OXFORD JOURNAL
OF STUDENT
SCHOLARSHIP

The Effects of Improper Generative AI Tool Use on High-School Students' Critical and Creative Thinking Abilities

Ishaan Vadlamani
ishaanvv10@gmail.com

ABSTRACT

As generative artificial intelligence (AI) becomes integrated into secondary education, concerns regarding "cognitive atrophy" and the erosion of student critical thinking have intensified. This research explores the intersection of AI over-reliance and the stunting of cognitive development in high school students. Utilizing a systems-thinking methodology, including Causal Loop Diagrams (CLDs) and qualitative stakeholder "listening sessions," this study identifies a cycle where students increasingly devalue human oversight in favor of AI-generated output. Findings suggest that while improper use leads to a decline in memory, organization, and original thought, a complete ban on technology is counterproductive. Instead, the study advocates for a shift from automation to augmentation. By benchmarking current solutions and analyzing the economic implications of AI integration, this report proposes a framework for "effective interaction", prioritizing a human's evaluation to mitigate AI's inherent biases and hallucinations. The research concludes that fostering human-centric evaluative skills is essential to ensuring that AI serves as a catalyst for, rather than a replacement of, intellectual growth in the digital age.

THE SUBPROBLEM AND TARGET AUDIENCE

Our initial question, which was "How does AI development limit people's critical thinking" emerged early on. As students ourselves, we were proximate to this problem. According to Pew Research (Sidoti, 2025), roughly one-quarter of U.S. teens used ChatGPT for schoolwork as of 2024, double the share from the year before. The extent to which these students used generative AI is unclear, yet teachers have responded by restricting Ai, fearing that students will plagiarise and stop learning. This made us wonder whether or not Ai actually damages students' critical thinking abilities, or can it strengthen them.

PRIOR RESEARCH AND WHERE THE STUDY FITS

Research on the effect on cognitive theory and more recent research distinguishing augmentation and automation especially in the educational world are vital to this study.

First, the **Cognitive theory and AI Dependency**, also called the Cognitive load theory (Sweller, 1988) established that working memory is finite and tools that externalize cognitive work, which at the time included calculators and search engines, can reduce the depth of mental processing when used passively. More recently, this framework has been applied directly to AI. Lee et al. 's (2025) large scale Microsoft Research survey had significant findings. They pointed out that knowledge workers, people whose main job is to think, analyze, and create using information, who expressed high confidence in generative AI were more skeptical of AI outputs. The study continued that users who trusted their own abilities more actually sharpened their critical thinking skills by comparing AI outputs against their own reasoning. Throughout this paper, it is identified that this usage of AI is clearly positive and is what students should strive for in their schools. For high school students in particular, our cognitive development is still in progress and therefore more vulnerable to what researchers point to as **automation bias**, the tendency to uncritically accept machine-generated outputs, which we mention multiple times throughout the paper. Research on AI's effects on learning and cognition has developed along three largely separate streams, each of which this study engages and extends.

Second, we identify throughout the paper that institutionally, responses to AI have been far too polarized and rash to benefit students. For example, teachers have blocked AI tools on computers or on the internet the school provides, or ban their use in assignments outright. Scholars such as Lipman and Distler (2023) have argued that prohibition is shortsighted, failing to build AI literacy while being easily circumventable. Students turn to other devices which do not have such restrictions or paraphrase AI's outputs as a means of circumventing restrictions which schools and teachers believe are putting a stop to AI generated work. Students in our listening sessions corroborate exactly what is aforementioned, and we add *why* prohibition simply fails, as our causal loop diagram demonstrates that banning AI severs the reinforcing feedback loop through which ethical AI use builds AI literacy, making students less equipped to engage with AI responsibly.

Finally, the third and potentially most important distinction to make is augmentation versus automation. The economics of AI literature (Acemoglu & Restrepo 2018; Autor, 2015) distinguishes automation (AI directly replacing tasks that humans are meant to complete) versus augmentation (AI enhancing what humans are meant to complete. Luckin (2017) applied this distinction to education, arguing that AI must be designed to scaffold human intelligence rather than substituting for it. Our study applies these frameworks to secondary schooling, where they remain insufficiently developed.

Our usage of causal loop diagrams as well as qualitative stakeholder interviews to model the feedback dynamics that determine how AI use in schools must evolve over time serves as grounds to our proposed solutions in systemic intervention.

