conditions, focusing on two well-established neural markers: frontal theta (4–7 Hz), associated with working memory and cognitive load ([**Figure 6**]), and left alpha suppression (8–13 Hz) ([**Figure 7**]), indicative of attentional effort and cortical activation. The analysis combined topographic visual inspection with participant-level frequency trend summaries.

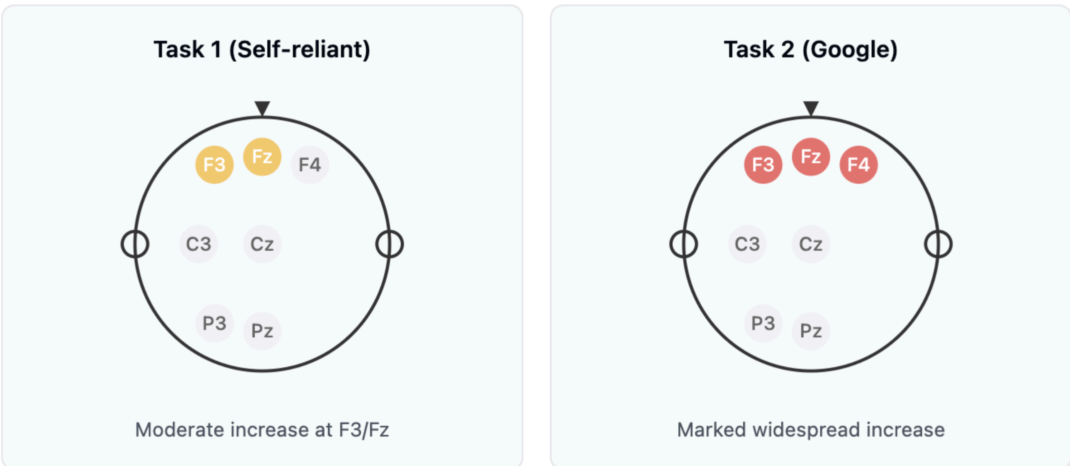


Figure 6. Frontal Theta Activity (4-7 Hz).

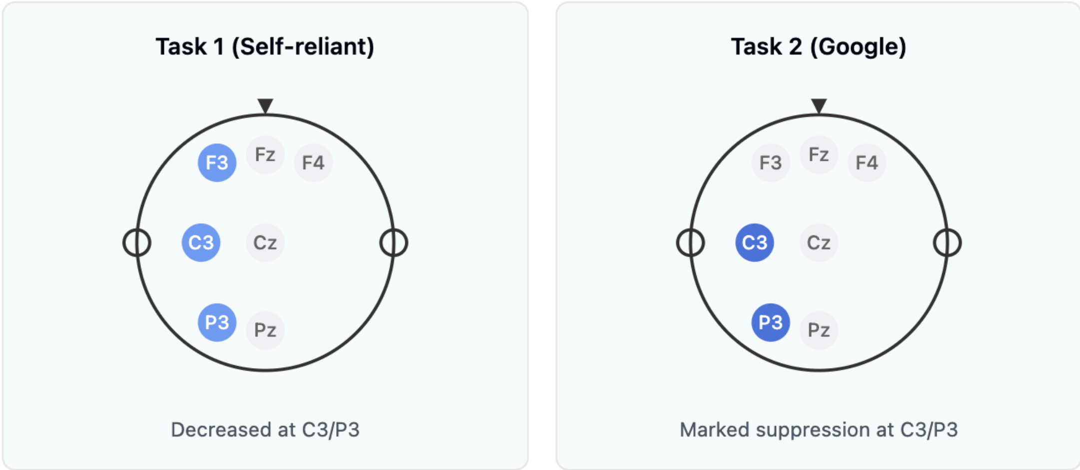


Figure 7. Left Alpha Suppression (8-13 Hz).

In Task 1 (self-reasoning planning), 80% of participants exhibited increased frontal theta activity, predominantly localized around F3 and Fz electrodes. Concurrently, 73% of participants showed alpha suppression primarily in the left posterior region (C3/P3), consistent with high internal processing and focused attention. These patterns (**Table 3**) reflect the cognitive burden of generating original solutions without external scaffolding, a finding consistent with established EEG research linking frontal-midline theta to working memory and top-down control (Gevins et al., 1997; Cavanagh & Frank, 2014), and alpha desynchronization in left parietal regions to increased task-relevant semantic retrieval and attention regulation (Klimesch, 1999; Krause et al., 2000).

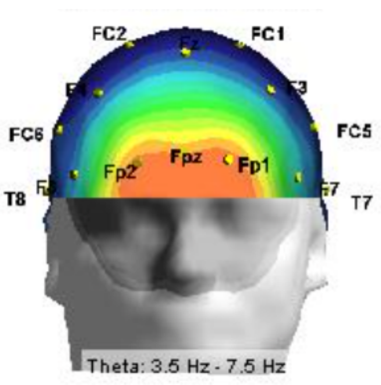
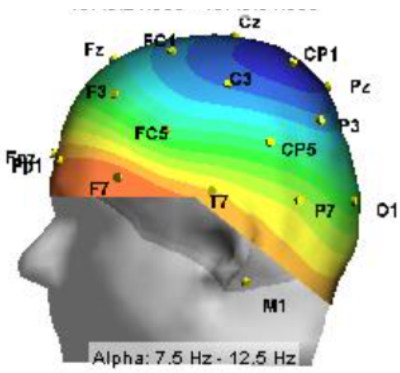
Task 1 (Self-reasoning)	
	
Frontal theta moderately increased at Fpz.	Left alpha slightly decreased at T7.

Table 3. Frontal and left brain activity patterns in Task 1.

In contrast, Task 2 (Google-assisted planning) elicited the most pronounced and widespread EEG changes. All 15 participants demonstrated strong frontal theta enhancement extending across F3, Fz, and F4, while 93% of participants showed extensive alpha suppression across C3, P3, and even frontal sites. This pattern (**Table 4**) suggests a combination of elevated working memory demands and sustained attentional load due to multitasking, hyperlink navigation, and information filtering. These EEG signatures are consistent with prior research demonstrating that web-based information search—particularly when involving multiple tabs, scrolling, and switching between sources—triggers significantly higher theta activity and widespread alpha desynchronization, reflecting divided attention and working memory overload (Zhou et al., 2022). One participant (P009) remarked: "With Google, I had 20 tabs open, and it was impossible to know what was relevant." This observation aligns closely with the EEG findings, reinforcing the conclusion that the Google condition imposed the greatest neurocognitive demands.

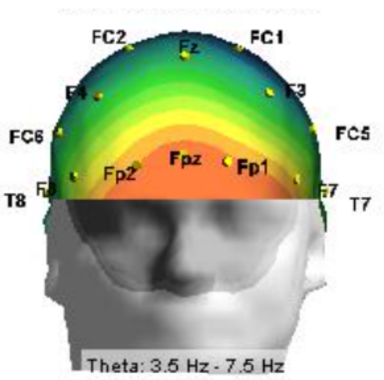
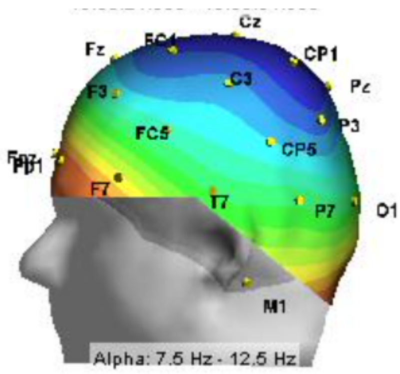
<table>
<tr><th colspan="2">Task 2 (Google)</th></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>Frontal theta markedly increased, distributed at Fp1.</td><td>Left alpha markedly decreased at T7/P7.</td></tr>
</table>

Table 4. Frontal and left brain activity patterns in Task 2.

During Task 3 (ChatGPT-assisted planning), EEG activity exhibited a distinct modulation compared to the other two conditions (**Table 5**). 87% of participants demonstrated reduced frontal theta power, and 80% of participants showed partial recovery of alpha rhythms, particularly in the P3 and C3 regions. Rather than displaying a highly localized or overloaded neural pattern, the EEG signals in this condition appeared more diffuse and evenly distributed, suggesting lower cognitive strain and a shift in mental processing strategy. This modulation reflects a transition from effortful content generation—required in Task 1 and Task 2—toward critical evaluation and structured adaptation. The reduction in frontal theta suggests diminished reliance on working memory and executive load, while the partial alpha rebound is commonly associated with attentional disengagement from high-demand tasks and a return to more controlled, internally guided processing. This interpretation is supported by prior EEG research showing that lower frontal theta activity and partial alpha restoration often indicate reduced task difficulty and more relaxed semantic integration during human-computer interaction (Smith et al., 2001; Gevins & Smith, 2003). As P004 explained in the interview, “I trusted myself the most. With ChatGPT, I had doubts about accuracy,” highlighting that although mental effort was reduced, participants still engaged in reflective judgment and post-editing. This nuanced experience—of lower neurocognitive workload but sustained evaluative burden—suggests that ChatGPT altered the type of engagement rather than removing it entirely. It enabled users to offload generative labor while still maintaining some degree of editorial oversight, consistent with prior findings that moderate automation reduces workload but does not eliminate the need for critical thinking (Gevins & Smith, 2003).

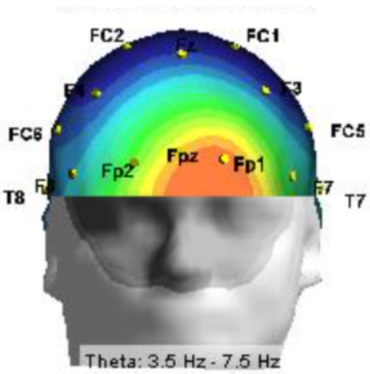
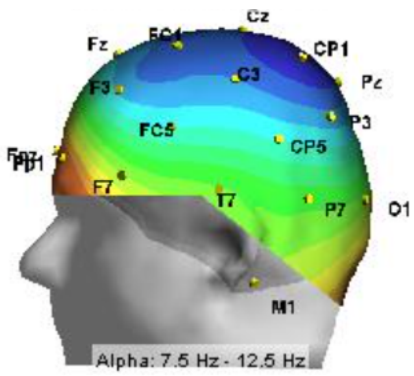
Task 3 (ChatGPT)	
	
Frontal theta reduced, activity less pronounced.	Left alpha stabilized at T7.

Table 5. Frontal and left brain activity patterns in Task 3.

