HHAI 2026
M. Alimardani et al. (Eds.)

© 2026 The Authors.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/FAIA260500

The Impact of AI Usage and Informativeness on Skill Development in Logical Reasoning

Shang WU [a,1], Hongyu YAO [b], Catarina BELÉM [a], Shuyuan FU [a], Mark STEYVERS [a], Padhraic SMYTH [a]

[a] *University of California, Irvine*
[b] *Massachusetts Institute of Technology*

Abstract. Artificial intelligence (AI) is being increasingly integrated into human problem-solving, yet its effects on individual skill development remain unclear. We examine how both AI usage and informativeness can shape learning in the context of a controlled logical reasoning task with on-demand access to AI assistance. We find that greater AI usage is associated with weaker skill development: heavy AI users underperform relative to comparable peers, whereas light AI users perform similarly to matched users who do not use AI. We also find in our study that these patterns are mediated by AI informativeness. Low-information AI neither improves immediate performance nor preserves performance after AI assistance is removed, and is linked to weaker learning overall. On the other hand, high-information AI was found to improve short-run performance without reducing post-AI outcomes on average in our experiments, but with heterogeneous effects. Our findings in general suggest that AI can, depending on context, either complement human skill development by amplifying independent reasoning or can act as a substitute that undermines such reasoning, with the implication that regulating AI access and usage will be important for promoting skill development in the presence of AI assistance.

Keywords. human-AI interaction, learning, individual heterogeneity, reliance

1. Introduction

Artificial intelligence (AI) assistance is being increasingly adopted in our everyday lives across a variety of contexts such as education [1,2], software development [3,4], autonomous driving [5], and medical diagnosis [6]. This has led to increased recent research attention on topics such as human-AI complementarity [7,8,9]. In parallel, there are concerns that easy access to AI may foster overreliance, reduce independent effort, and potentially hinder learning from the perspective of human users [10,11]. However, there have been relatively few studies to date that directly examine how AI assistance

[1]Corresponding Author: Shang Wu, University of California, Irvine, CA, 92697; E-mail:shangw13@uci.edu

affects individual skill development over time. In this paper, we study this issue with a particular focus on assessing skill development after AI assistance has been removed.

The literature to date on the impact of AI assistance has provided a number of alternative perspectives on the role of AI assistance in the context of human skill learning. Some studies argue that richer AI assistance, such as detailed guidance or uncertainty estimates, can improve human performance [12,13]. Other studies caution that limiting AI assistance may help reduce human overreliance and promote independent thinking [14,15]. Moreover, most prior work focuses on immediate performance gain rather than comparing pre-AI and post-AI learning outcomes [14,16,17]. To provide a more nuanced understanding, we investigate how both (i) the amount of AI usage, and (ii) the amount of information provided by AI, shape the trajectory of individual skill development.

We study how access to AI assistance of varying informativeness shapes individual skill development over time. Specifically, we address several related questions. First, we ask whether greater AI usage, as a behavioral pattern of reliance, is associated with weaker subsequent latent skill development. We then investigate how the informativeness of AI assistance moderates this relationship. Specifically, we examine whether the amount of information provided by the AI differentially alters immediate performance and subsequent latent skill development. Finally, we examine how individuals' AI usage strategy drives the differences in observed learning outcomes.

We investigate these questions through a controlled user study with time-constrained logical reasoning tasks. Participants solve puzzles under conditions that allow on-demand interaction with a simulated AI system. Each participant completes assessments both before and after AI exposure without any assistance, enabling us to isolate the effect (of interacting with AI) on skill development beyond the period of AI use (see Figure 1). The experimental design systematically manipulates the level of information provided by the AI, a central treatment dimension aligned with our theoretical focus on reliance and learning. Our goal is to understand how varying the informativeness of assistance changes AI usage behavior and, in turn, affects individual learning outcomes.

Our results show that AI assistance is associated with weaker skill development in logic-based problem-solving tasks. This effect is driven by usage behavior: earlier and more frequent reliance on AI leads to less independent problem-solving effort. As a result, heavy AI users underperform comparable peers in terms of skill development, whereas light AI users slightly outperform peers with similar initial ability who did not use AI. The magnitude and direction of these associations vary with AI informativeness. High-information AI substantially improves immediate performance and, on average, does not appear to impair skill growth. In contrast, low-information AI neither meaningfully improves immediate performance nor preserves subsequent performance. Moreover, high-information AI is linked to heterogeneous effects: higher-ability individuals tend to use it more selectively and later in the problem-solving process, and exhibit stronger skill growth. On the other hand, lower-ability individuals rely more heavily and earlier, and show comparatively weaker learning outcomes. Lower-ability participants report elevated perceived ability when AI is available, suggesting a possible role for miscalibrated self-assessment that may, in turn, be associated with weaker subsequent learning.