THE ORIGINAL CONSOLIDATED CAUSAL LOOP MAP

We began by creating a causal loop diagram (CLD) of the problem space. We each created individual CLDs and then compared them. We chose the largest and most extensive map, as the "base" map, and used its structure to guide the development of the consolidated CLD.

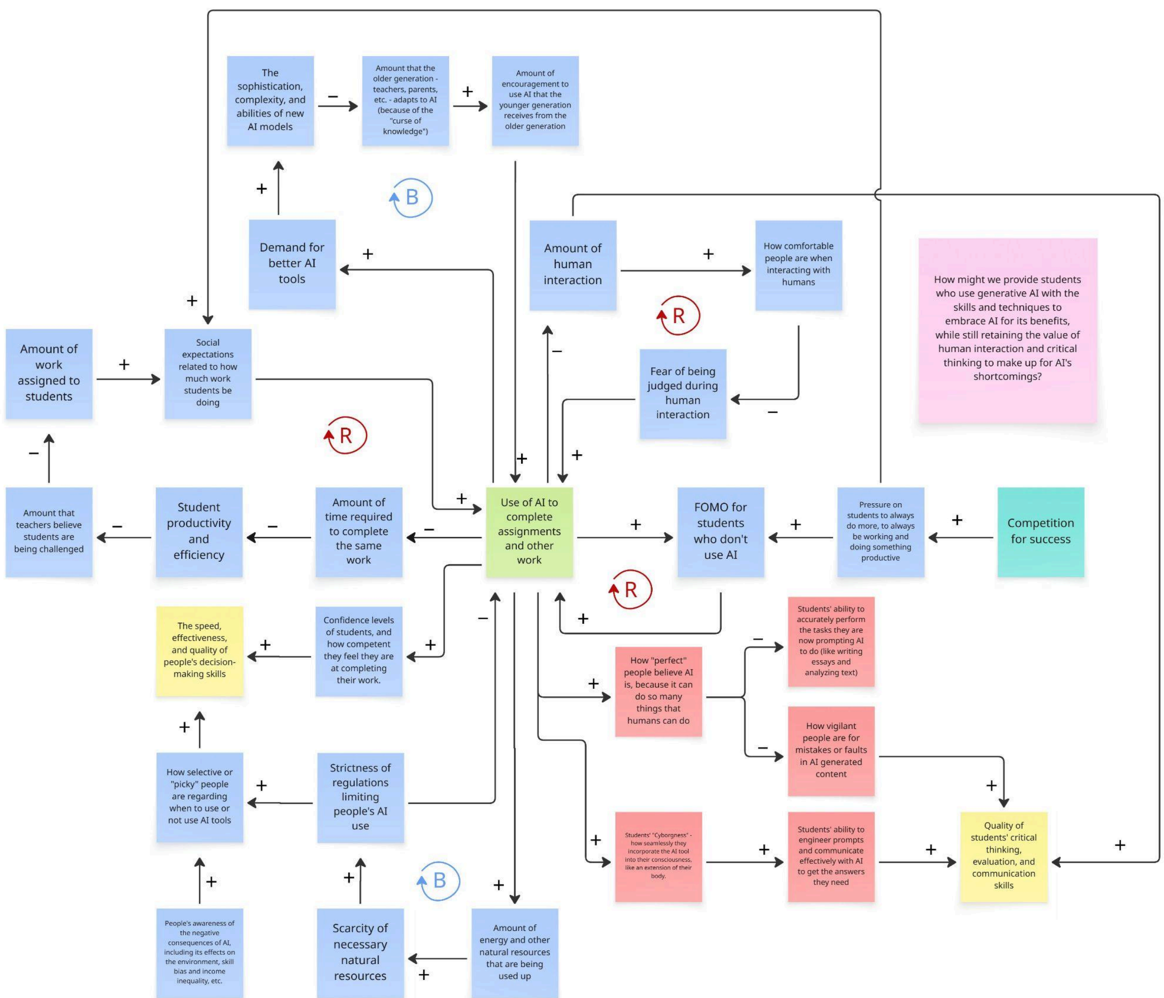

Figure 1: The base map

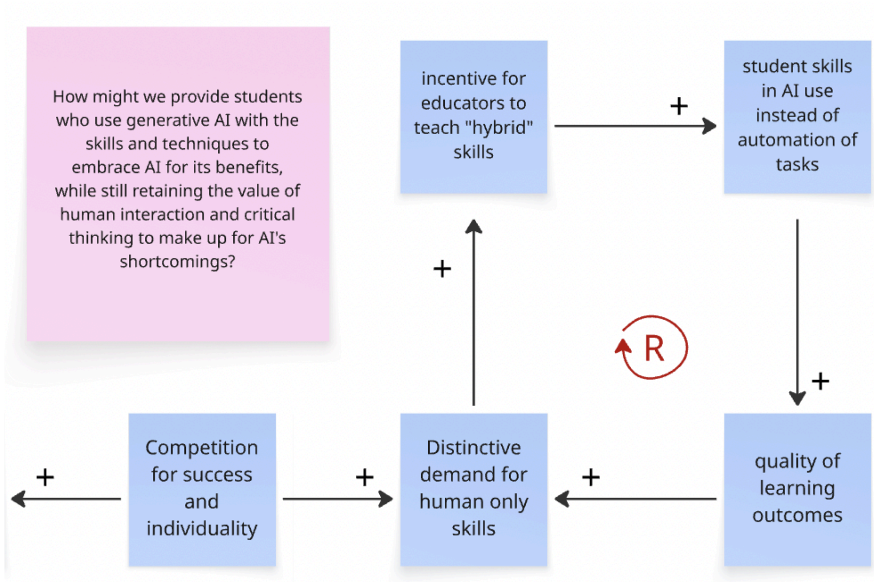

Figure 2: A section that was added to the base map.

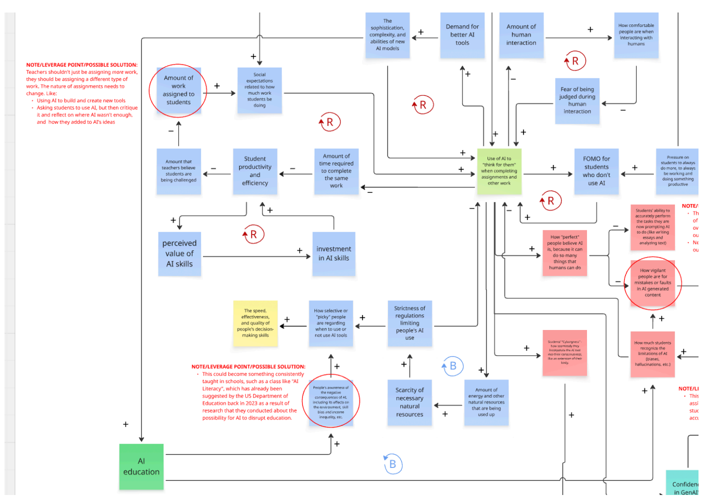

Figure 3: A section that was added to the base map.