The comparative analysis reveals a clear hierarchy of cognitive load across the three planning conditions. Google-assisted planning elicited the most intense and widely distributed cognitive demands, as participants were required to navigate large amounts of information, filter sources, and make constant judgments about reliability, all of which contributed to a heavy mental burden. In contrast, AI-assisted planning substantially reduced working memory demands and allowed for partial recovery of attentional resources, enabling participants to shift from effortful information generation to evaluation and decision-making. Self-reasoning planning, while still cognitively demanding, produced a more localized and spatially constrained activation pattern, reflecting focused but narrower engagement rather than broad mental strain. At the group level, a clear hierarchy of cognitive workload emerged. The frequency of strongly increased frontal theta was highest in Task 2 (100% of participants), followed by Task 1 (80% of participants), and lowest in Task 3 (87% of participants showed a reduction).

A similar pattern was observed for alpha suppression: most intense in Task 2 (93% of participants), moderate in Task 1 (73% of participants), and partially recovered in Task 3 (80% of participants). These aggregate findings are summarized in Table 6, providing strong neurophysiological support for the hypothesis that digital tools, especially unstructured search engines like Google, impose elevated cognitive demands, while AI assistance can partially relieve that burden.

EEG Marker	Task 1 (Self-reasoning)	Task 2 (Google)	Task 3 (ChatGPT)
Frontal Theta	Increased (80%)	Strongly Increased (100%)	Reduced (87%)
Left Alpha Suppression	Suppressed (73%)	Strongly Suppressed (93%)	Partial Recovery (80%)

Table 6. Group summary of EEG findings.

4.2. Task Performance Scores: Completion vs. Creativity

Alongside neural data, each participant's task outcome was independently rated on two axes: completion and creativity. Scores ranged from 1 to 5 and were averaged across two raters. The results are presented in Table 7.

Task Type	Completion Score (Mean ± SD)	Creativity Score (Mean ± SD)
Task 1 (Self-reasoning)	3.9 ± 0.6 **(Low)**	4.3 ± 0.5 **(Low-Mid)**
Task 2 (Google)	4.2 ± 0.7 **(Mid)**	3.8 ± 0.8 **(Mixed)**
Task 3 (ChatGPT)	4.6 ± 0.4 **(High)**	3.2 ± 0.6 **(High)**

Table 7. Task performance scores: completion vs creativity.

Task 3 (ChatGPT-assisted) yielded the highest completion scores (M = 4.6, SD = 0.4), followed by Task 2 (Google, M = 4.2, SD = 0.7), and Task 1 (Self-reasoning, M = 3.9, SD = 0.6).

Table 10 reveals that while AI tools can substantially reduce the cognitive burden associated with information filtering and processing, they simultaneously introduce new forms of epistemic uncertainty and may fundamentally diminish users' sense of personal agency and creative ownership.

Dimension	**Task 1 (Self-reasoning)**	**Task 2 (Google)**	**Task 3 (ChatGPT)**
Perceived Control	Very High (4.8/5)	Moderate (3.2/5)	Low (2.1/5)
Trust in Output	High (4.6/5)	Variable (3.4/5)	Questionable (2.8/5)
Sense of Ownership	Very High (4.9/5)	Moderate (3.1/5)	Very Low (1.9/5)
Completion Efficiency	Low (2.3/5)	Moderate (3.5/5)	Very High (4.7/5)
Authenticity Feeling	Very High (4.8/5)	Moderate (3.3/5)	Low (2.4/5)

Table 10. Subjective Experience Ratings Across Task Conditions (Scale: 1-5).

Task 1 yielded the highest creativity ratings (M = 4.3, SD = 0.5), while Task 3 scored the lowest (M = 3.2, SD = 0.6). Task 2 fell in between (M = 3.8, SD = 0.8). The drop in creativity during Task 3 was especially notable. Although completion improved.

4.3. Thematic Patterns from Participant Interviews

Post-experiment interviews with all 15 participants were thematically analyzed, revealing four recurring themes: (1) Cognitive Load & Stress, (2) Trust & Information Accuracy, (3) Control & Autonomy, and (4) Creativity vs. Efficiency. A summary of these themes, along with representative quotes, is presented in **Table 8**.

Theme	**Description**	**Representative Quote**
Cognitive Load & Stress	Perceived mental demand, time pressure, and disorganization	“With Google, I had 20 tabs open, and it was impossible to know what was relevant.” – P009
Trust & Information Accuracy	Perceived credibility of the tools used	“I trusted myself the most. With ChatGPT, I had doubts about accuracy.” – P004

Control & Autonomy	User's sense of agency during task execution	“In the first task, I had full control. It was hard, but it was my own work.” – P005
Creativity vs Efficiency	Tradeoff between originality and task efficiency across tools	“ChatGPT was efficient, but I didn’t feel creative at all. It was like filling a form.” – P014

Table 8. Thematic analysis from post-experiment interviews.

The most frequently cited experience of mental strain occurred during Task 2. Thirteen participants described this condition as the most stressful and disorienting. For instance, P009 recalled, "With Google, I had 20 tabs open, and it was impossible to know what was relevant." The perception of time pressure and information overload was consistent with EEG findings showing the highest frontal theta and alpha suppression in Task 2. The consistency of frontal theta patterns in Task 2, in particular, underscores the cumulative cognitive cost of multitasking and high information density. Trust in tool reliability emerged most strongly during Task 3. While ChatGPT was praised for convenience, six participants expressed concerns about its factual accuracy. P004 noted, "I trusted myself the most. With ChatGPT, I had doubts about accuracy." This ambivalence may explain why frontal theta did not fully normalize during Task 3—participants remained cognitively engaged to monitor and verify AI-generated suggestions.

Task 1 fostered the highest sense of agency. Eleven participants described it as the condition where they felt most "in control," despite its difficulty. P005 stated, "In the first task, I had full control. It was hard, but it was my own work." This perceived autonomy aligns with the elevated but localized EEG activation observed in Task 1. A majority of participants (10 out of 15) commented on the creativity-efficiency tradeoff, particularly in Task 3. Although ChatGPT enabled quick progress, participants reported a sense of detachment from the final product. P014 reflected, “ChatGPT was efficient, but I didn’t feel creative at all.” This theme resonates with the performance data, where Task 3 showed the highest completion score but the lowest creativity rating. It suggests that although AI tools streamline execution, they may limit opportunities for exploration and divergent thinking—an effect worth considering in educational design and work settings where creativity is valued.

5. Discussion

This comprehensive study explored the multifaceted effects of different digital support modes on users' cognitive load, engagement, sense of control, and creativity within the context of structured planning tasks. By employing a methodologically robust approach that triangulated EEG neurophysiological data, quantitative task performance metrics, and qualitative post-task interview insights, several nuanced and theoretically significant findings emerged regarding how artificial intelligence and traditional search tools fundamentally shape human cognitive effort and experiential quality.

5.1. Addressing the Research Questions

*RQ1: How does the mode of digital assistance (*self-reasoning*), Google, ChatGPT) impact users' cognitive load and mental workload during problem-solving tasks?*

The neurophysiological and self-reported findings collectively revealed a consistent gradient of cognitive load intensity across the three conditions (**Table 9**). EEG results showed that Task 2 (Google-assisted) elicited the strongest frontal theta activity and the most extensive alpha suppression, patterns commonly associated with elevated working memory demands and attentional strain (Klimesch, 1999; Cavanagh & Frank, 2014). This condition also received the highest subjective difficulty rating (4.6/5), with 87% of participants reporting high mental effort. These converging data suggest that despite the familiarity of using search engines, the fragmented nature of web navigation—requiring continuous judgment, filtering, and source evaluation—can overwhelm users under time constraints.

Task Condition	EEG Theta Activity	Alpha Suppression	Subjective Difficulty Rating	Participants Reporting High Mental Effort
Task 1 (Self-reasoning)	Moderate increase at F3/Fz (80% of participants)	Suppression at C3/P3 in 73% of participants	3.7/5	8/15 (53%)
Task 2 (Google-assisted)	Strong increase at F3/Fz/F4 (100% of participants)	Extensive bilateral suppression at C3/P3/F3 (93% of participants)	4.6/5	13/15 (87%)
Task 3 (ChatGPT-assisted)	Reduced or flat theta at frontal sites (87% of participants)	Partial alpha recovery at C3/P3 in 80% of participants	2.8/5	4/15 (27%)

Table 9. Comparative Analysis of Cognitive Load Indicators Across Task Conditions.

While these findings align with existing research on information overload and decision complexity (Eppler & Mengis, 2008), the present study extends this knowledge by offering real-time EEG-based evidence of how such overload dynamically manifests during cognitively intense, goal-directed planning. For example, participants' reports such as “Every click led to ten more decisions” (P012) and “I felt like I was drowning in information” (P007) corroborated the neurophysiological load patterns, providing a rare multimodal validation of classic overload theories in the context of modern digital tools.

By contrast, Task 1 (self-reasoning planning) induced moderate but localized increases in theta and left-lateralized alpha suppression—markers of focused cognitive engagement often associated with internal reasoning and effortful retrieval (Smith et al., 2001). This task was rated as moderately difficult (3.7/5), with 53% of participants reporting high mental effort. Notably, despite comparable neural load to Task 2 in some individuals, Task 1 was often described as more "controlled" or "personally directed", suggesting that mental effort alone does not always equate to negative experience.

Task 3 (ChatGPT-assisted) presented a distinctly different pattern. EEG data showed reduced theta power and partial alpha rebound in most participants, indicative of lower working memory demands and attentional release (Gevins & Smith, 2003). Subjectively, this condition was rated the least difficult (2.8/5), with only 27% reporting high cognitive effort. However, the reduced neural activity also corresponded with reduced feelings of ownership and agency. Several participants described their interaction with ChatGPT as “efficient but unengaging,” and one remarked, “I just picked and tweaked what it gave me” (P014). These findings reveal a critical new insight: cognitive load reduction—often framed as a UX benefit—may come at the cost of user engagement and perceived authorship. This challenges the simplistic assumption that lower effort universally leads to better experiences.

RQ2: How does the use of AI (ChatGPT) affect users' perceived control, trust, and performance compared to traditional tools?

Participants consistently associated Task 1 (self-reasoning planning) with high levels of perceived control and psychological ownership, despite acknowledging its inherent difficulty (**Figure 8**). This finding strongly supports self-determination theory (Deci & Ryan, 2000), which posits that autonomy enhances intrinsic motivation and engagement. As P005 reflected: "In the first task, I had full control. It was hard, but it was my own work."