Collectively, these findings underscore that the educational and productivity consequences of AI depend not only on access but also on its informativeness and user het-

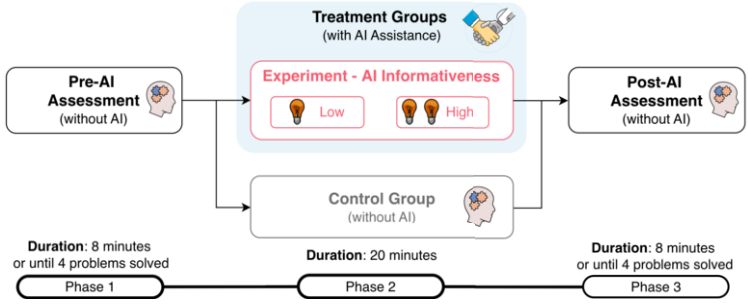


Figure 1. Study design with three phases: Phase 1 (pre-AI) and Phase 3 (post-AI) human performance was assessed without any assistance. In Phase 2, participants were assigned to either a control group (no AI) or a treatment group with optional AI assistance that varies in AI informativeness.

erogeneity. AI can augment learning when it complements sustained cognitive engagement, but may hinder skill development when it substitutes for it, particularly among individuals most inclined to rely on it.

2. Related Work

The concept of cognitive skill development refers to durable improvements in reasoning capacity, problem-solving ability, and learning transfer beyond immediate task performance [18]. In this context, a central question in recent human–AI interaction (HAI) research is whether AI assistance in general merely affects immediate performance or can also meaningfully affect longer-term underlying skill development (either positively or negatively). Research in HAI has largely emphasized human-AI complementarity and performance gains during AI use [7,8]. There has been less research on the evaluation of downstream learning outcomes, and the results of that work have been mixed in their conclusions. Some studies show that external assistance can scaffold reasoning and improve retention when appropriately designed [19,20]. Related work suggests that AI is beneficial when it complements rather than substitutes for human thinking [21]. In contrast, other work documents reduced cognitive effort, increased offloading, and weaker skill development [22,23]. Additional evidence suggests that AI assistance can improve immediate task performance without inducing skill degradation [24]. These studies reveal a consistent tension between immediate support and long-term skill development. In addition, much of the previous literature has had only limited emphasis on evaluating human performance after AI assistance is removed, making it difficult to distinguish genuine skill development from temporary reliance on external support. Pre- and post-assessment designs provide a clear approach by measuring baseline ability before AI exposure and reassessing performance after the AI is removed [25], making it possible to attribute changes to learning rather than to ongoing assistance.

Another underexplored dimension concerns individual heterogeneity. Prior work shows that reliance behaviors, such as help-seeking or AI adoption rate, vary substantially across individuals [26,27]. However, many studies either provide continuous AI access with limited user autonomy [28,29] or rely on self-reported measures of traits and trust [27,30], which are vulnerable to reporting biases [31]. Moreover, relatively little work

systematically connects objectively measured baseline ability to subsequent reliance patterns and post-assistance learning outcomes. As a result, it remains unclear whether AI engagement differences reflect stable individual skill differences or context-specific behavioral responses.

Additionally, prior research shows that design features such as timing and explanations displays influence reliance and performance [24,27,32], and some studies vary AI accuracy to study these effects [33]. However, manipulating accuracy introduces trust dynamics: early errors can shift beliefs and confound perceived reliability with the intrinsic AI assistance value [34]. Moreover, perspectives diverge on the role of informational depth. Some argue that richer guidance enhances understanding and performance [12,13], while others suggest that limited assistance encourages deeper engagement and mitigates overreliance [14,15]. Yet few studies cleanly isolate the effect of AI informativeness itself by assessing skill levels after AI assistance is removed.

We address these gaps by holding AI accuracy constant while experimentally varying informational depth. Combined with a pre- and post-assessment design and an objective measure of baseline ability, our framework allows us to examine how initial ability influences AI reliance and subsequent skill development, as well as how differences in AI informativeness affect learning once the AI is removed.

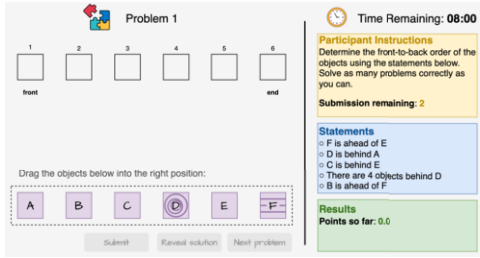
3. Experiment

We conducted a controlled user experiment to study how AI assistance affects individual skill development when help-seeking is self-directed, and tasks have time constraints. We manipulated the level of AI informativeness to examine how AI engagement influences learning and reliance.

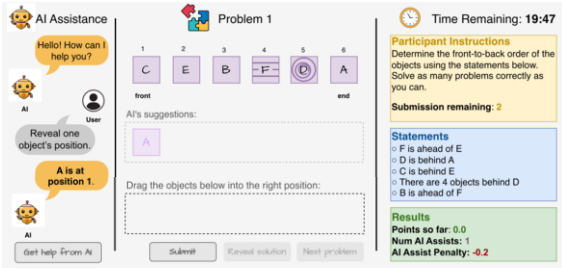
3.1. Task Description

We designed a time-constrained, logic-based puzzle in which participants determine the unique ordering of six objects subject to a set of constraints. Although this type of structured logic puzzle is an abstraction of real-world situations, it nonetheless offers a controlled environment that isolates the impact of AI assistance on problem-solving and learning.

