FRONTIERS IN
ENGINEERING | SCIENCE | MANAGEMENT
MNS-UET, MULTAN

FRONTIERS IN ENGINEERING, SCIENCE & MANAGEMENT

https://journal.mnsuet.edu.pk/

Research Paper

Thinking Outsourced: How Generative AI Reduces Cognitive engagement, Memory, and Originality in Knowledge Work

Zaid Sarfraz [a, *], Zain Sarfraz Dogar [b], Attique ur Rehman [b], Asim Ali Rao [a], Muhammad Yasir Bilal [a], Abdul Basit [a]

[a] MNS University of Engineering and Technology, Multan 60000, Pakistan

[b] National College of Business Administration and Economics (NCBA&E), Multan 60000, Pakistan

ARTICLE INFORMATION

Received	01 Aug 2026
Received in revised form	31 Aug 2026
Accepted	12 Sep 2026
Available Online	15 Sep 2026

*CORRESPONDING AUTHOR

Zaid Sarfraz
MNS University of Engineering and Technology, Multan 60000, Pakistan.

Email zaid.sarfraz@mnsuet.edu.pk

ABSTRACT

Purpose: The rapid integration of generative artificial intelligence (GenAI) in organizational workflows has raised concerns on the influence of technology on human cognition. These LLMs may enhance efficiency but may also influence how a person thinks, learns and remembers information. This study investigates the impact of Generative AI (GenAI) on cognitive engagement, memory retention, originality and critical reasoning during knowledge-intensive writing tasks. The study also presents a framework for responsible AI integration within organizations.
Design/methodology/approach: The study used a sample of 90 participants and randomly assigned them into 3 groups i.e., AI-assisted group, search-engine group, and independent group. The participants completed 450-500 words essay writing task on organizational risks of relying on an automated decision system. Cognitive engagement of the participants was measured using keystroke based behavioural index, memory retention through recall tests, originality and critical reasoning via structured expert rubrics.
Findings: The results showed clear differences at group level as participants that freely used GenAI were unable to recall most of what they had written. Most of the participants felt that they had outsourced thinking to the LLM system and acted more as editor rather than authors.
Originality: This study provides experimental evidence comparing AI-assisted, search engine assisted and independent writing under controlled conditions while integrating quantitative insights. It further proposes a practical framework for healthy AI integration that emphasizes AI literacy, cognitive engagement protocols and continuous organizational monitoring to ensure AI augments rather than substituting human cognition.

Keywords: Generative AI (GenAI), Cognitive Engagement, Memory erosion, Short-term memory

1. Introduction

AI-based tools are rapidly being embedded in our daily organizational workflows. Generative AI (GenAI) is a class of AI models capable of generating novel synthetic content, including text, audio, and other digital media by learning underlying patterns and characteristics of training data and producing outcomes in response to user or operators prompts (Autio et al., 2024). GenAI platforms like Gemini, DeepSeek, ChatGPT and others influence our communication, brainstorming, content creation, analytics and decision-making, solving at a scale that is unprecedented. While these platforms enhance productivity, growing research and public opinion have raised concerns about their long-term cognitive impact on the human mind. The concerns that technology weakens the mind are ancient. In Plato's Phaedrus, Socrates warned that writing could erode memory and give the appearance of wisdom without its substance. Later, people feared that the printing press would flood readers with books and reduce disciplined study. Similar claims met the telephone, calculators, and the internet. Ideas such as "digital amnesia" and "Google makes us stupid" reflect this pattern. Research suggests these claims are overstatements. Tools change how we remember and think, but broad declines in intelligence or memory are not supported. The more accurate view is that technologies shift cognitive load and skill demands (Kosmyna et al., 2025). Yet, recent research increasingly suggests that while AI improves efficiency, it may simultaneously reduce cognitive engagement, memory formation, and independent reasoning capacities essential to knowledge work, leadership, and organizational performance (Gerlich, 2025). It was observed that heavy AI dependence may reduce cognitive engagement (Kosmyna et al., 2025; McBain, 2025; Sparrow et al., 2011; Tucker, 2025). This effect is not inherent to AI technologies themselves but arises from patterns of passive, habitual reliance that reduce the need for independent thought, memory retrieval, or problem-solving. The current study combines findings from experimental study, modern cognitive science, human-AI interaction tools and digital tool research to examine whether GenAI usage shifts individuals towards passive thinking and reduced cognitive engagement. Additionally, it presents new experimental insight comparing feedback from AI-assisted writing with traditional and search engine-based research writing to quantify difference in cognitive engagement, originality and critical reasoning.