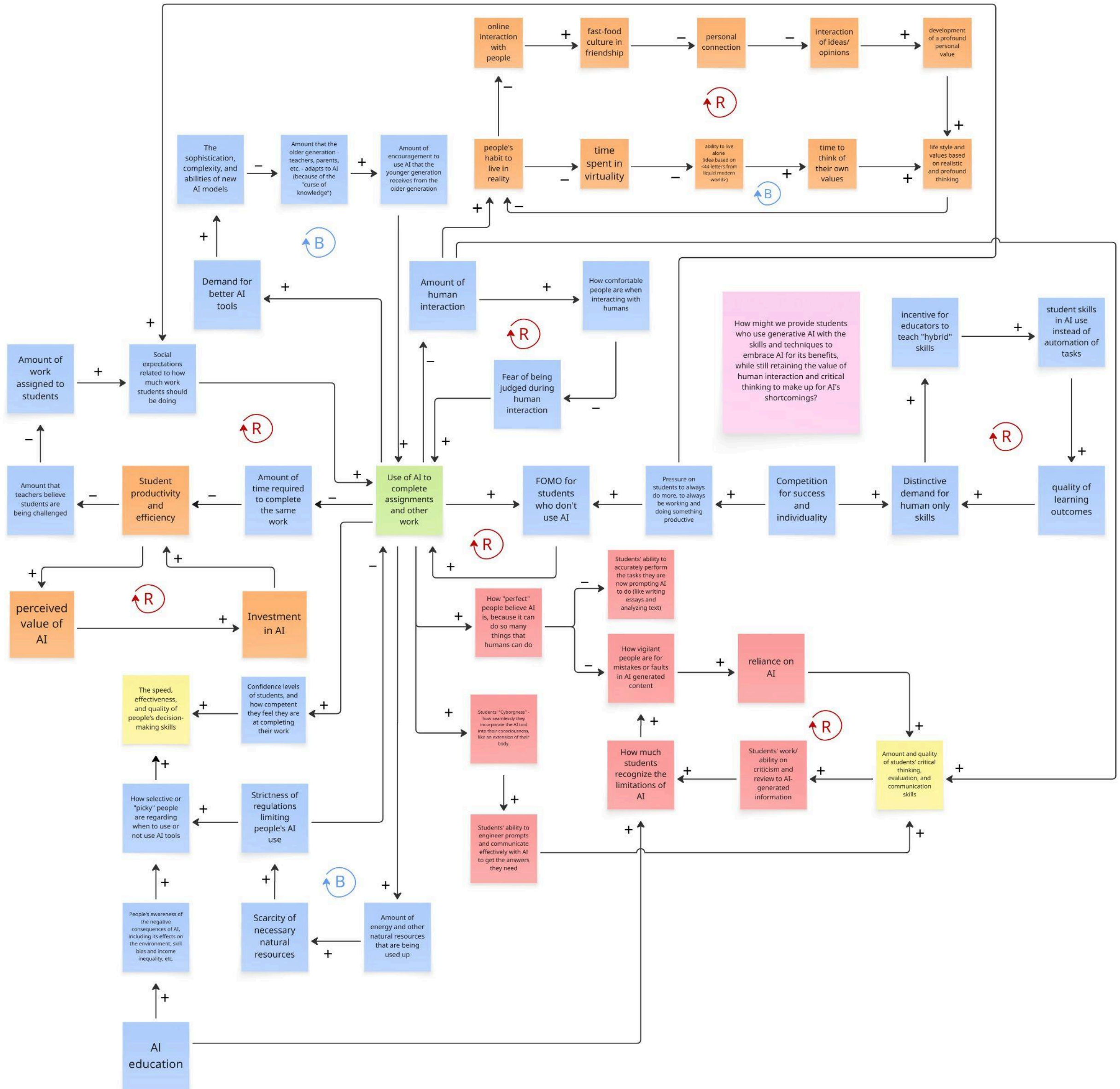

Figure 4: Our consolidated causal loop diagram

For simplicity purposes, we have elected to highlight the four most important structural insights the consolidated CLD reveals:

1. **Use of AI as a shortcut is the central hub of the system**. Nearly every other variable in the diagram either feeds into, or is affected by, the amount of students using AI to complete schoolwork. This centrality confirms that interventions must engage this hub to meaningfully change outcomes.

2. **There are more reinforcing loops than balancing loops**. This structural imbalance clearly shows that the system has a natural tendency to amplify AI use over time. AI use increases student productivity → productivity increases confidence in AI → confidence increases AI use → so on and so forth. This matches the empirical trend: Pew Research (2025) found that teen ChatGPT use for schoolwork doubled between 2023 and 2024.
3. **Academic competition is a root cause: it has no inputs, rather only outputs**. The factor "competition for success and individuality" has no arrows pointing into it. Yet, it drives pressure to use AI, FOMO among students who don't use AI, and ultimately AI dependency. This means interventions that only target students' knowledge and attitudes will be insufficient unless there is also a reform in the incentive structure created by academic competition.
4. **AI literacy is the most promising leverage point.** AI education is one of the few variables that activates a balancing loop, reducing overreliance by increasing awareness of AI's limitations. However, this loop is currently underactive; most schools lack structured AI literacy programs. Increasing AI education would absolutely help this area.

Based on the CLD structure, we also propose the following hypotheses for future empirical testing:

H1: Students who receive structured AI literacy education will demonstrate greater skepticism of AI outputs and stronger critical evaluation skills than those who do not.

H2: Students experiencing higher levels of academic competitive pressure will show greater AI dependency, regardless of their awareness of AI's limitations.

H3: Pedagogical interventions that embed AI use within assignment workflows (augmentation model) will produce better critical thinking outcomes than outright AI restrictions.

H4: Students who use AI primarily for cognitive offloading (automation) will show measurably lower scores on original synthesis tasks compared to students who use AI for scaffolding (augmentation).

METHODS

This section describes our research approach, including how participants were selected, how data were collected, and how findings were analyzed. Our methodology combined systems-thinking tools with qualitative stakeholder research.

