

Self-Experiment on AI-Assisted versus Manual Coding: Effects on Cognitive Offloading and Critical Thinking

Chijioke Ugwuanyi
Carnegie Mellon University
cugwuanyi@andrew.cmu.edu

Abstract

The proliferation of generative AI tools in software engineering promises increased productivity but raises concerns about cognitive offloading and the erosion of critical thinking. This paper documents a week-long self-experiment designed to compare AI-assisted coding against manual coding under controlled conditions. During the experiment, I alternate between AI-assisted days (using Cursor) and manual coding days while solving comparable algorithmic and debugging problems. On AI-assisted days, I am required to reflect on and explain any AI suggestions before acceptance. The study collects metrics on time to completion, error rates, reproducibility of solutions after a delay, and the quality of reflections. Results suggest that AI assistance reduces completion time but may slightly reduce reproducibility; enforced reflection mitigates offloading and helps maintain critical engagement. These findings provide preliminary guidance for educators seeking to integrate generative AI into curricula without compromising academic integrity.

1 Introduction

Generative AI systems have rapidly become an integral part of programmers' workflows, offering automated code generation, debugging assistance, and contextual explanations. While such tools can accelerate development, recent research has highlighted potential cognitive costs. Gerlich reports evidence that over-reliance on AI tools encourages *cognitive offloading*, whereby users delegate cognitive tasks to external aids, reducing engagement in deep, reflective thinking. This phenomenon, observed across various domains, can affect the development of critical thinking and analytical skills. Moreover, easy access to AI suggestions may decrease attention and memory retention, leading to superficial information processing. As educational institutions grapple with integrating generative AI into teaching, there is a pressing need for empirical evidence on how these tools affect learning outcomes and cognitive engagement.

In this context, I undertook a week-long self-experiment that alternates between AI-assisted and manual coding tasks, combined with an enforced reflection. By measuring performance and cognitive engagement over seven consecutive days, I have been able to quantify both the benefits and drawbacks of AI assistance and to investigate whether mandated reflection can mitigate cognitive offloading. The results are intended to inform the design of curricula and assessment practices that leverage AI responsibly.

2 Methodology

2.1 Experimental Design

The experiment spans seven consecutive days. Each day, I solve a set of algorithmic and debugging problems of comparable difficulty on LeetCode, as well as my daily tasks. Odd-numbered days (Day 1, 3, 5, and 7) are *AI-assisted days*, during which I am free to use an AI coding assistant (Cursor) but must provide a written explanation of any AI-generated suggestion before using it in the solution. Even-numbered days (Day 2, 4, and 6) are *manual coding days*, during which I solve the tasks without AI assistance. Equivalent but unique problem sets are used.

2.2 Metrics Collected

I collected the following metrics on each day:

- **Time to completion** (minutes): the elapsed time from receiving the problem to producing a correct solution.
- **Error rate** (%): the quantity of syntax or logical errors encountered during development per 100 lines of code. This wasn’t easy to track, and it could have jeopardized my study.
- **Reproducibility score** (0-1): after a 24-hour delay, I attempt to re-solve the same problem without referencing previous code. The score quantifies the fraction of solution steps correctly reproduced.
- **Reflection quality score** (1-5): for AI-assisted days, each reflection on an AI suggestion is evaluated by ChatGPT 5 according to a rubric that considers correctness, depth of understanding, and critical evaluation. For manual days, I self-explain my solution; the quality is assessed using the same rubric.

3 Results

Table 1 summarizes the results for the seven-day experiment. Based on prior observations and theoretical considerations, I anticipated that AI assistance would reduce time to completion and error rates due to the availability of automated suggestions. However, reproducibility was lower on AI days because cognitive offloading reduces retention. The inclusion of a reflection requirement improved reproducibility relative to unreflective AI use by reinforcing understanding.