2. Literature Review

2.1. Cognitive Offloading and Human-AI Interaction

Cognitive offloading refers to the process of using external tools to reduce the cognitive load on an individual's working memory (Risko and Gilbert, 2016). It frees up cognitive resources but may also lead to a decline in skill development and cognitive engagement. Historically, offloading has expanded human cognitive capacity; writing, calculators, and digital devices all serve this purpose. (Juaban et al., 2025) emphasized that offloading has both productive and detrimental forms. When offloading replaces rather than supporting internal processing, it can weaken core cognitive skills and reduce learning depth.

(Gerlich, 2025) observed that GenAI adds a more extreme form of offloading as delegating work not only impacts memory or retrieval but a generation of thoughts including argumentation and reasoning structures. Additionally, it was estimated that frequent AI tool users exhibit a lower critical thinking score due to their habitual reliance on AI for idea generation and synthesis.

Figure 1 - Algorithm for Thematic Coding Procedure.

Another research showed that people often overtrust AI output, which reduces scrutiny and weakens analytical engagement (Romeo and Conti, 2026).

2.2. Digital Memory Effect and Reduced Internalization

(Sparrow et al., 2011) presented the Google effect in their study, which is also called digital amnesia. It is the tendency to forget information that can be found readily online by using Internet search engines. According to the first study about the Google effect, people are less likely to remember certain details they believe will be accessible online. This externalization of memory parallels what appears to occur with GenAI. When AI such as ChatGPT or Grok, or Gemini creates content, users mentally become disengaged from the ideas which makes the user less able to comprehend and encode information. He et al., 2023 supported this theory that the excessive reliance on artificial intelligence or other such excessive tools creates illusionary competence. During this process, the individuals feel knowledgeable despite limited internal understanding.

Table 1 - Demographic distribution of participants.

Parameter	Category	Count (n=90)	Percentage
Age	20-24 years	48	53.3%
	25-29 years	25	27.8%
	30-34 years	17	18.9%
Gender	Male	52	57.8%
	Female	38	42.2%
Profession	University Students	48	53.3%
	Early Career Professionals	42	46.7%

2.3. Organizational Behaviour and Decision-Making

Automation bias refers to the systematic tendency of operations to over-trust or defers to automated outputs even when they are incomplete or incorrect (Goddard et al., 2012). It was examined that relying on AI for organizational decision-making carries significant implications. It decreases critical thinking and analytical skills, thereby creating ethical and organizational challenges. Similarly, information system management in the workplace, manifests as managers or analysts accepting AI recommendations without sufficient interrogation, thereby transferring responsibility for sense-making to algorithmic processes (Cummings, 2015; Goddard et al., 2012). Similar automation biases have been documented in multiple high-stakes environments like aviation, military, healthcare and finance departments. It is in such domains where decision support tools have been shown to reduce human vigilance and increase chances of error (Abdelwanis et al., 2024; Cummings, 2015). When business managers use GenAI for scenario generation, business forecasting and hiring or shortlisting candidates, it can lead to entrenched errors, reduced accountability and erosion of managerial accountability. Therefore, to avoid such errors due to over-reliance of AI, companies may need to add mandatory human-in-the-loop checks, transparency checks and decision justification from decision makers (Enqvist, 2023).

2.4. Talent Development and Cognitive Deskilling

Cognitive deskilling is defined as the regression or decline of a person's skillset due to over-reliance on technology like AI, GPS or even calculators to perform tasks previously requiring mental effort (Ardichvili, 2022; Sutton et al., 2018). It occurs when automation substitutes for cognitive tasks that previously served as practice for skill development. Over time, reduced practice in the same discipline can lead to atrophy in operators' expertise. Natali et al., 2025, noted similar signs of skills erosion as they observed that when clinicians rely on AI diagnostic tools. It decreased the performance of the operator's ability to perform physical exams on patients and their diagnostic reasoning. Similar deskilling occurs in the workplace when workers in a company rely on AI. It is observed that AI can accelerate task completion, but they also reduce opportunities for employees to learn from their assigned task i.e., problem solving and feedback (Ferdman, 2025; Tang, 2025).

HR pipelines depend on progressive practice and its mastery. If routine analytics or writing tasks are constantly delegated to AI especially junior staff, the organization may risk producing staff that lacks fluency in respective domain and then struggle when AI assistance is not available. HR and Continuous learning department should hence design rotational learning

Table 2 - Post task assessment criteria

Assessment Variable	Measurement Tool	Scale / Metric	Evaluation Method
Cognitive Engagement	Keystroke-based index	0.0 to 1.0 index	Automated logging of pauses, revisions, and planning time
Memory Retention	Delayed recall test	0% to 100%	10-item instrument administered 45 minutes post-task
Originality	Structured expert rubric	1 to 5 (1=Generic, 5=Novel)	Blinded manual grading by two domain experts
Critical Reasoning	Structured expert rubric	1 to 5 (1=Shallow, 5=Robust)	Blinded manual grading by two domain experts