For each problem, participants were presented with six randomly ordered objects along with several logical statements that jointly implied a correct and unique sequential ordering (Figure 2a). Participants could submit up to two proposed solution orderings per problem. After the first submission, the user is given feedback to let them know how many objects are in their correct location in the sequence. If all six objects are in the correct order, the user proceeds to the next problem. If one or more objects are not in the correct position, the participants can either (i) accept their initial attempt as final or (ii) revise their answer and submit the revision. In all cases, the user is always shown the correct solution before moving to the next problem, in order to provide feedback to support learning across problems. Additionally, certain objects were marked with hints linked to specific positions, for example, an object with a horizontal line was more likely



(a) Without AI assistance



(b) With AI assistance

Figure 2. Examples of user interfaces for the logic-based puzzle task

to be placed in Position 4 (see Figure 2a, object F), thereby facilitating learning and knowledge transfer.

The study consisted of three phases (Figure 1). Phase 1 functioned as a pre-AI assessment and lasted 8 minutes, with participants required to complete at least four problems; those who had not completed four within 8 minutes were allowed to continue until they did. Phase 2 (20 minutes) introduced the experimental conditions, including optional AI assistance, as described in the Conditions section below. Phase 3 replicated Phase 1 as a post-AI assessment. Problem order was randomized within each phase to control for difficulty-related learning effects. Participants earned one point per correctly solved problem.

3.2. Conditions

The study used a between-subjects design with three Phase 2 conditions: *No-AI*, *Low-information AI*, and *High-information AI*. AI assistance was available exclusively in Phase 2 and could be requested at most once per problem, allowing participants to decide whether and when to seek assistance. Upon requesting help, the AI assistant revealed the positions of randomly selected objects, with the number of revealed objects determined by the assigned condition. Each AI request incurred a 0.2-point deduction, applied only if the problem was solved correctly. Participants were informed of this cost in advance.

Participants were randomly assigned to one of the following conditions:

- ***No-AI***: No AI assistance was available to participants (see Figure 2a).
- ***Low-information AI***: The AI can randomly reveal, on request during Phase 2, the location of one object (see Figure 2b).

- ***High-information AI***: The AI can randomly reveal, on request during Phase 2, the locations of three objects.

To rule out AI accuracy as a potential confound when examining how informativeness affects human reliance and performance, we used a simulated AI agent with perfect (100%) accuracy. Simulated AI agents are commonly employed in human–AI interaction research (e.g., [35,36,37,38]), particularly in settings where AI is assumed to be fully accurate to isolate specific behavioral mechanisms [35,36]. Participants were not told about the AI's accuracy and were informed only that AI assistance would be available during Phase 2 at a cost.

3.3. Procedure

The study was conducted online via Prolific. After providing informed consent, participants were randomly assigned to an experimental condition and presented with instructions. They were required to pass two comprehension checks before moving on, and only those participants who answered both correctly proceeded to the main task. The study consisted of three phases, separated by brief screen breaks. An attention check was embedded in Phase 2 to maintain data quality. Upon completion of the final phase, participants filled out a short post-study survey.

3.4. Participants

We recruited 160 U.S.-based, English-speaking adults (18+) with at least an undergraduate degree. After excluding 28 participants due to attention failures, inattentiveness, inconsistent AI-use reports, and extreme outliers, the final sample comprised 132 participants (42 *No-AI*, 43 *High-information AI*, 47 *Low-information AI*). Participants had a mean age of 39.7 (SD = 11.9); 58 identified as male, and 74 as female; 67 had an undergraduate degree and 65 had a master's degree or higher.

Participants received $9 for participation (median duration: 60 minutes) and could earn up to $3.50 in performance-based bonuses. The experiment received Institutional Review Board (IRB) approval at the University of California, Irvine, protocol number #7206.

3.5. Measures

We report below seven primary metrics from our study:

- ***Response Time***: Time spent on each problem (seconds) from when the participant first sees the problem to when they submit their final solution and proceed to the next problem.
- ***Correctness***: Number of objects placed in the correct position (0–6).
- ***Reward Rate***: Correctness divided by response time (converted to minutes), capturing the speed–accuracy trade-off in cognitive modeling [39]. Higher values indicate better performance.
- ***Initial Ability***: Participants' Phase 1 reward rate. We further split participants into low- and high-ability groups based on the median of initial ability.
- ***Timing of Request for AI Assistance***: Time from problem start to the first AI request, conditional on AI assistance being requested.

- ***AI Usage Fraction***: The percentage of Phase 2 problems in which participants requested AI assistance. The upper bound for this metric is 1.0, indicating that a participant used AI assistance in every problem in Phase 2.
- ***Solo Thinking Ratio***: From a cognitive offloading perspective, one observable marker of engagement is the amount of time individuals devote to independent problem-solving before seeking external assistance. We use the ratio of *solo thinking time* to response time, where *solo thinking time* refers to the period participants worked independently before requesting AI assistance (i.e., same as *Timing of Request for AI Assistance* above), or, if no AI was used, the entire problem-solving time until submission. We also report ***Solo Share***, defined as the participant-level average of the *Solo Thinking Ratio* across all problems completed within a phase. Higher values indicate greater sustained independent problem-solving before seeking AI assistance.