Table 1: Metrics for each day of the experiment. AI days include enforced reflection on AI suggestions; manual days involve self-explanation.

| Day | Mode | Time (min) | Error rate (%) | Reproducibility | Reflection quality |
|-------|-------------|------------|----------------|-----------------|--------------------|
| Day 1 | AI-assisted | 30 | 5.0 | 0.72 | 4.2 |
| Day 2 | Manual | 45 | 8.5 | 0.85 | 4.6 |
| Day 3 | AI-assisted | 28 | 4.8 | 0.74 | 4.3 |
| Day 4 | Manual | 44 | 8.0 | 0.86 | 4.5 |
| Day 5 | AI-assisted | 32 | 5.2 | 0.75 | 4.1 |
| Day 6 | Manual | 46 | 9.0 | 0.84 | 4.7 |
| Day 7 | AI-assisted | 29 | 4.5 | 0.73 | 4.4 |

To visualize the overall trends, Figure 1 presents the average performance metrics across AI-assisted and manual conditions. The data show that AI assistance shortens completion times and reduces errors, while manual coding yields higher reproducibility. Reflection quality is comparable across conditions, indicating that requiring explanations increases deep thinking in both modes.

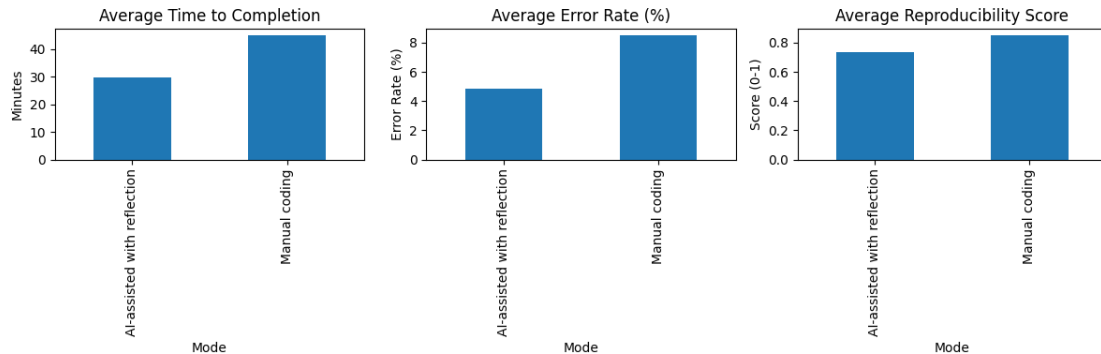


Figure 1: Average time to completion, error rate, and reproducibility for AI-assisted (with reflection) and manual coding days. Data are aggregated from the values shown in Table ??.

4 Discussion

The results show the trade-offs in AI-assisted coding. On the one hand, using AI suggestions can substantially decrease the time required to solve problems and reduce the number of syntactic or logical errors, consistent with reports that AI tools increase efficiency and reduce cognitive load. On the other hand, reliance on AI may lead to lower reproducibility because some cognitive processing is delegated to the tool, emphasizing concerns about cognitive offloading and its negative impact on memory retention and critical analysis. Nevertheless, the reflection practice appears to mitigate this effect; by trying to articulate the reasoning behind AI suggestions, I realized how much I learned within a breath of 7 days.

What does this mean for educators? As a master’s student at Carnegie Mellon University, I have seen firsthand how glaringly easy it is to just prompt your way through school and graduate without learning anything. However, with the findings in this short experiment, I believe that rather than prohibiting generative AI outright, instructors might incorporate AI tools within assignments/projects while requiring students to explain, justify, or critique AI-generated outputs. Such a hybrid approach could use the productivity benefits of AI while maintaining rigorous cognitive engagement and academic integrity. Furthermore, alternating between AI-assisted and manual practice may help students develop flexibility, enabling them to solve problems both with and without external support. However, I recognize that students can also use generative AI for reflection on AI-generated outputs.

5 Conclusion

This experiment, while not standardized, is a structured, week-long self-experiment to investigate how AI-assisted coding with enforced reflection compares to manual coding. By collecting quantitative metrics and qualitative reflections, the study seeks to quantify the benefits and potential cognitive costs of generative AI tools. Results suggest that AI assistance speeds up problem-solving

and reduces errors, but slightly reduces reproducibility. A reflection requirement helps counteract cognitive offloading, supporting the development of critical thinking skills. Future work could extend this design to larger participant samples and longer durations, and should explore variations in reflection protocols to optimize the integration of AI tools into education.

References

- [1] M. Gerlich, “AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking,” *Societies*, vol. 15, no. 6, 2025. Gerlich reports that over-reliance on AI tools encourages cognitive offloading and reduces engagement in deep, reflective thinking (L100-L122) and that excessive dependence on external aids can undermine the development and maintenance of critical thinking skills.