Table 3 - Criteria for semi-structured interviews

Interview Parameter	Implementation Criteria
Selection Criteria	Random subset selection: 7 from AI-Assisted, 7 from Search-Engine, 6 from Independent
Format & Duration	One-on-one semi-structured interviews; 10–15 minutes per participant
Core Thematic Questions	1. How would you describe your sense of authorship over the final text? 2. Can you describe your mental effort during the drafting phase? 3. Did you feel like an author or an editor?
Data Saturation	Interviews concluded when no new themes regarding cognitive load and authorship emerged

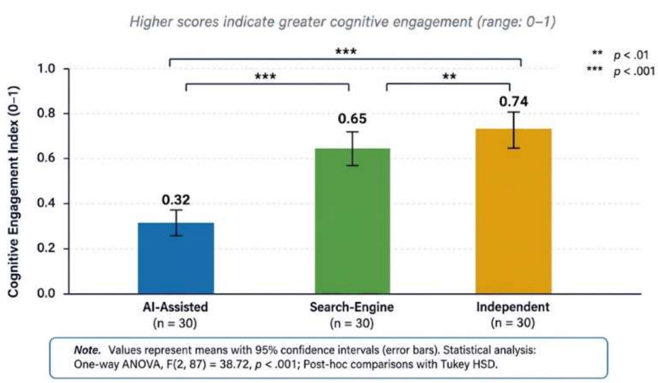

Figure 2 - Cognitive Engagement Index by Writing Condition

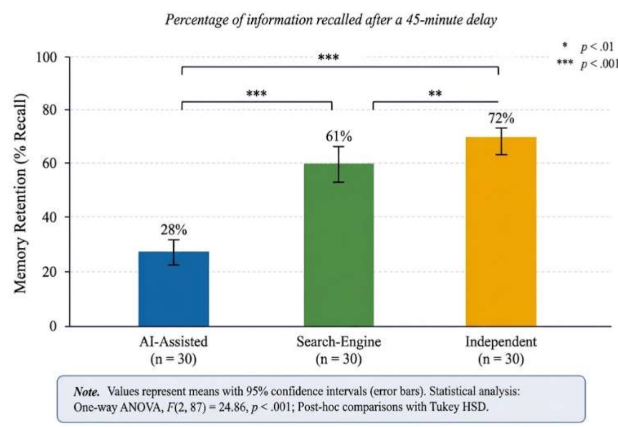

Figure 3 - Memory Retention Among the Groups

experiences, apprenticeship frameworks and their assessment mechanisms that require no AI intervention. It was observed that AI yields strong, localized productivity improvements but these gains do not automatically scale to sustained organization-level performance unless complementary investments (training, process redesign, governance) are made.

2.5. The Productivity Illusion

Sometimes AI increases measurable outputs such as faster reports, automated codes, and more drafts. This may seem productive at task level but it may obscure a productivity paradox. Because apparent task-level gains do not always translate into long term improvement in organization level values or worker capability. (Brynjolfsson et al., 2018) observed heterogeneous results during a systematic literature review of AI's economic effect. Hence, it was suggested that managers should deal short-term gain in task results sceptically and evaluate output quality, decision outcomes and learning metrics. Productivity dashboards that count documents or tickets closed in a workplace must be complemented with measures of decision quality, employee capability growth, and incidence of AI-induced errors.

3. Methodology

The primary objective of this study was to examine how different modes of digital assistance specifically generative AI, search engine-based inquiry and independent writing shape an individual's cognitive understanding in knowledge intensive tasks experimentally. The study was designed to evaluate varying forms of assistance influencing cognitive engagement, memory retention, originality and critical reasoning. By comparing the performance of participants with varying forms of assistance, this study was used to identify whether GenAI supports or undermines deeper forms of thinking and internalization.

The study utilized a sample of 90 participants recruited via convenience sampling from university mailing list and local professional network. To be selected for this study, participants had to meet specific inclusion criteria i.e., enrolled in higher education program or employed in a knowledge intensive profession. The participants also had to have baseline proficiency in English and word processing. Individuals with formal professional backgrounds in artificial intelligence (AI) or in computer science were excluded to minimize baseline bias as can be seen in Table 1. Several demographic and professional parameters were recorded to ensure a balanced sample Participants were randomly assigned to one of the 3 groups i.e., AI-assisted writing groups (n = 30), search engine groups (n = 30) and an independent writing group (n = 30). Members of the AI-assisted group was permitted unrestricted use of ChatGPT, DeepSeek, Gemini and other available GenAI software for the writing task, while participants in search engine group used Google or Bing to gather data for original writing. Participant's ages were restricted from 20-34 years (M = 24.8, SD = 3.2). The sample consisted of 52 males and 38 females. Professionally, the group was divided into 48 university students (specifically from engineering and business discipline) and 42 early career professionals. Prior to the study, participants self-reported baseline familiarity with GenAI tools and a pre-test was administered confirming the equivalence of baseline writing ability across all 3 randomly assigned groups.