The primary focus of our analysis is on reward rates at the participant level. We also examine response time and accuracy at the problem level to separate the effects of speed and accuracy. Additionally, we discuss the *solo share* metric, which reflects how participants balance working independently versus seeking external help over time, allowing us to explore how individual problem-solving effort relates to skill development.

4. Results

We begin by providing a high-level summary of skill development trajectories across individuals over the duration of the experiment. Figure 3 displays the distribution of individual-level changes in reward rate from Phase 1 to Phase 3. Most participants exhibit gains in skill (reward rate) over time, and the average improvement is statistically significant (mean improvement = 1.64, $p < 0.01$), providing evidence of overall skill development over time (the median reward rate increases from 1.89 in Phase 1 to 3.35 in Phase 3). The substantial spread of the distribution reveals substantial cross-individual variation in the magnitude of these learning gains.

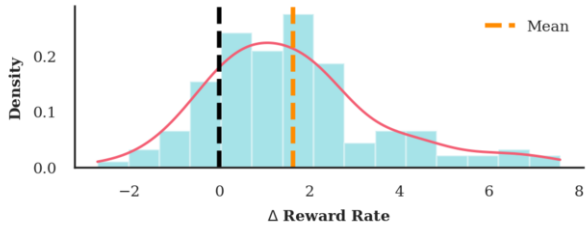
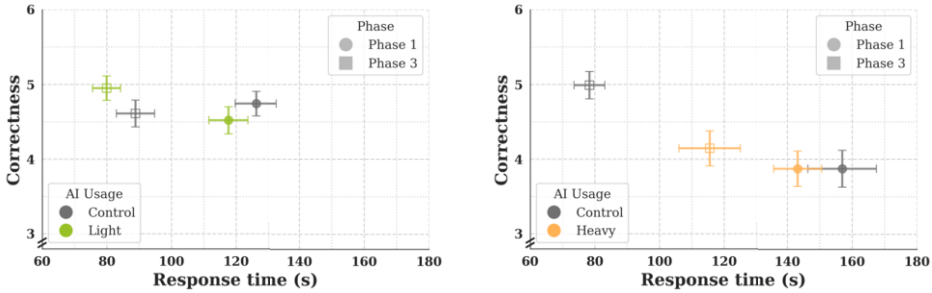


Figure 3. Distribution of individual changes in *Reward Rate* (correct objects per minute), defined as Phase 3 − Phase 1. The black dashed line marks zero-change reference, and the colored dashed line denotes the sample mean.

4.1. When AI assistance hinders skill development

We start by analyzing the relationship between individual AI usage and subsequent skill development. Participants are categorized according to their AI usage fraction (the frac-



(a) Light usage group vs. matched control group

(b) Heavy usage group vs. matched control group

Figure 4. Propensity score matching (PSM): average correctness and response time (with standard error bars) for light and heavy AI usage groups versus matched control groups in Phase 1 (pre-AI) and Phase 3 (post-AI).

tion of Phase 2 problems where participants requested AI assistance): Zero (no AI requests), Light (usage fraction $\in (0, 0.4]$), and Heavy (usage fraction $\in (0.4, 1]$), where 0.4 corresponds to the mean usage rate among those who requested AI at least once. To account for differences in participants' initial ability, we conduct propensity score matching (PSM) [40], matching Light and Heavy users separately to Zero-usage participants based on Phase 1 performance in the two-dimensional space of correctness and response time, using logistic regression with nearest-neighbor matching. Each light-AI user ($N = 36$) is matched with one control, and each heavy-AI user ($N = 24$) with one control. Results are robust to alternative thresholds, such as a median split. Figure 4 represents the results. In Phase 1, treatment and control groups do not differ significantly, as expected, since matching was based on initial performance. In Phase 3, however, relative to the matched controls, heavy AI users have significantly longer response time (115.54 (se = 9.49) vs. 78.31 (se = 4.76); $p < 0.01$) and lower correctness (4.15 (se = 0.23) vs. 4.99 (se = 0.18); $p < 0.01$), suggesting weaker individual skill development. Light AI users are statistically indistinguishable from their matched controls in response time, exhibiting slightly higher correctness (4.95, se = 0.16 vs. 4.61, se = 0.18; $p = 0.08$), suggesting that light AI use does not slow participants down and may modestly improve accuracy. These patterns strongly indicate, in the context of logic puzzle-solving, that greater reliance on AI is associated with weaker learning, whereas limited and selective use may be compatible with, or even supportive of, skill growth.

