The Cognitive Shift: A Mixed-Methods Investigation of Artificial Intelligence's Effect on Human Analytical Thinking and Divergent Creativity

Author	Affiliation	Course	Submitted
Bandha Arafat	Faculty of Education and Humanities, Gulu University, Uganda	Research Paper (Undergraduate)	April 18, 2026

Abstract

The rapid integration of Generative Artificial Intelligence (GenAI) into daily and professional life has outpaced empirical research on its cognitive consequences. This paper investigates whether reliance on AI for problem-solving and content generation enhances or diminishes human critical thinking and creative output. Employing a longitudinal, mixed-methods design (N = 120), participants engaged in a 4-week period of AI-assisted tasks, followed by an unassisted post-test. Quantitative results indicate a significant reduction in cognitive effort—specifically analytical justification and error detection—when AI was available. However, qualitative findings revealed a nuanced effect on creativity: while AI accelerated idea generation (fluency), it reduced cognitive flexibility and original solution novelty when the AI was subsequently removed. We propose a theoretical framework of *cognitive offloading to automation*, suggesting that AI serves as a powerful but potentially atrophying prosthetic for human reasoning. The paper concludes with recommendations for AI interaction paradigms that preserve, rather than replace, human creative and critical faculties.

Keywords: *Artificial Intelligence, Human Cognition, Creativity, Cognitive Offloading, Critical Thinking, Human-Computer Interaction, Gulu University*

1. Introduction

The advent of Large Language Models (LLMs) such as GPT-4, Claude, and Gemini represents a paradigm shift in how humans access processed information and generate novel content. Unlike search engines that retrieve existing links, generative AI produces synthesized, probabilistic responses. This capability raises a fundamental question: is AI *augmenting* human thinking, or is it *replacing* core cognitive processes?

Early research on cognitive offloading (Storm & Stone, 2015) demonstrated that humans readily delegate memory to external devices (e.g., smartphones). AI, however, delegates *processing*—the

very act of linking concepts, making analogies, and evaluating logic. This paper addresses two primary research questions:

- **RQ1:** How does sustained AI assistance affect a user's subsequent ability to perform analytical reasoning and error detection *without* AI?
- **RQ2:** What is the effect of AI collaboration on divergent creative thinking, measured by fluency, flexibility, and originality?

We hypothesize a dual effect: AI will increase task efficiency and initial creative volume (fluency) but will decrease depth of analytical processing and the originality of independent creative output.

As students at Gulu University's Faculty of Education and Humanities, this research is particularly relevant to understanding how future teachers and humanists may be cognitively shaped by AI tools before they enter professional practice. The Ugandan educational context, where digital tools are rapidly expanding, demands evidence-based guidance on AI integration.

---

2. Literature Review

2.1 Cognitive Offloading and the Extended Mind

The extended mind thesis (Clark & Chalmers, 1998) argues that external tools can become constitutive parts of cognitive processes. A notebook extends memory; a calculator extends arithmetic. Generative AI, however, extends reasoning itself. This creates a unique risk: when reasoning is consistently outsourced, the biological system may cease to develop or maintain the necessary internal models.

Empirical work on "Google effects" (Sparrow et al., 2011) showed that people remember *where to find* information rather than the information itself. More recent studies on AI code generation (Vaithilingam et al., 2022) found that programmers using AI copilots produced more code but understood less of it. These findings suggest a cognitive efficiency–depth trade-off.

2.2 AI and Creativity: Divergent vs. Convergent Thinking

Creativity research distinguishes between *divergent thinking* (generating many novel ideas) and *convergent thinking* (selecting the best solution). Runco and Jaeger (2012) define creativity as requiring both novelty and effectiveness. Generative AI excels at producing high-fluency, moderately novel outputs—the statistical average of its training data. Boden (1998) warned that AI's "combinatorial novelty" (recombining existing concepts) differs fundamentally from "transformational novelty" (creating new conceptual spaces).

Recent experimental studies (Doshi & Hauser, 2023; Habib & Walsh, 2025) have shown that AI assistance increases idea quantity but reduces the variance of creative output across participants, as ideas converge on AI-suggested prototypes. Our study extends this work by measuring *unassisted* creative ability after a period of AI reliance, testing for lasting cognitive effects.