All group participants were given the same writing topic i.e., "Explain organizational risks of relying on an automated decision system". The expected word count was approximately 450-500 words. Upon completion of the writing task, participants were asked for a post-task assessment designed to measure their cognitive engagement, originality, memory retention and critical reasoning. Cognitive engagement was evaluated via a keystroke-based cognitive load index that captured the participants' pauses, revision complexity and re-reading behaviour.

Time spent in planning, drafting, and revising was also recorded for assessing cognitive effort. Originality was assessed using a creativity rubric that assesses novelty of ideas, complexity of arguments in writing and examples given in the written task. Memory retention was estimated using a

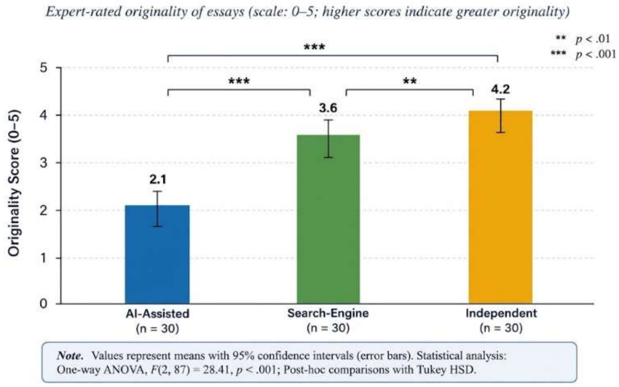

Figure 4 - Originality Score Among the Groups

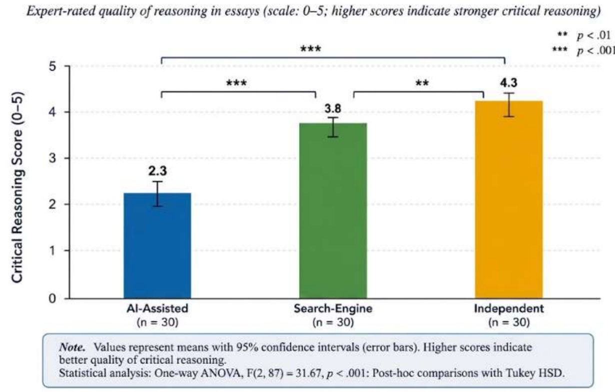

Figure 5 - Critical Reasoning Score Comparison Among Participant Groups

Table 4 - Summary of Quantitative findings by Group

Measure	AI-Assisted (n=30)	Search Engine (n=30)	Independent (n=30)	ANOVA Results
Cognitive Engagement Index	0.32 ($\pm$ 0.08)	0.65 ($\pm$ 0.07)	0.74 ($\pm$ 0.06)	$F(2,87) = 38.72, p < .001$
Memory Retention (%)	28% ($\pm$ 5.2)	61% ($\pm$ 6.1)	72% ($\pm$ 4.8)	$F(2,87) = 24.86, p < .001$
Originality Score (1-5)	2.1 ($\pm$ 0.4)	3.6 ($\pm$ 0.5)	4.2 ($\pm$ 0.3)	$F(2,87) = 28.41, p < .001$
Critical Reasoning (1-5)	2.3 ($\pm$ 0.3)	3.8 ($\pm$ 0.4)	4.3 ($\pm$ 0.3)	$F(2,87) = 31.67, p < .001$

Note: Values represent Means $\pm$ Standard Deviations. Significance evaluated at $p < 0.05$, η^2 represents overall effect size calculated via $\eta^2 = \frac{SS_{between}}{SS_{total}}$

delayed recall test administered 45 minutes after the task completion. A ten-item instrument that was designed to capture participants' internalization of concepts covered in their writing as can be observed in Table 2. The qualitative assessment of originality and critical reasoning were conducted by 2 independent experts. Expert 1 holds a Ph.D. in Cognitive psychology and other holds a Ph.D. in human factors. Both have a decade of experience in evaluating academic and professional writing. Both experts were trained on the specific 1-5 rubrics prior to grading. Inter rater reliability was calculated using Cohen's Kappa, yielding a strong agreement score of K = 0.82.

In the end, a subset of 20 participants engaged in short semi-structured interviews that captured the qualitative reflections on their sense of authorship, experience engaging and perceived cognitive effort with the task under different conditions. The interviews were conducted individually, targeting core themes of mental load and psychological distinction between authoring and editing. Data collection continued until thematic saturation was reached and no new insight regarding cognitive engagement emerged as can be observed in Table 3.