4.2. AI informativeness moderates immediate performance and learning outcomes

We next discuss how these usage patterns interact with AI informativeness. Figure 5 presents reward rates over time. The first and last points correspond to Phase 1 and Phase 3, respectively, while the first and second halves of Phase 2 are represented by two intermediate points.[2] In Phase 1, reward rates do not differ significantly across groups, indicating balance in baseline reward rate. During Phase 2, participants with access to high-information AI achieve higher reward rates ($p < 0.10$), with gains emerging immediately in the first half of Phase 2 and persisting into the second half. However, the short-run im-

[2] Since Phase 2 lasts 20 minutes, we divide it into two halves to capture more granular dynamics and to make the time intervals more comparable to Phases 1 and 3.

provement under low-information AI is comparable to that observed in the control group, indicating no additional gain beyond the baseline learning. In Phase 3, participants in the low-information condition exhibit significantly lower reward rates than the no-AI control group ($p = 0.08$), whereas the high-information group does not differ significantly from the control.These patterns suggest that while heavier AI usage is associated with weaker learning overall, the magnitude and even the direction of these effects depend on the quality of assistance provided: high-information AI improves immediate performance without reducing average post-AI outcomes, whereas low-information AI neither enhances immediate performance relative to the control group nor preserves subsequent skill growth.

We further examine Phase 3 reward rates by AI informativeness and initial ability (see Figure 6). Two patterns emerge. First, the low-information AI condition shows a uniform decline: both high- and low-ability participants exhibit lower Phase 3 reward rates than those in the no-AI group. The decline is statistically significant for low-ability participants ($p < 0.05$) but not for high-ability participants, indicating that low-information AI is associated with broadly negative post-AI outcomes, particularly among lower-ability participants. Second, the high-information AI condition displays a polarization effect: the post-AI performance gap between high- and low-ability participants is substantially larger under high-information AI (5.19 vs. 2.63; $p < 0.01$) than in the no-AI condition (4.82 vs. 3.36; $p = 0.06$). That is, in the context of logic puzzles, high-information AI assistance widens the ability gap. The next two subsections explore the mechanisms behind these results.

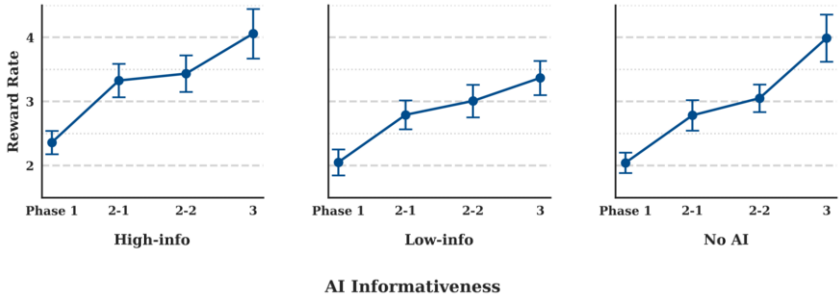


Figure 5. Reward rate trajectories across experimental phases by AI informativeness. The x-axis denotes Phase 1, Phase 2–first half (2-1), Phase 2–second half (2-2), and Phase 3. Error bars represent standard errors.

4.3. Low-information AI displaces independent effort without performance gains

Here we investigate the mechanism underlying the uniform decline observed under low-information AI. Unlike the high-information condition, this version of AI provided only limited assistance: each request revealed at most one object's location and, in some cases, repeated information already implied by the problem (e.g., the problem statement could specify that A is at position 1, and AI could then provide the same redundant suggestion). As a result, the AI had a limited effect on reducing the core complexity of the task. We conjecture that this limited assistance interrupted participants' independent reasoning and strategy formation without meaningfully simplifying the problem. Evidence from

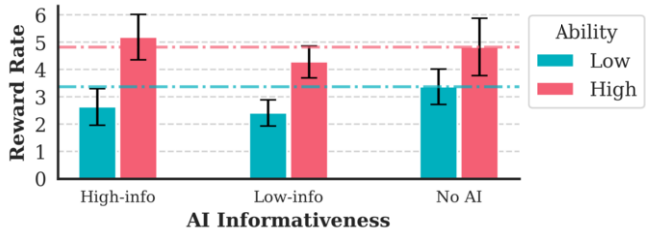


Figure 6. Post-AI average reward rates by initial ability level across different AI informativeness conditions. Error bars show 90% confidence intervals; dashed lines show post-AI average reward rates for high- and low-ability groups in the no-AI condition.

Phase 2 behavior supports this interpretation. The average Phase 2 *solo share*, which represents the proportion of independent thinking time within total response time across problems, declines from a baseline of 0.95 in the no-AI condition to 0.85 under low-information AI, and further to 0.65 for problems in which AI is actively used. This pattern indicates that even limited AI assistance is associated with a substantial reduction in independent problem-solving effort.

To examine whether this reduction in independent effort translates into immediate performance gains, Figure 7 shows average correctness *at the problem level*, broken down by condition and AI usage for Phase 2 (where AI assistance was available). We use correctness rather than reward rate because reward rate depends on response time, which in Phase 2 may reflect interruptions to the user due to accessing AI or additional review time, rather than true performance. Focusing on the low-information condition, problems assisted by AI exhibit lower correctness than both non-assisted problems within the same group and the no-AI control, while non-assisted problems in the low-info group perform relatively better. When the high-information AI is used, correctness is significantly higher than the no-AI control. These patterns suggest that low-information AI reduces independent effort without having compensating immediate gains.