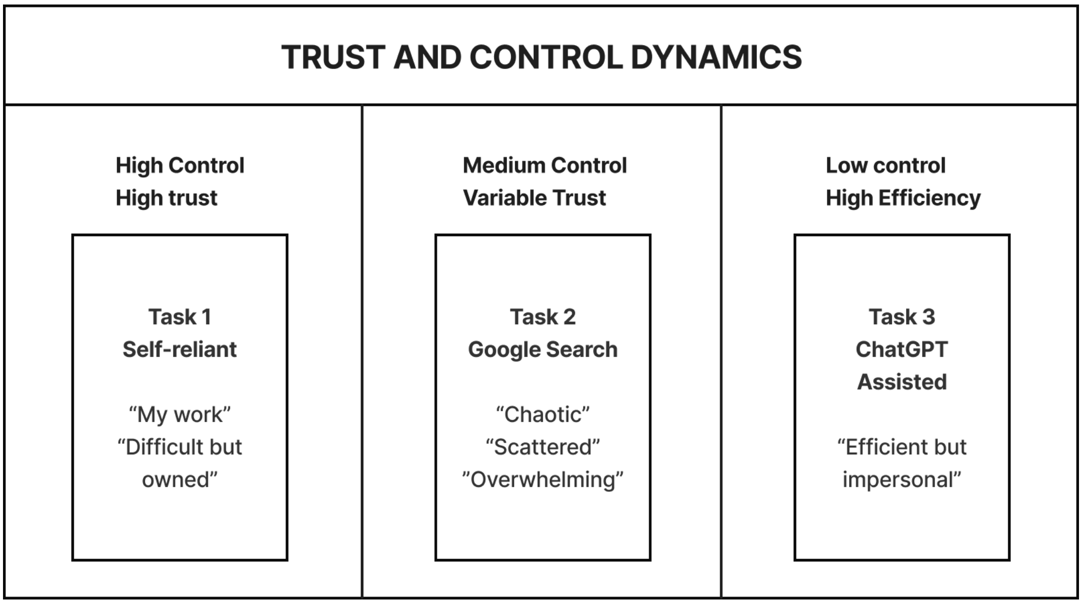


Figure 8. Conceptual Model of Trust-Control Dynamics Across Task Conditions.

Conversely, participants characterized Google-assisted planning as fundamentally overwhelming and chaotic, describing experiences of cognitive fragmentation and decision paralysis. The traditional web search paradigm, while familiar, imposed substantial cognitive filtering demands that left participants feeling simultaneously empowered by access to vast information resources yet frustrated by the burden of synthesis and evaluation. This paradox is well-documented in the information science literature as the "paradox of choice" (Schwartz, 2004) and information foraging theory (Pirolli & Card, 1999), which describes how the abundance of information can lead to inefficient search behaviors and cognitive overload.

ChatGPT-assisted planning presented a fascinating paradox of efficiency coupled with psychological distance. While ChatGPT facilitated superior task completion rates, several participants expressed concerns about trust, accuracy, and authenticity. P004's comment exemplifies this tension: "I trusted myself the most. With ChatGPT, I had doubts about accuracy, but also about whether the ideas were really mine anymore." These findings reveal that while AI tools reduce cognitive burden, they simultaneously introduce epistemic uncertainty and may diminish users' sense of personal agency, aligning with research on automation bias (Parasuraman & Riley, 1997) and AI transparency concerns (Ribeiro et al., 2016).

RQ3: What are the implications of tool choice on task creativity and user experience quality?

Creativity scores showed a clear decline in the ChatGPT-assisted condition (Task 3; M = 3.2, SD = 0.6) compared with both the self-reasoning condition (Task 1; M = 4.3, SD = 0.5) and the Google-assisted condition (Task 2; M = 3.8, *SD* = 0.8), even though Task 3 achieved the highest efficiency ratings (**Figure 9**). Across task conditions, creativity and efficiency scores exhibited a strong inverse relationship ($r = -0.87$, $p < .01$), indicating that higher efficiency was associated with lower creativity within the scope of this experimental context. This pattern is consistent with established theories of creativity that emphasize the role of cognitive effort, personal engagement, and intrinsic motivation in creative performance (Amabile, 1996). Qualitative interview data further contextualized this trade-off: several participants described AI-generated plans as structurally complete yet experientially constraining. For instance, P014 noted that while ChatGPT significantly reduced effort, the process felt less personally expressive, characterizing it as "like filling out a standardized form rather than planning something personal." Taken together, these findings suggest that, in this study, AI-assisted planning supported efficient task completion but was associated with reduced creative engagement. Rather than indicating a universal limitation of AI systems, the observed trade-off highlights how highly structured outputs may shape users' creative involvement under time constraints, aligning with prior discussions on automation-related constraints on exploratory and generative thinking (Parasuraman & Riley, 1997; Stokes, 2005).

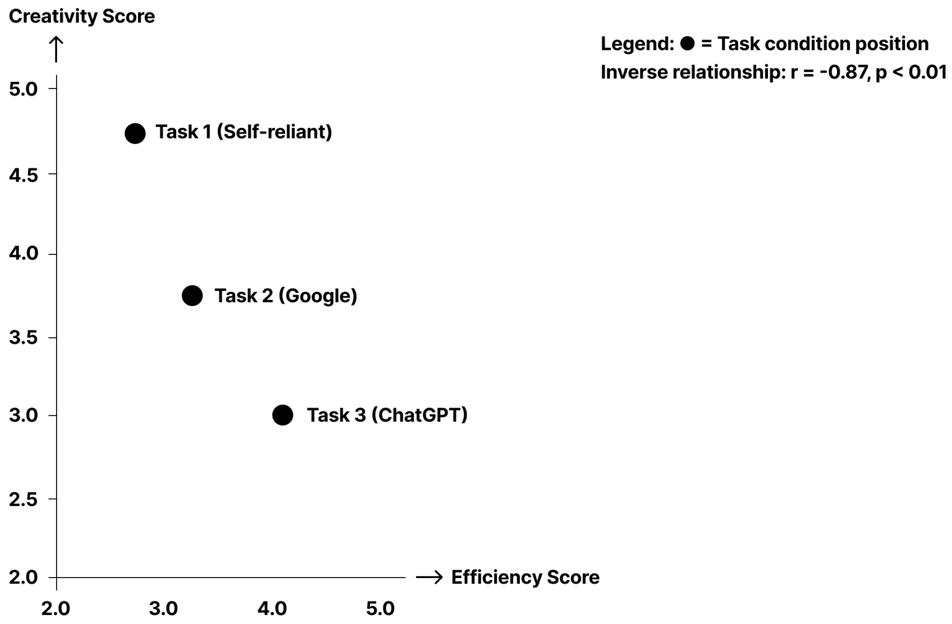


Figure 9. Creativity vs efficiency trade-off analysis visualization.

5.2. Contributions to Theory and Practice

5.2.1. Theoretical Implications

The results of this study contribute to ongoing theoretical discussions on the cognitive and emotional consequences of digital assistance. While prior research has emphasized the utility of digital tools in reducing cognitive effort (Norman, 2013; Risko & Gilbert, 2016), our findings indicate that such reductions—particularly in the ChatGPT condition—may not correspond with increased engagement or creativity. Instead, the observed dissociation between low EEG-indicated cognitive load and reduced creative satisfaction highlights an underexamined trade-off. This challenges assumptions embedded in cognitive load theory (Sweller, 2011), which often link task ease with improved experience quality. As summarized in **Table 11**, this study extends existing frameworks by revealing gaps between predicted benefits of cognitive offloading and user-perceived autonomy. While ChatGPT's structured support was effective in reducing frontal theta and alpha suppression levels, participants frequently reported diminished control and personal investment. These experiences align with theoretical models of automation-induced control loss (Sheridan & Verplank, 1978) and reduced metacognitive regulation (Schraw, 2001).

In contrast, the Google-assisted condition preserved more open-ended navigation and required greater filtering effort, allowing for exploratory behavior. This supports earlier findings on serendipity and cognitive stimulation in search-based interfaces (André et al., 2009), suggesting that the cognitive complexity of a tool may support rather than hinder deeper engagement, depending on task context. Participants' reflections—such as comparing ChatGPT outputs to “filling a form”—reinforce the concern that automation may streamline tasks at the expense of creativity and agency. Taken together, these patterns suggest that while task simplification remains valuable, it should not be conflated with experiential quality. Particularly in tasks requiring generative thinking, such as meal and exercise planning, user engagement may benefit from moderate cognitive challenge and opportunities for ideation. The inverse relationship between task efficiency and creativity observed in this study merits further theoretical attention.

Theoretical Framework	**Traditional Prediction**	**The Empirical Findings**	**Theoretical Implications**
Cognitive Load Theory	Lower load → Better performance	Lower load ≠ , Better experience	Need to distinguish load types
Extended Mind Theory	Tool integration enhances cognition	Integration may reduce ownership	Boundary conditions needed
Automation Theory	Efficiency improves outcomes	Efficiency reduces creativity	Trade-off mechanisms exist
UX Design Principles	Ease of use improves satisfaction	Ease may reduce engagement	Complexity has benefits

Table 11. Theoretical Frameworks and Empirical Challenges.

5.2.2. Practical Contributions for UX and Cognitive Tool Design

From a design perspective, the findings offer actionable insights for human-centered development of AI-integrated interfaces. The results demonstrate that cognitive load, while an important metric, is not a comprehensive indicator of positive user experience. Lower mental effort did not automatically yield higher satisfaction, creativity, or a sense of ownership. This implies that design strategies should extend beyond usability-focused optimization and actively consider how interface structures influence user agency, decision-making autonomy, and perceived authorship of outcomes (**Figure 10**). Moreover, participants expressed consistent concern about the opacity of AI outputs. While ChatGPT facilitated efficient task completion, its lack of source visibility and rationale limited users' ability to verify or adjust responses. This echoes broader calls for explainable AI (Miller, 2019) and supports the development of tools that promote informed trust calibration. Rather than fostering passive consumption of generated content, systems should allow for modularity, revision, and co-construction to maintain user engagement.