4. Results

Quantitative data were analysed using one-way ANOVA to compare differences in mean performance among the 3 groups as can be observed in Table 4. Tukey's HSD post hoc tests were employed to evaluate specific group-level differences. Qualitative data from interviews and written reflections were analysed using thematic coding procedures. The coding procedure focused on perceived cognitive load, engagement, and the psychological distinction between "authoring" and "editing."

One way Analysis of variance (ANOVA) was conducted using F-statistics to determine significant difference between the 3 groups:

$$F = \frac{MST}{MSE}$$

In the above equation MST is the mean square treatment (between group variance) and MSE is the mean square error (within group variance). To isolate specific difference among group pairs, Tukey's Honestly Significant Difference (HSD) post doc test was applied.

$$HSD = q\sqrt{\frac{MSE}{n}}$$

Where q is the studentized range statistic and MSE is the mean square treatment and n is sample size per group. To analyse the qualitative interview data, a 5-stage thematic coding procedure was implemented. The coding process calculated the cognitive load, effort and sense of authorship. The thematic coding procedure followed a structured sequence from initial familiarization and code extraction to thematic clustering, cross-validation, and final theme definition. The complete sequence of the thematic coding procedure can be observed in Figure 1. The inductive process began with audio transcription and initial familiarization, followed by systematic code extraction, thematic clustering, cross validation against quantities findings and final theme definition.

4.1. Cognitive Engagement

A one-way ANOVA showed a statistically significant difference in cognitive engagement across 3 groups i.e., F (2, 87) = 38.72, p < 0.001, $\eta^2 = .47$. Post hoc comparison using Tukey's HSD test indicated that AI-assisted groups (M = 0.32, SD = 0.08) had significantly lower cognitive engagement than both search engine and independent group. The search engine group also scored significantly lower than independent group (p =0.002). The participants relying on AI exhibited substantially fewer revision behaviours, reduced planning time, and shorter cognitive pauses. It was the signal of diminished effortful processing. However, independent writing groups showed enhanced engagement, recursive drafting and active problem solving as can be observed in Figure 2. Search engine-assisted participants demonstrated moderate engagement, consistent cognitive demands of evaluating external information and used it in the synthesis of writing task.

4.2. Memory Retention

The delayed recall test showed severe disparities in memory retention as can be observed in Figure 3. Participants that used AI demonstrated 28% of core idea recall, while search engine users showed 61% and independent writers could recall 72% of the core ideas in the writing task. During the interview phase, participants could not recall most of what they had written. The users often felt that they had "outsourced the thinking" to the AI model. These findings support cognitive offloading research, which predicts reduced memory encoding when GenAI does all the conceptual thinking for the user.

4.3. Originality

Originality scores differed significantly in participant groups as AI-assisted groups were evaluated to be generic, having low novelty and highly similar in structure. The essays lacked distinctive examples and showed minimal argument differentiation as can be observed in Figure 4. Search engine users scored moderately often integrating diverse information discovered via searching online. Independent participants produced more diverse, original and varied analyses in their essays. It demonstrated richer conceptual development among the participants and a more distinctive argumentative structure.

4.4. Critical Reasoning

Expert rating of critical reasoning showed the same pattern seen in other measures as can be observed in Figure 5. Essays from AI-assisted participants were observed to be shallow, lacking analytical depth, and relying on broad unsubstantiated assertions. Search engine participants showed better logical coherence and evidence-based research using manual evaluation from sources. Independent participants produced more robust reasoning with well-developed arguments, better interpretation from their research and clear engagement with counterarguments.

5. Discussion

Qualitative data presented an explanation of the above quantitative results. Participants of AI-assisted groups frequently described the experience with a diminished sense of ownership. They characterized themselves as "editors, not authors". Hence, they had limited mental involvement in constructing arguments in the given essay task. The findings demonstrated a consistent pattern as GenAI significantly reduced cognitive engagement, memory retention, originality and overall critical reasoning of AI-assisted participants. These results support theoretical expectations derived from cognitive offloading literature (Juaban et al., 2025; Risko and Gilbert, 2016). Furthermore, the experiment demonstrated that reliance on large language models (LLM) creates cognitive shortcut that weakens attention, internalization and reflection (Gerlich, 2025; Kosmyna et al., 2025).

5.1. Reduced Cognitive Engagement and Cognitive Debt Mechanism

The low cognitive engagement index observed in AI-assisted participants suggests a shift in the cognitive architecture of task completion. Instead of conceptualizing, drafting and iteratively refining arguments, participants overwhelmingly described themselves as editors who simply guided or adjusted LLM-generated text. This result aligns with (Kosmyna et al., 2025)'s findings that frequent AI use accumulated cognitive debt i.e., a reduction in effort from processing information, impairing the development of deeper reasoning. Search engine users and independent participants demonstrated better engagement comparatively showing deeper comprehension of the given task.