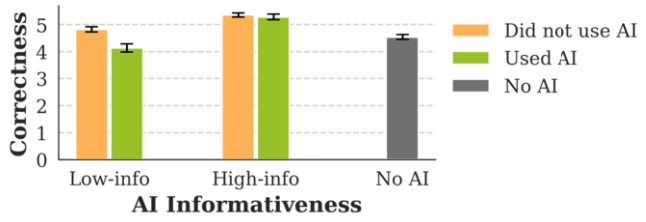


Figure 7. Phase 2 average problem-level correctness by AI informativeness and usage. Error bars represent standard errors.

Post-study survey responses reinforce this interpretation. Participants rate the low-information AI as significantly less helpful than high-information AI (mean 3.44 vs. 4.00; $p < 0.05$). Nevertheless, despite recognizing its limited usefulness, participants still chose to use the low-information AI, where, on average, the AI usage fraction is 0.22. This suggests that even objectively weak AI tools can attract engagement and divert attention, making low-quality assistance potentially distracting and detrimental to effective learning.

4.4. High-info AI induces ability-based divergence

As mentioned above, high-information AI does not affect Phase 3 performance on average, but it widens the performance gap between high- and low-ability participants. Figure 8 further shows that high-ability participants in the high-info AI condition benefit not only in long-run skill growth but also in immediate Phase 2 performance, particularly in the first half, relative to the no-AI group. In contrast, low-ability participants do not exhibit comparable short-run gains. To understand the mechanism, we examine independent problem-solving time (*solo share*) and AI usage within the high-information AI condition. We find that low-ability participants spend significantly less time solving problems independently than high-ability participants (*solo share* = 0.81 vs. 0.91, $p < 0.01$). Low-ability participants also use AI more frequently (0.37 vs. 0.22), reflecting greater reliance. Heavier reliance seems to hinder learning: lower-ability participants tend to turn to AI earlier and use it more intensively, which ultimately reduces the potential benefits of high-quality assistance. Moreover, high-information AI appears to inflate confidence among lower-ability participants. Despite weaker objective performance, in the post-study survey, low-ability participants report a stronger belief that they have "found a strategy" (mean 4.37 vs. 4.04 for high-ability participants). In the control group, self-assessments are similar across ability levels (4.13 vs. 4.06), and high-ability participants' ratings remain stable across conditions. This pattern suggests that high-information AI may increase the divergence between perceived and actual competence among lower-ability individuals, increasing the gap between perceived and actual competence.

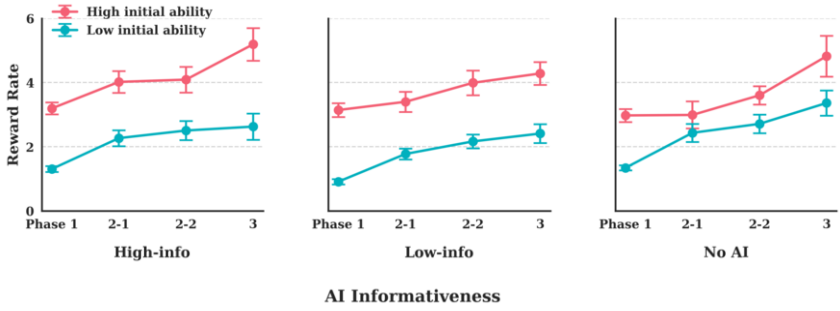


Figure 8. Reward rate trajectories across experimental phases by AI informativeness and initial ability. The x-axis denotes Phase 1, Phase 2–first half (2-1), Phase 2–second half (2-2), and Phase 3. Error bars represent standard errors.

5. Discussion

Our findings both align with and extend prior work in human–AI interaction and the learning sciences. Consistent with research showing that excessive external support can undermine independent reasoning [14,41], we find, in the context of our logic puzzle study, that heavy reliance on AI reduces independent problem-solving effort and is associated with worse post-AI performance Specifically, heavy AI users and individuals with lower initial ability exhibit the largest post-AI performance declines, echoing concerns that automation may substitute for internal cognitive processes rather than support

them [11]. Framed within the broader literature on AI and skill acquisition [42,43], our findings are consistent with upskilling inhibition, where reliance on automated support limits the development of underlying competencies.

Our results add nuance by showing that AI's impact on learning might depend not only on access, but also on how users choose to engage with AI and the informativeness of the assistance, consistent with evidence that cognitively engaged AI use, rather than passive delegation, preserves learning outcomes [23]. In our setting, a limited amount of information fails to improve immediate performance while still interrupting independent reasoning, thereby harming subsequent learning. This pattern suggests that AI assistance with low information value may end up being more distracting than helpful, drawing attention away from the effort that could be spent on independent learning. In contrast, more informative AI assistance can potentially support skill development, especially for users who engage with it selectively and strategically. For individuals with lower initial ability, however, access alone may be insufficient and potentially counterproductive without the type of additional guidance, monitoring, or incentive mechanisms that encourage sustained cognitive engagement. Taken together, our results highlight the importance of AI usage design and scaffolding, consistent with emerging work on AI systems that promote user engagement and learning [44,45], suggesting that effective AI deployment in learning contexts requires careful attention to system capability, user heterogeneity, and usage design.