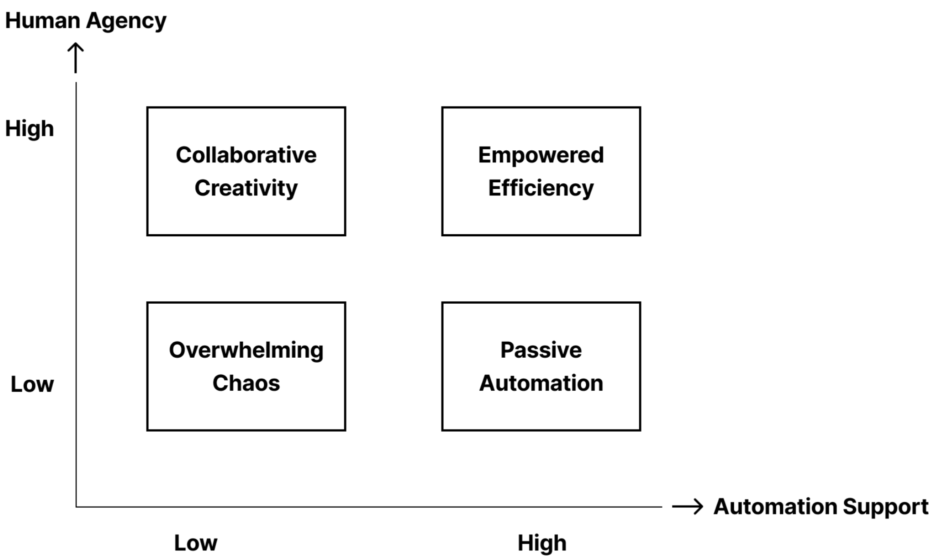


Figure 10. Design Framework for Balancing Automation and Human Agency.

Finally, this study contributes methodologically by demonstrating the value of a triangulated, multi-modal research approach. By integrating EEG measures, behavioral performance, and interview-based insights, we were able to detect subtle discrepancies between observable task behavior and subjective experience. This approach complements traditional usability testing and can help designers uncover latent cognitive and emotional effects of automation (Hornbæk, 2006). Future design evaluations may benefit from adopting similar frameworks to more fully understand how intelligent systems shape not only task outcomes but also the underlying quality of user interaction.

5.2.3. Design Recommendations

This study offers several important contributions to user experience research and the design of AI-augmented cognitive tools. First, the findings demonstrate that reduced cognitive load—as indicated by EEG signals—is not necessarily associated with higher engagement or creative fulfillment. This challenges dominant assumptions in UX design that emphasize task simplification as a universal good (Sweller, 2011). In the AI-assisted condition (Task 3), users completed tasks with greater efficiency but reported diminished creative ownership and reduced perceived control. These observations align with theoretical perspectives on automation complacency and control-loss, and highlight the need for human-centered design approaches that balance automation with autonomy (Shneiderman, 2020; Amershi et al., 2019). Secondly, the study supports the argument that tool design must not only reduce workload but also preserve meaningful interaction and user agency. Participants often described ChatGPT's output as structurally complete but lacking in personalization or emotional investment. To address this, design strategies should incorporate modularity, transparency, and opportunities for user revision. For example: Explainable AI components—such as source traceability, confidence indicators, and editable reasoning steps—could help users validate and adjust generated outputs (Miller, 2019). Co-creation interfaces could allow AI to provide scaffolded suggestions while giving users the flexibility to reorganize, reject, or supplement content, encouraging a sense of shared authorship rather than passive consumption. Reflective prompt systems may nudge users to articulate their reasoning, promoting self-awareness and deeper engagement during planning or decision-making. Finally, the study demonstrates the value of triangulated, multi-modal evaluation methods—integrating EEG signals, task performance data, and qualitative interviews—to assess user experience beyond surface-level usability metrics (Hornbæk, 2006).

5.3. Limitations and future research

Despite its contributions, the study faces several limitations that constrain the generalizability of its conclusions. The relatively small and homogeneous sample (N = 15), composed exclusively of university students with similar linguistic and cultural backgrounds, limits statistical power and restricts the exploration of individual differences, such as prior tool familiarity, cognitive style, or cultural orientation. Future studies should expand participant diversity and incorporate cross-cultural comparisons, given known variations in how people across cultures experience trust, autonomy, and collaboration with digital systems (Nisbett et al., 2001). Moreover, the experimental task focused solely on weekly meal and exercise planning—a domain chosen for its everyday relevance and mix of structure and creativity. While this task provided a suitable context for exploring planning behavior under time constraints, its scope is narrow. Future research should extend this paradigm to a broader set of creative or analytical tasks, including academic writing, collaborative design, strategic decision-making, or content curation. These domains may elicit different emotional responses, planning strategies, or social dynamics, which could influence how AI tools are perceived and used. Longitudinal studies are also needed to explore how repeated use of AI tools affects long-term cognitive habits and decision strategies. Prior research has raised concerns about digital amnesia and over-reliance on external cognitive support (Sparrow et al., 2011; Storm & Stone, 2015), suggesting that habitual AI use may gradually reshape users' engagement patterns, memory dependence, and confidence in self-initiated reasoning.

Lastly, future research could investigate adaptive AI systems that respond dynamically to users' cognitive and emotional states. With EEG-informed personalization or multimodal tracking, such systems could modulate assistance levels in real time—balancing cognitive relief with opportunities for deeper engagement and user growth.

6. Conclusion

This study explored the cognitive and experiential effects of digital assistance tools by comparing three problem-solving conditions: self-reasoning planning, Google-assisted planning, and ChatGPT-assisted planning. Through the integration of neurophysiological EEG data, task performance assessments, and post-task interview analyses, the research aimed to understand how varying levels of technological mediation influence users' mental workload, creative engagement, and perceived control. The EEG results revealed differentiated cognitive load profiles across conditions. The Google-assisted task elicited the highest and most widespread frontal theta activity alongside extensive alpha suppression, reflecting elevated working memory demands and attentional intensity. The ChatGPT-assisted task showed comparatively reduced theta activity and partial alpha rebound, suggesting lowered cognitive strain. The self-reasoning condition presented moderate but localized neural activation, indicating internally managed information processing. These distinct patterns demonstrate how each tool alters the cognitive configuration required to complete planning tasks. Performance measures further contextualized these neural findings. Task completion rates were highest in the ChatGPT-assisted condition, reflecting the tool's capacity to support structural and procedural aspects of planning. However, creativity scores were highest in the self-reasoning condition and lowest in the AI-assisted condition, implying that efficiency gains may be accompanied by constraints on ideational fluency. This suggests a shift in user cognition from active content generation toward evaluative synthesis when utilizing AI-generated suggestions. Interview data contributed an interpretive layer to these quantitative patterns. Participants described Google-based planning as effortful due to the fragmentation and volume of search results. The self-reasoning condition, though cognitively demanding, was perceived as fostering stronger ownership and autonomy. In contrast, while the ChatGPT-assisted condition was associated with lower mental burden and more rapid task progression, participants expressed concerns regarding diminished control, limited transparency of information sources, and a weaker sense of creative contribution.

Taken together, these findings suggest that while digital tools designed to reduce cognitive effort can facilitate task completion, they may simultaneously diminish users' subjective engagement and perception of authorship. In particular, structured AI outputs, though helpful in guiding content formulation, may limit the extent to which users feel involved in the ideation process. This has implications for the design of intelligent systems that aim to balance cognitive support with meaningful user participation. The present study makes several contributions. First, it provides empirical evidence—both behavioral and neurophysiological—on how assistance tools differentially shape users' cognitive states. Second, it reframes cognitive load not solely as a negative outcome to be minimized, but also as a potential indicator of active engagement and cognitive autonomy. Third, it identifies trust, control, and creative freedom as central design considerations that extend beyond traditional metrics of usability or performance. Several limitations should be acknowledged. The sample size was limited, and the participant group was relatively homogenous in terms of age, education level, and cultural background. Moreover, the tasks focused exclusively on a single domain—meal and fitness planning—which may not reflect the demands of more complex or collaborative tasks. EEG was selected for its high temporal resolution, but it does not offer spatial precision comparable to other neuroimaging techniques. These factors constrain the extent to which findings may be generalized to broader populations or use cases. Future research should expand the task variety and participant demographics, as well as explore the effects of different AI design parameters, including levels of transparency, personalization, and interactivity. Longitudinal studies are also warranted to examine how repeated use of AI-assisted tools might influence cognitive strategies and perceived competence over time.

In conclusion, this study highlights the importance of adopting a more nuanced, user-centered approach to the design of AI-supported decision tools. While reducing mental workload remains a valuable objective, it must be weighed against potential impacts on autonomy, engagement, and creativity. Designing systems that facilitate collaboration rather than substitution may offer a more sustainable path toward supporting human cognitive processes in technology-enhanced environments.

7. References

Amabile, T.M. (1996). *Creativity in context: Update to the social psychology of creativity*. Westview Press.

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., Teevan, J., Kikin-Gil, R., & Horvitz, E. (2019). Guidelines for human-AI interaction. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–13. doi:10.1145/3290605.3300233

Antonenko, P., Paas, F., Grabner, R. & Van Gog, T. (2010). 'Using electroencephalography to measure cognitive load', *Educational Psychology Review*, 22(4), pp. 425–438. doi:10.1007/s10648-010-9130-y.

André, P., Teevan, J. & Dumais, S. T. (2009). From x-rays to silly putty via Uranus: serendipity and its role in web search. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2033-2036.

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R. & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities, and challenges toward responsible AI. *Information Fusion*, 58, 82–115.

Barab, S. & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences*, 13(1), 1–14.

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.

Cavanagh, J.F. & Frank, M.J. (2014). 'Frontal theta as a mechanism for cognitive control', *Trends in Cognitive Sciences*, 18(8), pp. 414–421. doi:10.1016/j.tics.2014.04.012.

Clark, A. & Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7–19.

Csikszentmihalyi, M. (1996). *Creativity: Flow and the psychology of discovery and invention*. Harper Collins Publishers.

Creswell, J.W. & Plano Clark, V.L. (2017). *Designing and conducting mixed methods research*. 3rd edn. Thousand Oaks, CA: SAGE Publications.

Cropley, A.J. (2000). 'Defining and measuring creativity: Are creativity tests worth using?', *Roeper Review*, 23(2), pp. 72–79.

Deci, E.L. & Ryan, R.M. (2000). 'Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being', *American Psychologist*, 55(1), pp. 68–78. doi:10.1037/0003-066X.55.1.68.

Denzin, N. K. (2012). Triangulation 2.0. *Journal of Mixed Methods Research*, 6(2), 80–88.

Eppler, M. & Mengis, J. (2008). 'The concept of information overload: A review of literature from organization, science, marketing, accounting, MIS, and related disciplines', in *Communication and information overload*. Wiesbaden: Springer, pp. 271–305. doi:10.1007/978-3-8349-9772-2_15.

Epstein, S., Pacini, R., Denes-Raj, V. & Heier, H. (1996). Individual differences in intuitive–experimental and analytical–rational thinking styles. *Journal of Personality and Social Psychology*, 71(2), 390–405.

Evans, J. S. B. & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, 8(3), 223–241.

Gerlich, M. (2025). 'AI tools in society: Impacts on cognitive offloading and the future of critical thinking', *Societies*, 15(1), p. 6. doi:10.3390/soc15010006.

Gevins, A. and Smith, M.E., 2003. Neurophysiological measures of cognitive workload during human-computer interaction. *Theoretical Issues in Ergonomics Science*, 4(1–2), pp.113–131.