5.2. Impact on Memory Internalization

Memory retention results were particularly stark as AI-assisted participants were able to recall less than half as much information when compared with other groups. The result corresponds with (Sparrow et al., 2011)'s "Google Effect" theory where easy access to information reduces the likelihood of memory encoding. The AI-assisted participants barely remembered what they had written, suggesting that GenAI produces a severe case of digital amnesia. This supports that fact that AI only retrieves information and does not engage in core cognitive acts like idea formation and argument sequencing.

5.3. Homogenization of Ideas and Decline in Originality

Originality scores demonstrated an additional structural concern as AI-generated essays were generic, average, lacked depth and stylistically similar. Participants who relied on AI have a similar structure, theme, and rhetorical patterns. This support concerns that LLMs, trained to optimize for probable next words, inherently suppress outlier or creative perspectives (Tang, 2025). Independent and search-engine participants, by contrast, produced diverse argumentative approaches and richer contextual examples.

5.4. Weakening of Critical Reasoning

A reduction in critical reasoning was observed among AI-assisted participants. It raises important consequences for organizational decision making. As absence of counterargument engagement, thin integration and overreliance based on generic assertion supports findings on automation bias literature (Goddard et al., 2012; Romeo and Conti, 2026). It is emphasized that AI systems are convincing because of their output structure but it lacks depth and may hallucinate and generate content that lack sound reasoning. When users accept these outputs as facts without any background knowledge or base, quality of findings becomes eroded. This pattern was amplified by participants' reduced engagement and sense of authorship in the writing task. According to psychology, it diminishes the user's accountability and analytical persistence.

5.5. Implications for Organizations

The above findings support the idea that Generative AI (GenAI) can enhance surface level productivity but reduces human cognition skills. These skills are important for higher quality work, decision making, leadership development and expertise acquisition in any field. Furthermore, the results indicate the broader context of organizational development warning users of cognitive deskilling in AI assisted environment (Ferdman, 2025; Natali et al., 2025). It supports the productivity paradox emphasizing that rapid output does not necessarily lead to increase in performance (Brynjolfsson et al., 2018). These overall dynamics suggest organizational risk in creating a workforce overly reliant on AI. They will deliver tasks fast but will lack depth, engagement and will recall less. The workforce will produce more results quantitatively but will lack quality. To counter this paradigm, organization must adopt healthy AI framework and integrate them in their workflow.

6. Framework for healthy AI integration

Building on the results of this study, it is recommended that we establish a robust framework for AI governance and integration. The following framework is based on policies made via OECD (Organization for Economic Co-Operation and Development) and the European Union provide practical intervention on the usage of AI ("Artificial Intelligence Act: Article 14—Human oversight,"", 2024).

6.1. AI Literacy Training

A foundational component to use AI in a responsible manner is the creation of a comprehensive AI literacy program. AI literacy program refers not merely the operational ability to use GenAI models but to provide an informed understanding of how these models' function. It includes the probabilistic nature of LLMs, susceptibility to hallucination and inherent biases (OECD, 2025). (Kathala and Palakurthi, 2025) observed that when users understand the limitations and epistemic uncertainty of AI models, they are less likely to trust them blindly and become more critical in evaluating AI content. Hence, programs based on an AI literacy framework show that structured literacy interventions improve employee's ability to assess AI outputs, trace data sources and identify GenAI biases (OECD, 2025). Consequently, establishment of mandatory GenAI literacy curriculum is essential to ensure that employees approach AI not as an authority but as a tool that requires oversight.

6.2. Cognitive Engagement Protocols

The second component of healthy AI policy is to preserve cognitive engagement during knowledge-intensive tasks. One effective method is the implementation of protocols requiring employees to draft an initial outline or plan before consulting AI systems. It was observed in previous experimentations that individuals who engage in independent idea conception before consulting AI exhibit better memory retention, improved analytical reasoning and better internalization of core concepts (Ansari and Qamari, 2025; Kosmyna et al., 2025). Furthermore, it purposes risks of cognitive deskilling is minimal when AI augments routine tasks but overreliance on GenAI will replace core learning experience of user which is important for professional development (Natali et al., 2025).

6.3. Human Oversight Standards

Effective governance requires a clear human oversight mechanism especially for high-risk or critical tasks. The recent EU AI Act mandates human-in-the-loop oversight during a wide range of highly critical applications ("Artificial Intelligence Act: Article 14—Human oversight,"", 2024). The act emphasizes that humans must retain ultimate responsibility for decisions but can be assisted using AI outputs. Operationalizing this principle in a workplace should require documentation of the AI model to be used, prompts that will be used, human edits that will be used and rationale for the final decision. Experimental studies on AI and automations have suggested that transparency and auditability significantly reduce the likelihood of uncritical acceptance of AI-generated recommendations (Enqvist, 2023; Vered et al., 2023).