Our findings should be interpreted within the study's scope. First, we focus on short-term effects after limited exposure to AI, rather than long-term skill development. It remains unclear whether these patterns persist or change with sustained use, which calls for future longitudinal work. Second, while our controlled, logic-based task allows for precise measurement, it differs from real-world settings that often involve open-ended problems and conversational AI. Future research should examine these dynamics in more high-stakes and domain-specific contexts to better understand how well they generalize.

6. Conclusion

In conclusion, our findings highlight a key trade-off in AI-assisted problem-solving: while AI can boost short-term performance, heavy reliance can reduce independent effort and hinder learning outcomes, especially for lower-ability individuals. These effects are influenced by the type and amount of assistance provided. Overall, the impact of AI on human learning depends not just on access, but on how it is used. The results suggest, in general, that we need to pay closer attention to AI's effects on cognitive effort, reliance patterns, and long-term learning, as well as to its impact in complex, high-stakes, and specialized fields.

Acknowledgments

We thank the reviewers for their helpful suggestions in improving the paper. This work was supported by National Science Foundation under award NSF 2505006, by the Hasso Plattner Institute (HPI) Research Center in Machine Learning and Data Science at UCI, and by funding support from Google and from SAP.

References

[1] Noy S, Zhang W. Experimental evidence on the productivity effects of generative artificial intelligence. Science. 2023;381(6654):187-92.

[2] Kasneci E, Seßler K, Küchemann S, Bannert M, Dementieva D, Fischer F, et al. ChatGPT for good? On opportunities and challenges of large language models for education. Learning and individual differences. 2023;103:102274.

[3] Peng S, Kalliamvakou E, Cihon P, Demirer M. The impact of AI on developer productivity: Evidence from Github Copilot. arXiv preprint arXiv:230206590. 2023.

[4] Vaithilingam P, Zhang T, Glassman EL. Expectation vs. experience: Evaluating the usability of code generation tools powered by large language models. In: CHI conference on human factors in computing systems extended abstracts; 2022. p. 1-7.

[5] Kalra N, Paddock SM. Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability? Transportation research part A: policy and practice. 2016;94:182-93.

[6] Budzyń K, Romańczyk M, Kitala D, Kołodziej P, Bugajski M, Adami HO, et al. Endoscopist deskilling risk after exposure to artificial intelligence in colonoscopy: a multicentre, observational study. The Lancet Gastroenterology & Hepatology. 2025.

[7] Steyvers M, Tejeda H, Kerrigan G, Smyth P. Bayesian modeling of human–AI complementarity. Proceedings of the National Academy of Sciences. 2022;119(11):e2111547119.

[8] Wilder B, Horvitz E, Kamar E. Learning to complement humans. arXiv preprint arXiv:200500582. 2020.

[9] Fügener A, Walzner DD, Gupta A. Roles of artificial intelligence in collaboration with humans: Automation, augmentation, and the future of work. Management Science. 2026;72(1):538-57.

[10] Gerlich M. AI tools in society: Impacts on cognitive offloading and the future of critical thinking. Societies. 2025;15(1):6.

[11] Macnamara BN, Berber I, Çavuşoğlu MC, Krupinski EA, Nallapareddy N, Nelson NE, et al. Does using artificial intelligence assistance accelerate skill decay and hinder skill development without performers' awareness? Cognitive Research: Principles and Implications. 2024;9(1):46.

[12] Zhang Y, Liao QV, Bellamy RK. Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making. In: Proceedings of the 2020 conference on fairness, accountability, and transparency; 2020. p. 295-305.

[13] Senoner J, Schallmoser S, Kratzwald B, Feuerriegel S, Netland T. Explainable AI improves task performance in human–AI collaboration. Scientific reports. 2024;14(1):31150.

[14] Buçinca Z, Malaya MB, Gajos KZ. To trust or to think: cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making. Proceedings of the ACM on Human-computer Interaction. 2021;5(CSCW1):1-21.

[15] de Jong S, Paananen V, Tag B, van Berkel N. Cognitive forcing for better decision-making: reducing overreliance on AI systems through partial explanations. Proceedings of the ACM on Human-Computer Interaction. 2025;9(2):1-30.

[16] Lai V, Tan C. On human predictions with explanations and predictions of machine learning models: A case study on deception detection. In: Proceedings of the conference on fairness, accountability, and transparency; 2019. p. 29-38.

[17] Brynjolfsson E, Li D, Raymond L. Generative AI at work. The Quarterly Journal of Economics. 2025;140(2):889-942.

[18] Vygotsky LS. Mind in society: The development of higher psychological processes. vol. 86. Harvard university press; 1978.

[19] Yan L, Martinez-Maldonado R, Jin Y, Echeverria V, Milesi M, Fan J, et al. The effects of generative AI agents and scaffolding on enhancing students' comprehension of visual learning analytics. Computers & Education. 2025:105322.

[20] Gajos KZ, Mamykina L. Do people engage cognitively with AI? Impact of AI assistance on incidental learning. In: Proceedings of the 27th International Conference on Intelligent User Interfaces; 2022. p. 794-806.