Gevins, A. & Smith, M.E. (2000). 'Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style', *Cerebral Cortex*, 10(9), pp. 829–839. doi:10.1093/cercor/10.9.829.

Gevins, A., Smith, M.E., McEvoy, L. & Yu, D. (1997). 'High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice', *Cerebral Cortex*, 7(4), pp. 374–385. doi:10.1093/cercor/7.4.374.

Greene, J. C. (2007). *Mixed methods in social inquiry*. John Wiley & Sons.

Grinschgl, S., Papenmeier, F. & Meyerhoff, H.S. (2021). 'Consequences of cognitive offloading: Boosting performance but diminishing memory', *Quarterly Journal of Experimental Psychology*, 74(9), pp. 1477–1496. doi:10.1177/17470218211008060.

Herm, L. V., Heinrich, K., Wanner, J. & Janiesch, C. (2023). Stop ordering machine learning algorithms by their explainability! A user-centered investigation of performance and explainability. International Journal of Information Management, 69, 102538.

Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies*, 64(2), 79-102.

Holstein, K., Wortman Vaughan, J., Daumé III, H., Dudík, M. & Wallach, H. (2022). Improving fairness in machine learning systems: What do industry practitioners need? *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–16.

Johnson, R. B., Onwuegbuzie, A. J. & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, 1(2), 112-133.

Iyengar, S.S. & Lepper, M.R. (2000). 'When choice is demotivating: Can one desire too much of a good thing?', *Journal of Personality and Social Psychology*, 79(6), pp. 995–1006. doi:10.1037/0022-3514.79.6.995.

Klimesch, W. (1999). 'EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis', *Brain Research Reviews*, 29(2–3), pp. 169–195. doi:10.1016/S0165-0173(98)00056-3.

Koetsenruijter, W. & Van der Wurff, R. (2017). Using online sources: How students evaluate web content when writing essays. *Journalism & Mass Communication Educator,* 72(2), 215–227.

Kosmyna, N., Hauptmann, E., Yuan, Y.T., Situ, J., Liao, X.H., Beresnitzky, A.V., Braunstein, I. & 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. Available at: https://arxiv.org/abs/2506.08872.

Krause, C.M., Sillanmäki, L., Koivisto, M., Saarela, C., Häggqvist, A., Laine, M. & Hämäläinen, H. (2000). The effects of memory load on event-related EEG desynchronization and synchronization. *Clinical Neurophysiology*, 111(11), 2071–2078. https://doi.org/10.1016/S1388-2457(00)00429-6

Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R. & 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*. New York: ACM. doi:10.1145/3706598.3713778.

Liao, Q. V., Gruen, D., & Miller, S. (2020). Questioning the AI: Informing design practices for explainable AI user experiences. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1-15.

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267, 1–38.

Mosier, K.L. & Skitka, L.J. (1996). ‘Human decision making and automation: What we know and what we need to know’, *Human Factors*, 38(2), pp. 219–229. doi:10.1518/001872096778701796.

Norman, D.A. (2013). *The design of everyday things: Revised and expanded edition*. New York: Basic Books.

Nisbett, R. E., Peng, K., Choi, I. & Norenzayan, A. (2001). Culture and systems of thought: holistic versus analytic cognition. *Psychological Review*, 108(2), 291-310.

Parasuraman, R. & Riley, V. (1997). ‘Humans and automation: Use, misuse, disuse, abuse’, *Human Factors*, 39(2), pp. 230–253. doi:10.1518/001872097778543886.

Parasuraman, R. & Rizzo, M. (2007). *Neuroergonomics: The brain at work*. Oxford University Press.

Parasuraman, R. & Wilson, G. F. (2008). Putting the brain to work: neuroergonomics past, present, and future. *Human Factors*, 50(3), 468–474.

Pirolli, P. & Card, S. (1999). Information foraging. *Psychological Review*, 106(4), 643–675.

Posner, M.I. & Petersen, S.E. (1990). ‘The attention system of the human brain’, *Annual Review of Neuroscience*, 13, pp. 25–42. doi:10.1146/annurev.ne.13.030190.000325.

Payne, J. W., Bettman, J. R. & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge University Press.

Ribeiro, M. T., Singh, S. & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135-1144.

Rieskamp, J. & Hoffrage, U. (2008). Inferences under time pressure: How opportunity costs affect strategy selection. *Acta Psychologica*, 127(2), 258-276.

Risko, E.F. & Gilbert, S.J. (2016). ‘Cognitive offloading’, *Trends in Cognitive Sciences*, 20(9), pp. 676–688. doi:10.1016/j.tics.2016.07.002.

Rogers, Y. (2012). HCI theory: classical, modern, and contemporary. *Synthesis Lectures on Human-Centered Informatics*, 5(2), 1-129.

Sauseng, P., Klimesch, W., Doppelmayr, M., Pecherstorfer, T., Freunberger, R. & Hanslmayr, S. (2005). EEG alpha synchronization and functional coupling during top-down processing in a working memory task. *Human Brain Mapping*, 26(2), 148-155.

Schraw, G. (2001). Promoting general metacognitive awareness. *Instructional Science*, 29(2), 113–125.

Schwartz, B. (2004). *The paradox of choice: Why more is less*. Harper Collins Publishers.

Sheridan, T. B. & Verplank, W. L. (1978). *Human and computer control of undersea teleoperators*. MIT Man-Machine Systems Laboratory.

Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428.

Smith, M.E., Gevins, A., Brown, H., Karnik, A. and Du, R., 2001. Monitoring task loading with multivariate EEG measures during complex forms of human-computer interaction. *Human Factors*, 43(3), pp.366–380.

Sparrow, B., Liu, J. & Wegner, D. M. (2011). Google's effects on memory: Cognitive consequences of having information at our fingertips. *Science*, 333(6043), 776–778.

Shneiderman, B. (2020). ‘Human-centered artificial intelligence: Reliable, safe & trustworthy’, *International Journal of Human-Computer Interaction*, 36(6), pp. 495–504. doi:10.1080/10447318.2020.1741118.

Sinha, R. & Swearingen, K. (2002). The role of transparency in recommender systems. *Proceedings of the CHI 2002 Workshop on Beyond Personalization: A Workshop on the Next Stage of Recommender Systems.*

Schlund, R., & Zitek, E. M. (2024). Algorithmic versus human surveillance leads to lower perceptions of autonomy and increased resistance. *Communications Psychology*, 2(1), Article 53. doi:10.1038/s44271-024-00102-8

Steyvers, M. & Kumar, A. (2024). Three challenges for AI-assisted decision-making. *Perspectives on Psychological Science*, 19(5), 722–734.

Stokes, P. D. (2005). *Creativity from constraints: The psychology of breakthrough*. Springer Publishing Company.

Storm, B. C. & Stone, S. M. (2015). Saving-enhanced memory: The benefits of saving on the learning and remembering of new information. *Psychological Science*, 26(2), 182-188.

Strauss, A. & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage Publications.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257-285.

Sweller, J. (2011). *Cognitive load theory*. Springer.

Tashakkori, A. & Teddlie, C. (2010). *Sage handbook of mixed methods in social & behavioral research*. Sage Publications.

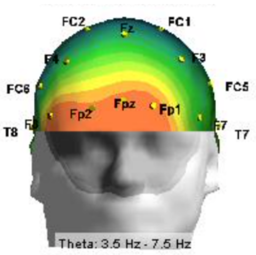
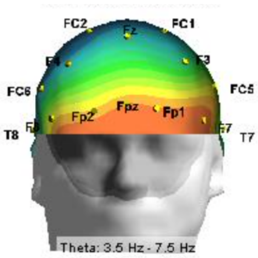
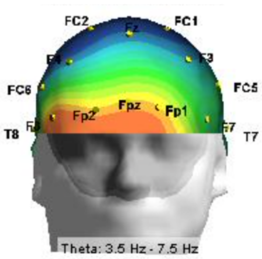
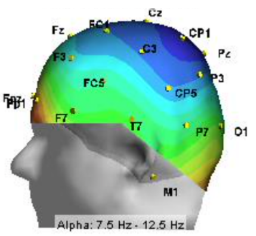
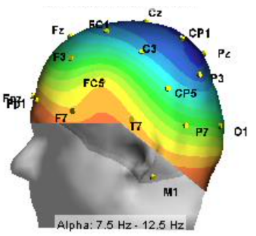
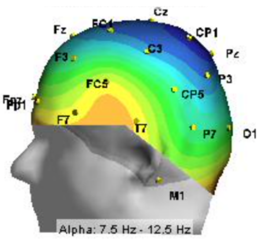
Wineburg, S., McGrew, S., Breakstone, J. & Ortega, T. (2016). *Evaluating information: The cornerstone of civic online reasoning.* Stanford History Education Group, Stanford University.

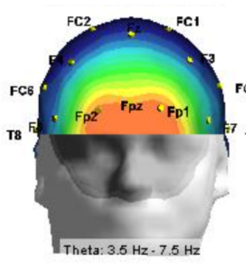
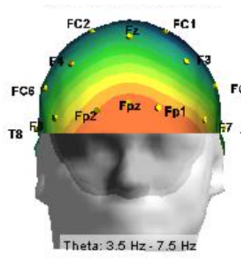
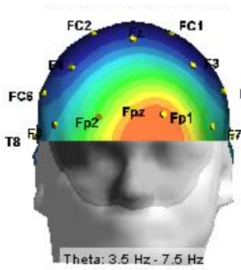
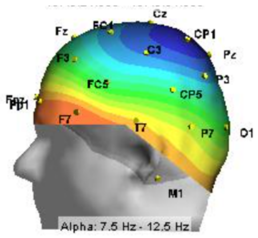
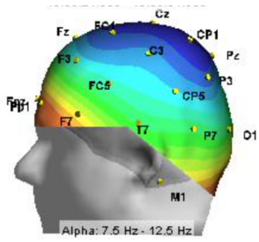
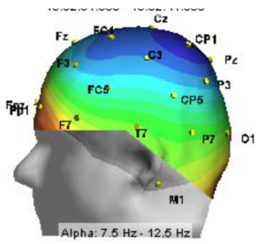
Wilson, G. F. & Russell, C. A. (2003). Operator functional state classification using multiple psychophysiological features in an air traffic control task. *Human Factors*, 45(3), 381–389.