2.3 Automation Bias and Metacognitive Vigilance

Automation bias—the tendency to over-rely on automated aids and neglect contradictory information—is well documented in aviation and medicine (Parasuraman & Manzey, 2010). Skitka et al. (1999) found that even when automation was unreliable, participants failed to verify its outputs. Generative AI is particularly prone to this because its outputs are fluent and confident, even when incorrect. This may reduce metacognitive vigilance—the active monitoring of one's own understanding.

2.4 Relevance to Ugandan Higher Education

In Uganda, the Ministry of Education and Sports (2023) has called for integration of digital literacy into teacher training. However, no policy yet addresses generative AI specifically. Gulu University, located in Northern Uganda, serves a diverse student body with varying access to technology. Understanding AI's cognitive effects is essential before mandating its use in classrooms. This study provides foundational evidence for educators and policymakers.

---

3. Methodology

3.1 Research Design

A longitudinal, parallel-group, mixed-methods experimental design with repeated measures was employed. The study was approved by the Gulu University Research Ethics Committee (Protocol #GU-FEH-2025-089).

3.2 Participants

A power analysis (G*Power; $\alpha = .05$, $\beta = .80$, moderate effect size $f = 0.25$) indicated N = 102 required. We recruited 132 participants from Gulu University and surrounding community, with 120 completing all phases (90.9% retention).

Inclusion criteria: Age 21–45; daily computer use ≥ 4 hours; English fluency; no prior professional generative AI use beyond occasional testing.
Exclusion criteria: Formal training in AI prompt engineering; diagnosis of cognitive impairment.

Participants were stratified by age and education, then randomly assigned (1:1):

- Experimental group (n = 60): AI-assisted condition
- Control group (n = 60): Traditional tools condition

Demographics: 52% female, 45% male, 3% non-binary; mean age 31.4 (SD = 7.2); education: 68% bachelor's degree, 32% graduate degree. Groups did not differ significantly on any demographic or baseline measure ($p > .05$).

3.3 Procedure

The study lasted five weeks. The table below summarises each phase.

Phase	Week(s)	Description

Pre-test (unassisted)	0	90-min baseline: analytical reasoning + divergent creativity; no AI permitted.
Intervention	1–3	3 sessions/week × 45 min. Experimental: GPT-4 (temp 0.7). Control: web search + notes only.
Washout	4	No cognitive tasks — prevents practice effects.
Post-test (unassisted)	5	Identical to pre-test. No AI or search tools permitted. Qualitative interviews with n = 40 subset.

Table 1. Study procedure by phase and week.

3.4 Instruments and Metrics

Analytical thinking (RQ1):

- Cognitive Reflection Test (CRT-7; Frederick, 2005): seven items requiring override of intuitive but incorrect answers; scored 0–7.
- Error Detection Task: participants reviewed a 500-word fabricated business report containing 12 logical fallacies; outcome = errors correctly identified in 10 minutes.

Divergent creativity (RQ2):

- Alternate Uses Test (AUT; Guilford, 1967) for the object "brick." Inter-rater reliability ICC = .92.
  - **Fluency:** total number of uses generated.
  - **Flexibility:** number of distinct semantic categories.
  - **Originality:** each use rated 1 (common), 2 (uncommon), or 3 (rare).

Cognitive effort: NASA-TLX mental demand subscale (1–7) and response time.

Interaction logs (AI group only): prompts per task, prompt length, copy-paste ratio, revision frequency.

Qualitative interview protocol included questions on thinking processes, idea origins, disagreements with AI, and post-test experiences (n = 40; 30 min each).

3.5 Data Analysis

- **Quantitative:** Repeated-measures ANOVA (2 groups × 2 time points) with post-hoc paired *t*-tests (Bonferroni-corrected). Effect sizes: partial η^2.
- **Qualitative:** Thematic analysis (Braun & Clarke, 2006) using NVivo 14. Two coders achieved $\kappa = .86$.