[21] Poulidis S, Ge H, Bastani H, Bastani O. Action vs. attention signals for human-AI collaboration: Evidence from chess. The Wharton School Research Paper. 2025.

[22] Kosmyna N, Hauptmann E, Yuan YT, Situ J, Liao XH, Beresnitzky AV, et al. Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task. arXiv preprint arXiv:250608872. 2025.

[23] Shen JH, Tamkin A. How AI impacts skill formation. arXiv preprint arXiv:260120245. 2026.
[24] Karny S, Mayer LW, Ayoub J, Song M, Su H, Tian D, et al. Learning with AI assistance: A path to better task performance or dependence? In: Proceedings of the ACM Collective Intelligence Conference; 2024. p. 10-7.
[25] Dimitrov DM, Rumrill PDJ. Pretest-posttest designs and measurement of change. Work. 2003;20(2):159-65.
[26] Gu F, Wongkamjan W, Boyd-Graber JL, Kummerfeld JK, Peskoff D, May J. Personalized help for optimizing low-skilled users' strategy. In: Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 2: Short Papers); 2025. p. 65-74.
[27] Swaroop S, Buçinca Z, Gajos KZ, Doshi-Velez F. Accuracy-time tradeoffs in AI-assisted decision making under time pressure. In: Proceedings of the 29th International Conference on Intelligent User Interfaces; 2024. p. 138-54.
[28] Cao S, Gomez C, Huang CM. How time pressure in different phases of decision-making influences human-AI collaboration. Proceedings of the ACM on Human-computer Interaction. 2023;7(CSCW2):1-26.
[29] Fogliato R, Chappidi S, Lungren M, Fisher P, Wilson D, Fitzke M, et al. Who goes first? Influences of human-AI workflow on decision making in clinical imaging. In: Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency; 2022. p. 1362-74.
[30] Lee HP, Sarkar A, Tankelevitch L, Drosos I, Rintel S, Banks R, et al. 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; 2025. p. 1-22.
[31] Joinson A. Social desirability, anonymity, and Internet-based questionnaires. Behavior research methods, instruments, & computers. 1999;31(3):433-8.
[32] Vasconcelos H, Jörke M, Grunde-McLaughlin M, Gerstenberg T, Bernstein MS, Krishna R. Explanations can reduce overreliance on AI systems during decision-making. Proceedings of the ACM on Human-Computer Interaction. 2023;7(CSCW1):1-38.
[33] Tejeda Lemus H, Kumar A, Steyvers M. How displaying AI confidence affects reliance and hybrid human-AI performance. In: HHAI 2023: Augmenting Human Intellect. IOS Press; 2023. p. 234-42.
[34] Kahr PK, Rooks G, Snijders C, Willemsen MC. The trust recovery journey. The effect of timing of errors on the willingness to follow AI advice. In: Proceedings of the 29th International Conference on Intelligent User Interfaces; 2024. p. 609-22.
[35] Panigutti C, Beretta A, Giannotti F, Pedreschi D. Understanding the impact of explanations on advice-taking: a user study for AI-based clinical Decision Support Systems. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems; 2022. p. 1-9.
[36] Srivastava DK, Lilly JM, Feigh KM. Improving human situation awareness in AI-advised decision making. In: 2022 IEEE 3rd International Conference on Human-Machine Systems (ICHMS). IEEE; 2022. p. 1-6.
[37] Cao S, Liu A, Huang CM. Designing for appropriate reliance: The roles of AI uncertainty presentation, initial user decision, and user demographics in AI-assisted decision-making. Proceedings of the ACM on Human-Computer Interaction. 2024;8(CSCW1):1-32.
[38] Eisbach S, Langer M, Hertel G. Optimizing human-AI collaboration: Effects of motivation and accuracy information in AI-supported decision-making. Computers in Human Behavior: Artificial Humans. 2023;1(2):100015.
[39] Standage D, Wang DH, Heitz RP, Simen P. Toward a unified view of the speed-accuracy trade-off. Frontiers in Neuroscience. 2015;9:139.
[40] Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behavioral Research. 2011;46(3):399-424.
[41] Wecks JO, Voshaar J, Plate BJ, Zimmermann J. Generative AI usage and exam performance. arXiv preprint arXiv:240419699. 2024.
[42] Rafner JF, Dellermann D, Hjorth A, Verasztó D, Kampf CE, Mackay W, et al. Deskilling, upskilling, and reskilling: a case for hybrid intelligence. Morals & Machines. 2021;1(2):24-39.
[43] Natali C, Marconi L, Dias Duran LD, Cabitza F. AI-induced deskilling in medicine: a mixed-method review and research agenda for healthcare and beyond. Artificial Intelligence Review. 2025;58(11):356.
[44] Ma Q, Zhou Y, Kaushik S, Joshi A, Majumdar A, Apthorpe N, et al. Learning password best practices

through in-task instruction. arXiv preprint arXiv:260106650. 2026.
[45] Tran K, Gao G, Lombard A, Yu T, Jiang H, Yeh TY. Pacing for mastery: Optimizing LLM interactions for learning. In: Proceedings of the 57th ACM Technical Symposium on Computer Science Education V. 1; 2026. p. 1068-74.