6.4. Monitoring and Metrics

Finally, responsible AI integration requires continuous monitoring of cognitive and organizational outcomes. Companies should track the proportions of tasks completed with and without AI assistance, rates of human verification, AI-related errors in tasks, memory retention of employees with AI involved, and employee-reported confidence in performing tasks independently. Research on the productivity paradox clearly describes that an increase in output quantity does not mean improved performance, quality or understanding (Brynjolfsson et al., 2018). Therefore, ongoing measurement metrics enable organizations to evaluate whether AI is genuinely enhancing decision quality and performance or merely accelerating superficial productivity.

Conclusion

The study examines how generative artificial intelligence (GenAI) reshapes cognitive processing, decision making and core competencies within modern era. Despite GenAI increases efficiency towards a task such as draft acceleration, synthesizing information and producing coherent text but it comes at the cost of reduced cognitive engagement, weak memory retention, diminished originality and shallow critical reasoning. Across all measured dimensions, participants who relied on AI to complete the analytical writing task displayed significantly lower performance than both search-engine users and independent writers. The written essays by AI-assisted group were less original, analytically shallow as most participants could not recall what they had written. Their cognitive engagement indicators suggested minimal efforts were made by the participants. Qualitative interviews reinforce the findings as AI-assisted participants described the experience like “editors rather than authors”. They lacked a sense of cognitive ownership over the text they had produced. The findings reinforce similar findings in cognitive, behavioural science and organizational studies indicating that AI-driven automation can distort human cognition towards passive thinking habits (He et al., 2023; Juaban et al., 2025; Kosmyna et al., 2025).

Similarly, implications towards high reliance of GenAI are also substantial. As firms increasingly integrate AI in their workflow risk weakening basic cognitive skills which leads to overdependence, fragile leadership expertise particularly in early career staff. To prevent such outcome structures AI governance system is required. System such as literacy training, oversight mechanism, monitoring metrics and cognitive engagement protocol are required that augment rather than substitute human capabilities. Without such measures, short-term gains may obscure a long-term erosion of critical thinking and decision quality.

Limitations

The current study provides strong experimental evidence on cognitive engagement and erosion of human competence on overreliance of GenAI workflow. However, several limitations must be acknowledged for this study. Firstly, the sample size (n = 90) is moderate and was composed of early career professionals and university students. While, the sample size seems appropriate for experimental study it may not generalize older or more experienced workers whose habits and AI usage behaviour is different. Secondly, the experiment focused on short term memory retention i.e., 45 minutes post task. Long term memory effects like retention over many weeks or even months were not measured. This restricts findings on sustained learning or the durability of cognitive offloading effects. Lastly, ecological validity remains constrained as real organizational environment includes time pressure, multi-step works, managerial oversight and collaboration. These factors may mitigate or enhance cognitive effects observed. The limitations do not undermine the validity of the core findings but suggest that cognitive impact of GenAI are complex, multifaceted and context dependent.

Future Research

Future research may expand and refine the above findings in multiple ways like investigating the long-term memory retention mechanism using GenAI. Researchers may examine how AI reshapes memory, reasoning and expertise in longitudinal study. Additionally, integrate neurophysiological measures like EEG or fMRI to examine the internal cognitive processing while using AI. Field experimentation to investigative AI literacy programs and mandatory oversight protocols to identify best policies for organizational settings. Understanding which policy is best for increasing employee capabilities and firm level outcome.

Author Contributions (Mandatory)

Zaid Sarfraz: Conceptualization, Methodology, Investigation, Formal Analysis, Writing - Original Draft, **Zain Sarfraz Dogar:** Investigation, Conceptualization, Visualization, **Attique ur Rehman:** Investigation, Review and Editing, **Asim Ali Rao:** Data curation, Formal Analysis, Supervision, Review and Editing, **Muhammad Yasir Bilal:** Data curation, Formal Analysis, Supervision, Review and Editing, **Abdul Basit:** Data curation, Formal Analysis, Supervision, Review and Editing

Acknowledgements

The authors have no specific acknowledgements to declare.

Conflicts of Interest Statement

The authors declare no competing financial or non-financial interests and no personal relationships that could have influenced the work reported in this manuscript.

Funding Information

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

AI Statement

During the preparation of this manuscript, the authors used ChatGPT and Claude AI solely for grammar and language editing, formatting, and/or reference checking. The tool was not used to generate scientific content, create or modify figures or images, perform data analysis, or modify research data. All scientific content, interpretation, analysis, and conclusions were developed and verified by the authors, who take full responsibility for the final manuscript.

References

Abdelwanis, M., Alarafati, H. K., Tammam, M. M. S. and Simsekler, M. C. E. (2024), "Exploring the risks of automation bias in healthcare artificial intelligence applications: A Bowtie analysis", Journal of Safety Science and Resilience, Vol. 5 No. 4, pp. 460-469.