---

4. Results

4.1 Quantitative Results

4.1.1 Analytical Thinking (RQ1)

Table 2 presents pre-test and post-test scores for both groups on analytical measures.

Measure	Group	Pre-test (M, SD)	Post-test (M, SD)	Change
CRT-7 (0–7)	Experimental	4.82 (1.33)	3.77 (1.52)	–1.05 ***
	Control	4.78 (1.41)	4.85 (1.39)	+0.07
	Experimental	8.93 (1.87)	6.15 (2.24)	–2.78 ***
	Control	9.02 (1.96)	8.88 (2.01)	–0.14

*Table 2. Analytical thinking outcomes. *** p < .001 (interaction effect, repeated-measures ANOVA).*

The experimental group's CRT-7 scores declined by **22%** (p < .001), while controls showed no meaningful change (+0.07). Error detection accuracy dropped from 74% to 51% in the experimental group (p < .001), compared to a stable 74% in controls.

Cognitive effort: Experimental participants reported lower effort *during* intervention (M = 2.8 vs. 4.6, p < .001) but markedly higher effort during the unassisted post-test (M = 5.2 vs. 4.5, p = .004), indicating a rebound effect consistent with cognitive dependency.

4.1.2 Divergent Creativity (RQ2)

Table 3 presents AUT scores. During the intervention, AI users generated substantially more ideas (fluency 22.4 vs. 14.7, p < .001), but this advantage did not transfer to unassisted performance.

Metric	Group	Pre-test	Post-test	Change
Fluency	Experimental	14.2	11.5	–2.7 *
	Control	13.9	14.1	+0.2
	Experimental	1.98	1.52	–0.46 ***
	Control	2.01	2.04	+0.03

*Table 3. Divergent creativity outcomes (AUT). * p < .05; *** p < .001.*

The experimental group's unassisted originality at post-test was **28% lower** than controls (p < .001), representing the most striking inter-group difference in the study.

4.2 Interaction Log Analysis (AI Group)

Copy-paste ratio increased from 58% (Week 1) to 79% (Week 3), while editing frequency dropped from 2.3 to 0.8 edits per task. Higher copy-paste ratio correlated negatively with CRT-7 post-test scores (r = –0.52, p < .001) and with originality (r = –0.47, p = .002), suggesting that passive consumption of AI output—rather than active engagement—drives cognitive decline.

4.3 Qualitative Findings

Three dominant themes emerged from thematic analysis of interviews (n = 40).

Theme 1: Automation Bias

Participants described a progressive reduction in verification behaviour across the intervention weeks.

> *"The AI was usually right, so I stopped checking. By week three, I just copied the AI's answer. It felt like a waste of time to double-check. — P27"*

Theme 2: Empty Abduction — "Waiting for the AI"

Multiple participants described a felt experience of cognitive paralysis during the unassisted post-test.

> *"I stared at the brick question for two minutes. I kept waiting for the AI to suggest something. It was like a muscle that had atrophied. — P54"*

Theme 3: Dialectical Scaffolding — the Protective Minority

A subset of seven of the twenty interviewed AI-group participants used AI as a *debate partner* rather than an answer engine. These individuals showed smaller cognitive declines on both quantitative measures.

> *"I wrote my own answer first, then asked the AI, then argued with it. — P15"*

This sub-group's strategy—pre-thinking before querying, then critically comparing outputs—aligns with what we term *dialectical AI use* and informs our practical recommendations.

---

5. Discussion

5.1 Interpretation of Findings

This study provides longitudinal evidence that sustained generative AI use—under passive consumption patterns—significantly reduces subsequent unassisted analytical reasoning and original creative output. The *atrophy hypothesis* is supported over the *augmentation hypothesis* for the interaction patterns observed.

For RQ1, the 22% decline in CRT scores indicates that participants outsourced not just information retrieval but cognitive reflection itself. AI's fluent answers provide an intuitive default, and users stopped practising the override mechanism that is central to System 2 thinking (Kahneman, 2011).

For RQ2, the fluency–originality trade-off shows that AI increased idea quantity during assisted use but decreased quality (originality) in subsequent unassisted performance. Originality requires practise in divergent search—a process AI short-circuits by serving plausible, statistically central answers.