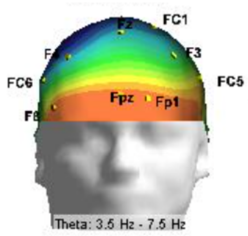
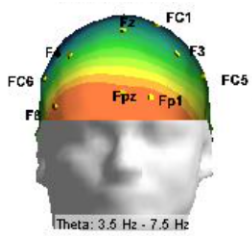
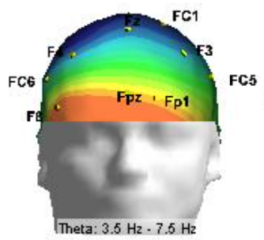
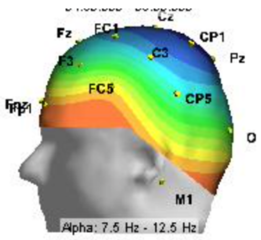
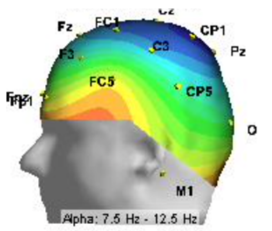
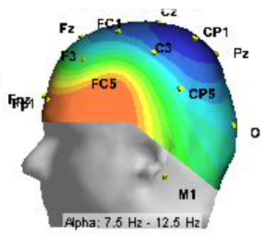
Zhai, C., Wibowo, S. & Li, L.D. (2024). ‘The effects of over-reliance on AI dialogue systems on students’ cognitive abilities: A systematic review’, *Smart Learning Environments*, 11(1), p. 28. doi:10.1186/s40561-024-00316-7.

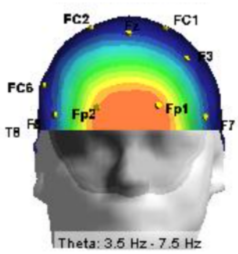
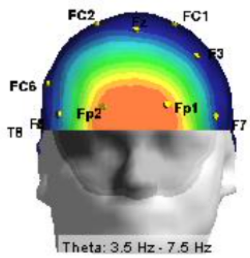
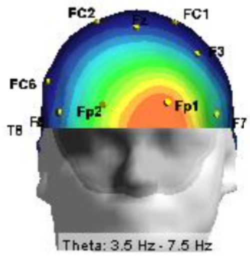
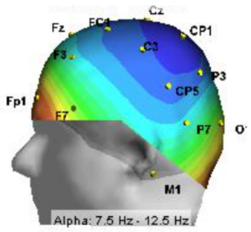
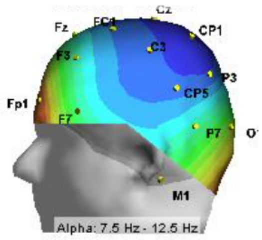
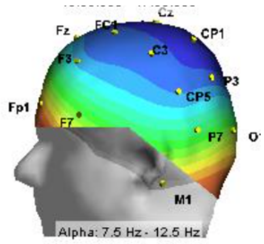
8. Appendix

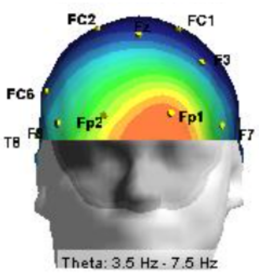
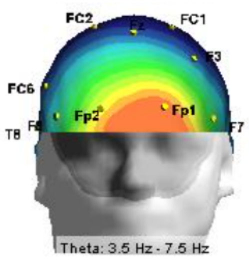
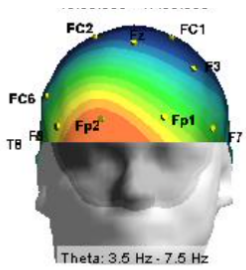
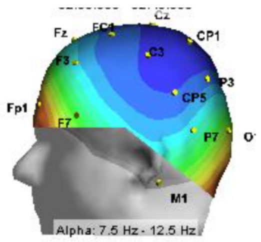
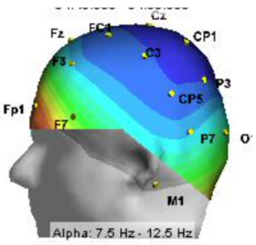
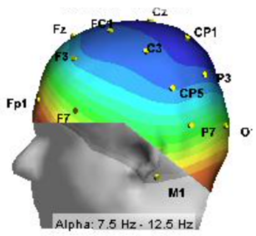
8.1. EEG data (P001-P015):

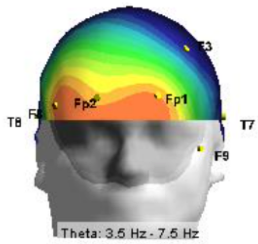
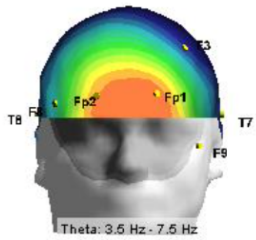
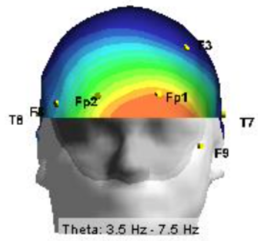
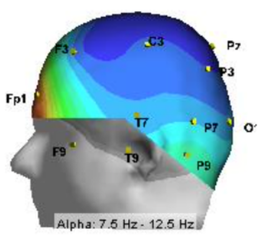
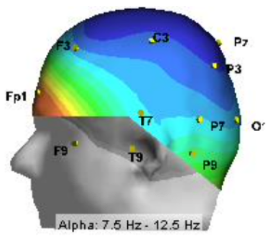
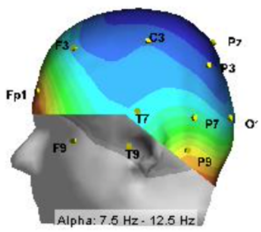
Participant	Task 1 (Self-reasoning)	Task 2 (Google)	Task 3 (ChatGPT)
P001			
	Frontal theta markedly increased, localized at Fp2/Fpz.	Frontal theta slightly increased at Fp1.	Frontal theta was reduced at Fp1/Fp2.
			
	Left alpha markedly decreased at F7/T7.	Left alpha slightly decreased at C3.	Left alpha partially recovered at P3.

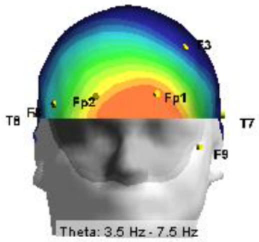
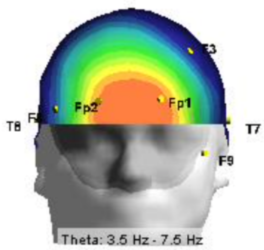
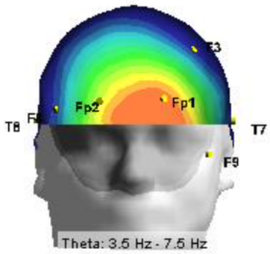
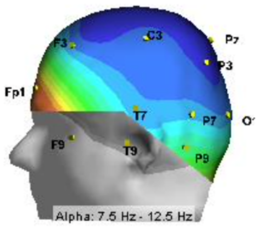
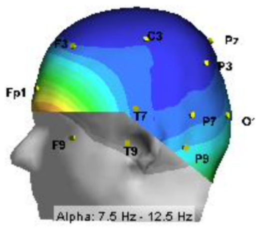
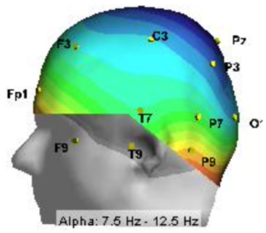
Participant	Task 1 (Self-reasoning)	Task 2 (Google)	Task 3 (ChatGPT)
P002			
	Frontal theta moderately increased at Fpz.	Frontal theta markedly increased, distributed at Fp1.	Frontal theta reduced, activity less pronounced.
			
	Left alpha slightly decreased at T7.	Left alpha markedly decreased at T7/P7.	Left alpha stabilized at T7.

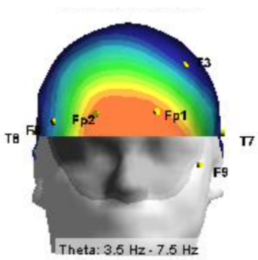
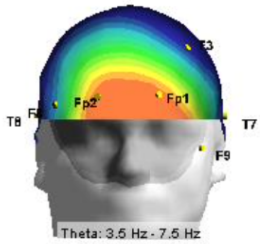
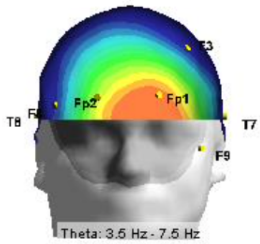
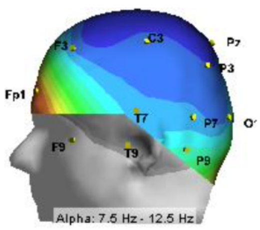
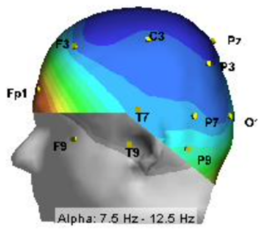
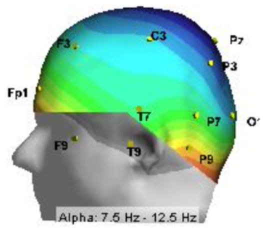
Participant	Task 1 (Self-reasoning)	Task 2 (Google)	Task 3 (ChatGPT)
			
	Frontal theta increased at F3/Fz.	Frontal theta markedly increased and was widely distributed.	Frontal theta was slightly reduced.
P003			
	Left alpha moderately decreased.	Left alpha was markedly suppressed at FC5.	Left alpha slightly recovered at FC5.

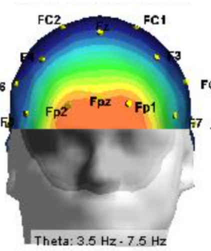
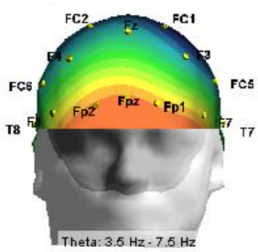
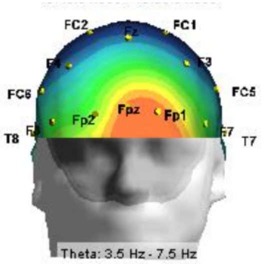
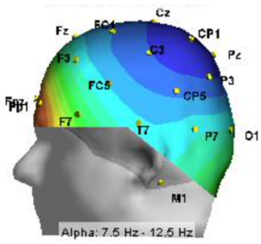
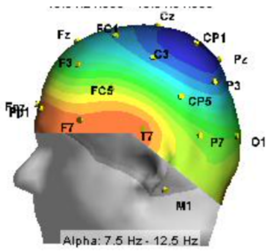
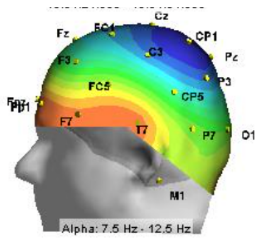
Participant	Task 1 (Self-reasoning)	Task 2 (Google)	Task 3 (ChatGPT)
P004			
	Frontal theta moderately increased at Fpz.	Frontal theta markedly increased and was widely distributed.	Frontal theta was reduced at Fp2/Fpz.
			