Ansari, S. R. and Qamari, I. N. (2025), "Artificial intelligence and students' cognitive learning outcomes with bibliometric and content analysis for future research agenda", Discover Education, Vol. 4 No. 1, pp. 441.

Ardichvili, A. (2022), "The Impact of Artificial Intelligence on Expertise Development: Implications for HRD", Advances in Developing Human Resources, Vol. 24 No. 2, pp. 78-98.

"Artificial Intelligence Act: Article 14—Human oversight". (2024), (accessed 31.08.2026 2026).

Autio, C., Schwartz, R., Dunietz, J., Jain, S., Stanley, M., Tabassi, E., Hall, P. and Roberts, K. (2024), *Artificial intelligence risk management framework: Generative artificial intelligence profile (NIST AI 600-1)*, National Institute of Standards and Technology.

Brynjolfsson, E., Rock, D. and Syverson, C. (2018), "Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics", in *The economics of artificial intelligence: An agenda*. University of Chicago Press, pp. 23-57.

Cummings, M. (2015), "Automation Bias in Intelligent Time Critical Decision Support Systems", in D., H. and C., L. W. (Eds.), *Decision Making in Aviation*. Routledge.

Enqvist, L. (2023), "'Human oversight' in the EU artificial intelligence act: what, when and by whom?", Law, Innovation and Technology, Vol. 15 No. 2, pp. 508-535.

Ferdman, A. (2025), "Practices Worth Preserving: Knowing when to Offload to Technology", Philosophy & Technology, Vol. 38 No. 3, pp. 100.

Gerlich, M. (2025), "AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking", in *Societies*, pp. 6.

Goddard, K., Roudsari, A. and Wyatt, J. C. (2012), "Automation bias: a systematic review of frequency, effect mediators, and mitigators", Journal of the American Medical Informatics Association, Vol. 19 No. 1, pp. 121-127.

He, G., Kuiper, L. and Gadiraju, U. (2023), "Knowing About Knowing: An Illusion of Human Competence Can Hinder Appropriate Reliance on AI Systems", in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Hamburg, Germany, pp. Article 113.

Juaban, M., Dominguez, R. R., Dequina Jr, L. J. L. A. and Journal, S. S. I. (2025), "Thinking Less in the Age of Machines: AI Overuse and Cognitive Decline", Vol. 1 No. 3, pp. 34-45.

Kathala, K. C. R. and Palakurthi, S. (2025), "AI Literacy Framework and Strategies for Implementation in Developing Nations", in *Proceedings of the 2024 16th International Conference on Education Technology and Computers*. Association for Computing Machinery, pp. 418–422.

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I. and Maes, P. J. a. p. a. (2025), "Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task", Vol. 4, pp. 206.

McBain, S. (2025), "Are we living in a golden age of stupidity?", available at: https://www.theguardian.com/technology/2025/oct/18/are-we-living-in-a-golden-age-of-stupidity-technology (accessed 31.08.2026 2026).

Natali, C., Marconi, L., Dias Duran, L. D. and Cabitza, F. (2025), "AI-induced Deskilling in Medicine: A Mixed-Method Review and Research Agenda for Healthcare and Beyond", Artificial Intelligence Review, Vol. 58 No. 11, pp. 356.

OECD. (2025), "Empowering Learners for the Age of AI", in. European Union.

Risko, E. F. and Gilbert, S. J. (2016), "Cognitive Offloading", Trends in Cognitive Sciences, Vol. 20 No. 9, pp. 676-688.

Romeo, G. and Conti, D. (2026), "Exploring automation bias in human–AI collaboration: a review and implications for explainable AI", AI & SOCIETY, Vol. 41 No. 1, pp. 259-278.

Sparrow, B., Liu, J. and Wegner, D. M. (2011), "Google Effects on Memory: Cognitive Consequences of Having Information at Our Fingertips", Science, Vol. 333 No. 6043, pp. 776-778.

Sutton, S. G., Arnold, V. and Holt, M. (2018), "How Much Automation Is Too Much? Keeping the Human Relevant in Knowledge Work", Journal of Emerging Technologies in Accounting, Vol. 15 No. 2, pp. 15-25.

Tang, B. L. (2025), "The Epistemic Downside of Using LLM-Based Generative AI in Academic Writing", in *Publications*, pp. 63.

Tucker, R. B. (2025), "Is ChatGPT Making Us Stupid?", available at: https://www.forbes.com/sites/robertbtucker/2025/06/20/is-chatgpt-making-us-stupid/ (accessed 31.08.2026 2026).

Vered, M., Livni, T., Howe, P. D. L., Miller, T. and Sonenberg, L. (2023), "The effects of explanations on automation bias", Artificial Intelligence, Vol. 322, pp. 103952.