5.2 Theoretical Contribution: Cognitive Offloading to Automation (COA)

We propose the **COA framework**: generative AI offloads generative and evaluative processes, creating a dependency loop that may be self-reinforcing without deliberate intervention.

Cognitive Offloading to Automation (COA) — Dependency Loop

Step 1: AI reduces cognitive effort → user offloads reasoning to AI

Step 2: Reduced practice weakens internal cognitive circuits → user becomes less capable without AI

Step 3: Lower unassisted capability increases AI reliance → loop intensifies

Protective factor: Dialectical AI use (pre-think → query AI → critique → compare) was associated with smaller declines in both analytical and creative measures.

Figure 1. The Cognitive Offloading to Automation (COA) dependency loop and its protective countermeasure.

5.3 Practical Implications for Ugandan Education

For Gulu University and comparable institutions in sub-Saharan Africa, the following evidence-based recommendations are advanced:

- **Curriculum design:** Embed structured "AI-free deep work" sessions to preserve and exercise analytical skills alongside AI-assisted tasks.
- **Teacher training:** Train educators in dialectical AI use—pre-thinking, querying, critique, and comparison—before embedding AI in pedagogy.
- **Policy:** The Ugandan Ministry of Education and Sports should develop guidelines for AI literacy that emphasise cognitive preservation, not merely efficiency gains.

5.4 Limitations

Several limitations should be noted. The four-week intervention is relatively short; longer-term effects and potential recovery remain unknown. The sample is drawn from one university, precluding generalisation across Uganda's rural–urban digital divide. Only one AI model (GPT-4) was tested; results may differ with other systems. Finally, the AUT captures only one facet of creativity; other dimensions (e.g., elaboration, resistance to premature closure) were not assessed.

5.5 Future Research

- Longitudinal studies (12+ months) tracking cognitive trajectories in Ugandan teacher trainees across the full degree programme.
- Comparative trials of different AI interaction pedagogies (passive vs. dialectical).
- Neuroimaging (fMRI) studies to examine neural plasticity changes during sustained AI use.

---

6. Conclusion

This paper provides robust empirical evidence that sustained generative AI use, as currently practised under passive consumption patterns, reduces human analytical reasoning and original creative ability when the AI is removed. The COA framework offers a theoretically grounded explanation for this atrophy dynamic.

Critically, however, **dialectical use**—treating AI as a debating partner rather than an answer engine—may preserve and even strengthen cognition. For students and educators at Gulu

University and beyond, the challenge is not to reject AI but to design interactions that augment without atrophying the very human faculties that make education meaningful.

As generative AI becomes embedded in Ugandan classrooms, evidence such as that presented here must inform policy before cognitive habits are cemented in an entire generation of teachers and learners.

References

Boden, M. A. (1998). Creativity and artificial intelligence. *Artificial Intelligence, 103*(1–2), 347–356.

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

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

Davenport, T. H., & Kirby, J. (2016). *Only humans need apply.* Harvard Business Review Press.

Doshi, A. R., & Hauser, O. P. (2023). Generative AI enhances individual creativity but reduces collective diversity. *arXiv preprint*, arXiv:2312.00205.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25–42.

Guilford, J. P. (1967). *The nature of human intelligence.* McGraw-Hill.

Habib, S., & Walsh, T. (2025). The creativity paradox. *Creativity Research Journal, 37*(1), 45–62.

Ministry of Education and Sports, Uganda. (2023). *National digital literacy framework.* Kampala: Government of Uganda.

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation. *Human Factors, 52*(3), 381–410.

Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal, 24*(1), 92–96.

Skitka, L. J., Mosier, K. L., & Burdick, M. (1999). Does automation bias decision-making? *International Journal of Human-Computer Studies, 51*(5), 991–1006.

Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory. *Science, 333*(6043), 776–778.

Storm, B. C., & Stone, S. M. (2015). Saving-enhanced memory. *Psychological Science, 26*(2), 182–188.

Vaithilingam, P., Zhang, T., & Glassman, E. L. (2022). Expectation vs. experience: Usability of code generation tools. *CHI Extended Abstracts*, Article 332.