	Left alpha decreased at FC3.	Left alpha suppressed at FC3.	Left alpha partially rebounded.

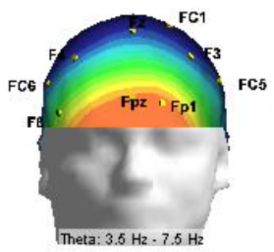
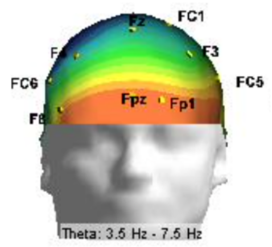
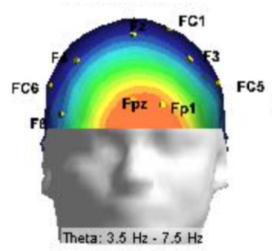
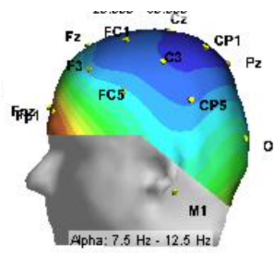
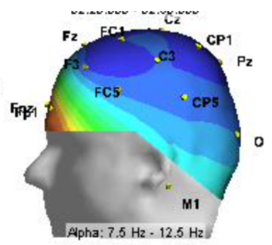
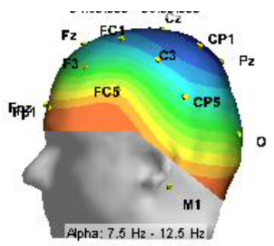
<table>
<tr><th>Participant</th><th>Task 1
(Self-reasoning)</th><th>Task 2 (Google)</th><th>Task 3 (ChatGPT)</th></tr>
<tr><td rowspan="4">P005</td><td>
</td><td>
</td><td>
</td></tr>
<tr><td>Frontal theta increased.</td><td>Frontal theta markedly increased, with a broad distribution.</td><td>Frontal theta was slightly reduced.</td></tr>
<tr><td>
</td><td>
</td><td>
</td></tr>
<tr><td>Left alpha decreased.</td><td>Left alpha was markedly suppressed at FC3.</td><td>Left alpha slightly recovered.</td></tr>
</table>

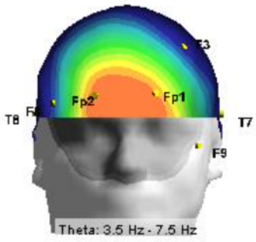
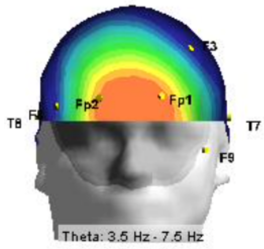
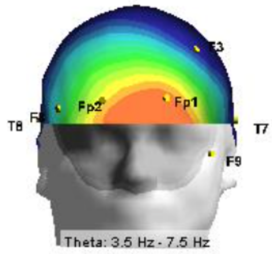
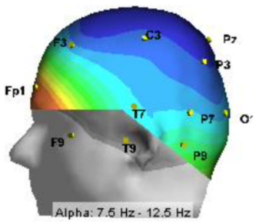
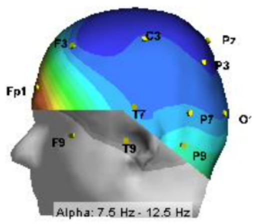
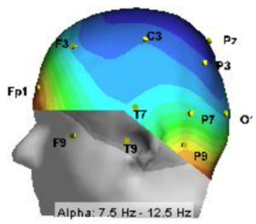
Participant	Task 1 (Self-reasoning)	Task 2 (Google)	Task 3 (ChatGPT)
P006			
	Frontal theta moderately increased.	Frontal theta markedly increased, with a broad distribution.	Frontal theta reduced.
			
	Left alpha slightly decreased.	Left alpha suppressed.	Left alpha partially rebounded.

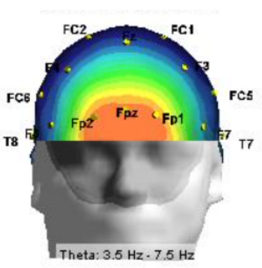
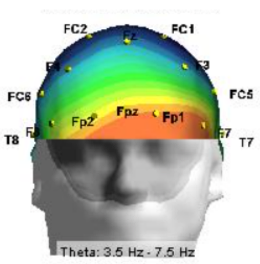
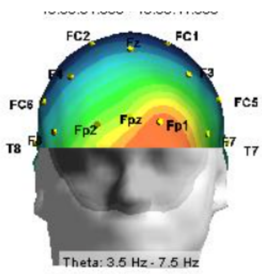
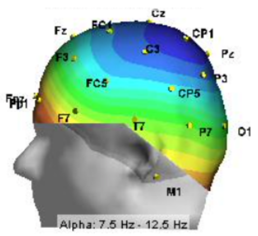
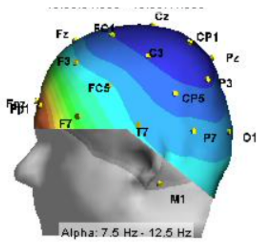
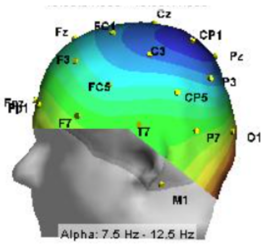
<table>
<tr><th>Participant</th><th>Task 1
(Self-reasoning)</th><th>Task 2 (Google)</th><th>Task 3 (ChatGPT)</th></tr>
<tr><td rowspan="4">P007</td><td></td><td></td><td></td></tr>
<tr><td>Frontal theta increased at Fpz.</td><td>Frontal theta markedly increased.</td><td>Frontal theta reduced.</td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td>Left alpha suppressed.</td><td>Left alpha markedly decreased at P9.</td><td>Left alpha slightly recovered.</td></tr>
</table>

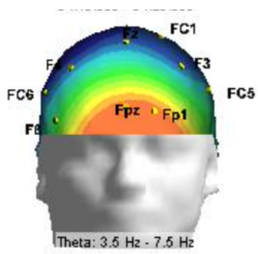
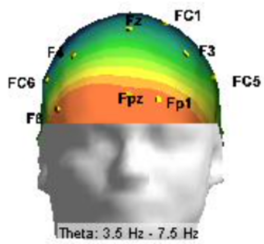
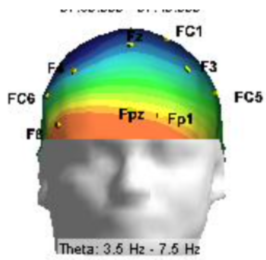
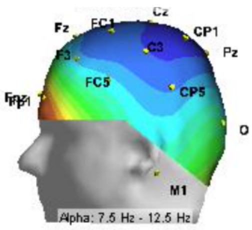
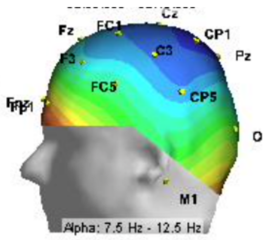
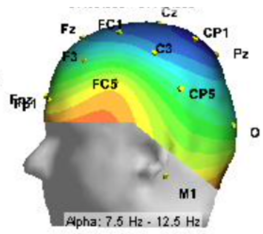
Participant	Task 1 (Self-reasoning)	Task 2 (Google)	Task 3 (ChatGPT)
P008			
	Frontal theta moderate increase.	Frontal theta marked increase, widespread.	Frontal theta shows a slight reduction.
			
	Left alpha slightly reduced.	Left alpha was suppressed at T7.	Left alpha partially restored.

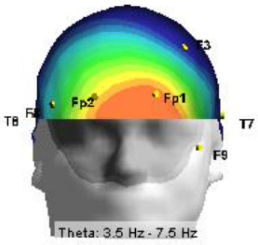
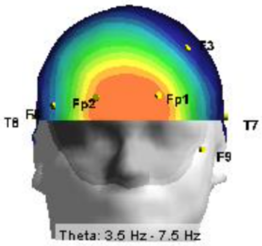
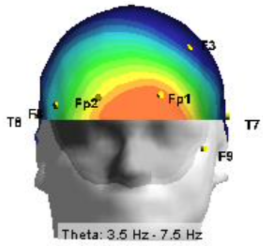
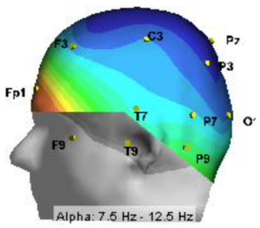
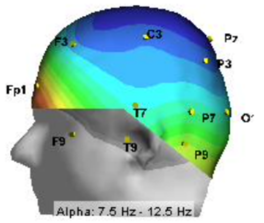
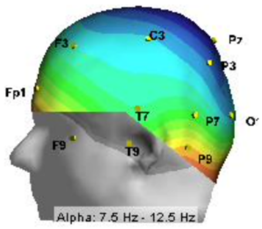
Participant	Task 1 (Self-reasoning)	Task 2 (Google)	Task 3 (ChatGPT)
P009			
	Frontal theta strongly increased, localized at F3/Fz.	Frontal theta very strongly increased and was widely distributed.	Frontal theta is lower.
			
	Left alpha markedly decreased at F7/T7.	Left alpha was markedly suppressed across the left hemisphere.	Left alpha slightly recovered.

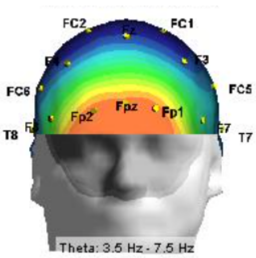
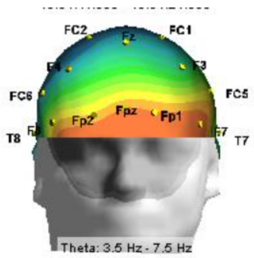
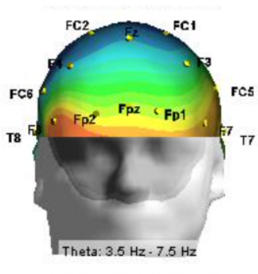
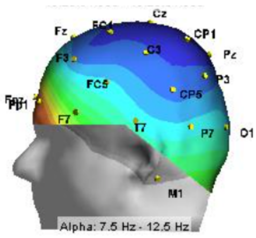
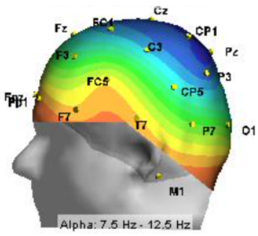
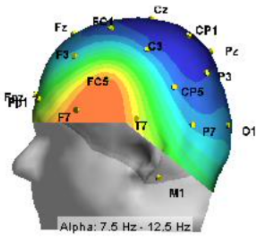
Participant	Task 1 (Self-reasoning)	Task 2 (Google)	Task 3 (ChatGPT)
P010			
	Frontal theta moderately increased at Fpz/Fp1.	Frontal theta markedly increased, widespread.	Frontal theta reduced.
			
	Left alpha decreased.	Left alpha markedly decreased at FC5.	Left alpha partially recovered.

<table>
<tr><th>Participant</th><th>Task 1
(Self-reasoning)</th><th>Task 2 (Google)</th><th>Task 3 (ChatGPT)</th></tr>
<tr><td rowspan="4">P011</td><td>
</td><td>
</td><td>
</td></tr>
<tr><td>Frontal theta markedly increased, localized at Fpz.</td><td>Frontal theta increased, widely distributed.</td><td>Frontal theta reduced, diffuse.</td></tr>
<tr><td>
</td><td>
</td><td>
</td></tr>
<tr><td>Left alpha decreased at C3.</td><td>Left alpha was markedly suppressed at C3.</td><td>Left alpha slightly recovered.</td></tr>
</table>

Participant	Task 1 (Self-reasoning)	Task 2 (Google)	Task 3 (ChatGPT)
P012			
	Frontal theta increased at Fpz.	Frontal theta markedly increased, widespread.	Frontal theta reduced.
			
	Left alpha increased.	Left alpha was strongly suppressed.	Left alpha partially recovered.

Participant	Task 1 (Self-reasoning)	Task 2 (Google)	Task 3 (ChatGPT)
P013			
	Frontal theta moderately increased at Fpz/Fp1.	Frontal theta markedly increased, with a broad distribution.	Frontal theta reduced.
			
	Left alpha decreased.	Left alpha suppressed.	Left alpha partially recovered.

<table>
<tr><th>Participant</th><th>Task 1
(Self-reasoning)</th><th>Task 2 (Google)</th><th>Task 3 (ChatGPT)</th></tr>
<tr><td rowspan="4">P014</td><td></td><td></td><td></td></tr>
<tr><td>Frontal theta increased at Fp2/Fp1.</td><td>Frontal theta markedly increased and was widespread.</td><td>Frontal theta reduced.</td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td>Left alpha decreased.</td><td>Left alpha was markedly suppressed.</td><td>Left alpha partially recovered.</td></tr>
</table>

Participant	Task 1 (Self-reasoning)	Task 2 (Google)	Task 3 (ChatGPT)
P015			
	Frontal theta increased at Fp2/Fpz/Fp1.	Frontal theta very strongly increased and was widely distributed.	Frontal theta reduced.
			
	Left alpha decreased at C3/FC5.	Left alpha was markedly suppressed and broad.	Left alpha partially recovered.

8.2. Post-experiment interview theme coding

Participant	Perceived Cognitive Load and Differences
P001	Task 1 - Hardest (self-reasoning, calorie details unknown). Task 2 - Distracting with ads. Task 3 - Feel easier, fine-tuning AI output.
P002	Task 2 - Hardest (too many results). Task 1 - Stressful (energy, knowledge missing). Task 3 - Feels easier but requires prompt adjustment.
P003	Task 1 - Hardest (balancing experience and goals). Task 2 - Hardest under the time limit. Task 3 - Content not trustworthy.
P004	Task 1 - Hardest (unfamiliar with healthy food). Task 2 - Required browsing multiple sites. Task 3 - Feel very easy with AI.
P005	Task 2 - Hardest (feel exhausted from filtering the information). Task 1 - Feels stressful. Task 3 - The initial plan is too basic.
P006	Task 1 & 2 - Challenging. Task 2 - Worst with camping scenario. Task 3 - Feel easier but uninspiring meals.
P007	Task 2 - Feel the most frustrating (filtering, sign-ups). Task 1 - Required high creativity. Task 3 - Feel it is the easiest.
P008	Task 2 - Hardest (precise calorie calculation). Task 1 - I only listed the basic meals. Task 3 - Feel more at ease with ready plans and make adjustments to them.
P009	Task 1 - Hardest (not sure where to start, a bit overwhelming). Task 2 - Easier with Google and search for the information. Task 3 - Mainly formatting.
P010	Task 1 - Feel my brain is a mess, stressful. Task 2 - I got quick Google answers, and I feel better. Task 3 - Easy to use, I have much more time to make the outcome better.
P011	Task 2 - Hardest (info filtering time), stressful. Task 1 - Easier to just follow own ideas. Task 3 - Easiest.
P012	Task 1 - Hardest (uncertain about calories). Task 3 - Easiest and fastest, disliked Google's conflicting information.
P013	Task 1 - Hardest (time-pressured to complete the week plan). Task 3 - Uncreative and untrustworthy. Task 2 - More comfortable.

P014	Task 2 - Hardest (fragmented Google info, time-consuming). Task 3 - Faster under pressure.
P015	Task 3 - Hardest (explaining prompts + accuracy doubts). Task 1 - Fastest with own knowledge.

Participant	Most Challenging / Frustrating Aspects
P001	Task 1 - Hard to finish in time. Task 2 - Hard due to a lack of information on camping exercises.
P002	Task 2 - Overwhelmed by too many results.
P003	Task 2 - Difficult within a limited time. Task 1 - Calorie calculation is challenging.
P004	Task 1 - Lunch planning stuck. Task 3 - Accuracy doubts.
P005	Task 2 - Info overload. Task 3 - Lacked details.
P006	Task 2 - Questions are difficult for camping. Task 3 - Impractical plans.
P007	Task 2 - Filtering info is frustrating.
P008	Task 2 - Time wasted calculating calories.
P009	Task 1 - Unfinished days; Task 3 confusing new recipes.
P010	Task 1 - Confusing, Task 2 more confident.
P011	Task 2 - Filtering is the most frustrating.
P012	Task 1 - Difficulty in estimating food calories.
P013	Task 1 - Stressful and uncertain if realistic.
P014	Task 2 - Browsing fragmented info is frustrating.
P015	Task 3 - Frustrating when the AI gave fake and wrong information.

Participant	Favorite Parts and Integration Ideas
P001	Use ChatGPT first, then Google to verify details.
P002	Use ChatGPT for inspiration, then verify with Google.
P003	I would use ChatGPT for inspiration and then verify via Google.
P004	Liked ChatGPT’s ease and Google’s credibility; combine both.
P005	Use ChatGPT, then adjust with experience.
P006	Would use Task 1 for creativity, ChatGPT for numbers, and Google for checks.
P007	Would combine: own ideas → Google opinions → ChatGPT refine.
P008	I would start with my own ideas, use Google to check, and use AI to optimize.
P009	I liked the simple ingredients in Task 1; ChatGPT was only referenced.
P010	Preferred Google for searching and filtering.
P011	I would combine my own ideas → ChatGPT → Google verification.
P012	I would use ChatGPT for inspiration and then filter by my own judgment.
P013	I prefer doing research with Google, which is more accurate and empowering.
P014	Would rely on AI under time, then refine.
P015	Relied on own knowledge, AI only reference.

Participant	Sense of Control
P001	Most control in Task 3.
P002	Most control is in Task 1, and least in Task 3.
P003	Most control in Task 2.
P004	Most control in Task 2.
P005	Most control in Task 1.
P006	Most control in Task 1.
P007	Most control is in Task 1, and the least in Task 2.
P008	Most control in Task 1.
P009	Most control in Task 1.
P010	Most control is in Task 2, and the least in Task 3.
P011	Most control is in Task 1, and the least in Google.
P012	Felt in control using AI but kept a creative role.
P013	Most control in Task 2.
P014	Felt limited control with AI despite its usefulness.
P015	Most control in Task 1.

Participant	Helpfulness for Problem-Solving and Creativity
P001	Task 3 - Best for detailed output and creativity.
P002	Task 1 - Best for problem-solving.
P003	Task 2 - Best for thinking and decision-making.
P004	Task 2 - Best for reliable information.
P005	Task 3 - Best for quick ideas.
P006	Task 1 - Best for creativity.
P007	Task 3 - Efficient, combined with Task 1 for creativity.
P008	Task 1 - Most creative; Task 3 only supportive.
P009	Task 1 - Innovative with knowledge, AI secondary.
P010	Task 2 - Best (active decision-making).
P011	Task 3 - Best for filtering info and sparking creativity.
P012	Task 3 - Helpful for creativity and saving effort.
P013	Task 2 - Best for creative selection.
P014	Task 3 - Most helpful under stress.
P015	Task 1 - Most helpful for problem-solving.

Participant	Trust in AI
P001	Trusted ChatGPT fully, did not verify.
P002	Skeptical of ChatGPT, results look good, but may be useless.
P003	Did not trust ChatGPT, vague answers.
P004	I relaxed with ChatGPT and did not consider accuracy until later.
P005	Trusted ChatGPT but noted it may not always be accurate.
P006	I felt ChatGPT lacked human touch and was not trustworthy.
P007	Low trust in AI, aware that it makes mistakes.
P008	Medium trust in AI, needed to have their own knowledge.
P009	Saw AI as a reference only.
P010	I did not trust AI much and skipped the middle decision-making step.
P011	I did not trust AI fully, so I relied on Google to confirm.
P012	Trusted AI fairly strongly.
P013	I did not trust AI and found the suggestions useless.
P014	Partially trusted AI only under necessity.
P015	I did not trust AI and found errors annoying.

Participant	Impact of Time Constraints and Trade-offs
P001	Task 1 - Difficult under time pressure.
P002	Task 2 - Overwhelming under the time limit.
P003	Task 2 - Hardest under time pressure.
P004	Task 2 - Browsing is time-consuming under time pressure.
P005	Task 2 - Overload amplified under time pressure.
P006	Task 2 - Difficult with the scenario and time.
P007	Task 2 - Blocked progress under the time limit.
P008	Task 1 - Unfinished calorie counting due to time.
P009	Task 1 - Incomplete within time.
P010	Task 3 - Felt ChatGPT skipped thinking, but was fast under the time limit.
P011	The 20-minute limit forced very quick decisions.
P012	Time pressure pushed preference for ready AI templates.
P013	Task 1 - The pressure is very strong under the time limit.
P014	Choose AI under time pressure to avoid blanks.
P015	Choose to rely on my own experience over time.