Research Design

We used a mixed qualitative design consisting of three components: (1) Causal Loop Diagram development, drawing on secondary literature and individual reflection; (2) semi-structured stakeholder interviews ("listening sessions") with three participant groups; and (3) thematic analysis of interview data, synthesized into empathy maps and a Needs Table. This design was chosen because our research question concerns a dynamic, feedback-driven process that is better captured through systemic and qualitative methods than through surveys alone.

Participant Selection

We recruited participants across three groups, selected purposively to represent the key stakeholders in secondary AI use:

Group 1 — Teachers, Administrators, and Educators. Participants included Dr. Jessica Potts (Curriculum Coordinator, Davidson Academy Online) and Jamie Shushan (Professor, Harvard University). These participants were selected because educators set the conditions within which students make decisions about AI use. Their insights on student usage of AI but also how they are dealing with it were important to our project.

Group 2 — High-School Students Who Rely Heavily on AI. A boarding-school student from China with documented heavy AI use for academic tasks was interviewed. This participant was selected through the personal networks of team members and consented to participate, with privacy protections in place (no identifying details were disclosed). This group represents the core population at risk for cognitive atrophy through AI overuse.

Group 3 — High-School Students Who Avoid AI. Students who primarily complete schoolwork independently were interviewed to understand the perspective of those already practicing the behaviors our solutions aim to encourage, and to identify any unintended consequences of AI integration for this group.

We also conducted a listening session with Quinine Gao, a PhD student in Education at UNC Chapel Hill, who provided additional insight into pedagogical strategies for AI integration. Participants were recruited through personal and professional networks of team members, and all gave verbal consent to participate. We believed listening sessions were vital in our creation of this paper.

Data Collection

Listening sessions were conducted as semi-structured interviews lasting approximately 30–45 minutes each, held via video call. Each session followed a flexible guide designed around the participant's role, with questions covering: current AI use behaviors and perceptions; the effects of AI on thinking and learning; experiences with AI policies; and ideas for solutions. Team members took notes during sessions, which were not recorded in order to protect participant privacy.

To complement the interview data, we conducted a review of academic and journalistic literature on AI and education, drawing on sources including Microsoft Research (Lee et al., 2025), Pew Research Center (Sidoti, 2025), UNESCO (Holmes et al., 2022), and Polytechnique Insights (Caliman, 2025). We also analyzed publicly available documentation on existing AI tools, including ChatGPT Study Mode (OpenAI, 2025) and AI literacy guidelines from the U.S. Department of Education.

Data Analysis

Interview notes were analyzed using inductive thematic analysis (Braun & Clarke, 2006). Each team member independently coded their own listening session notes, identifying recurring ideas, tensions, and needs. We then consolidated these codes in a group session by:

(1) Building individual empathy maps for each PPP group (Says, Does, Thinks, Feels), which made implicit attitudes and unspoken assumptions visible.

(2) Clustering all individual empathy map items across groups into thematic categories (e.g., "distrust of AI," "stress as a driver," "need for AI literacy").

(3) Constructing a Needs Table that translates themes into specific unmet needs for each stakeholder group, directly informing our solution design.

We also incorporated secondary data, including economic studies on automation versus augmentation (Acemoglu & Restrepo, 2018) and an MIT survey on worker confidence in AI (Lee et al., 2025), to validate or complicate findings from the listening sessions. Discrepancies between participant accounts and empirical data were noted and resolved through group discussion.

Our CLD was updated iteratively: a base version was constructed at the outset, then revised after each listening session as new variables and relationships emerged. Leverage points were identified by locating factors that appeared at junctions of multiple reinforcing loops. These points would indicate that a change to this portion of the diagram would significantly alter the rest and possibly point to a solution.

THEMES FROM INSPIRATION ACTIVITIES (PPPS AND NEEDS ANALYSIS)

After interviewing several PPPs, including students, professors, and curriculum coordinators, we consolidated the findings into three empathy maps, one for each PPP group: Teachers/Administrations/Educators, Students Who Rely Heavily on AI To Complete Schoolwork, and Students Who Don't Rely on AI To Complete Schoolwork.

<u>PPP 1: Teachers/Administrations/Educators</u>

Figure 5: PPP Empathy Map of Teachers, Administrators, and Educators

PPP 2: Students Who Rely Heavily on AI

Figure 6: PPP Empathy Map of Students Who Rely Heavily On AI

<u>PPP 3: Students Who Don't Rely on AI</u>

Figure 7: PPP Empathy Map of Students Who Don't Rely on AI

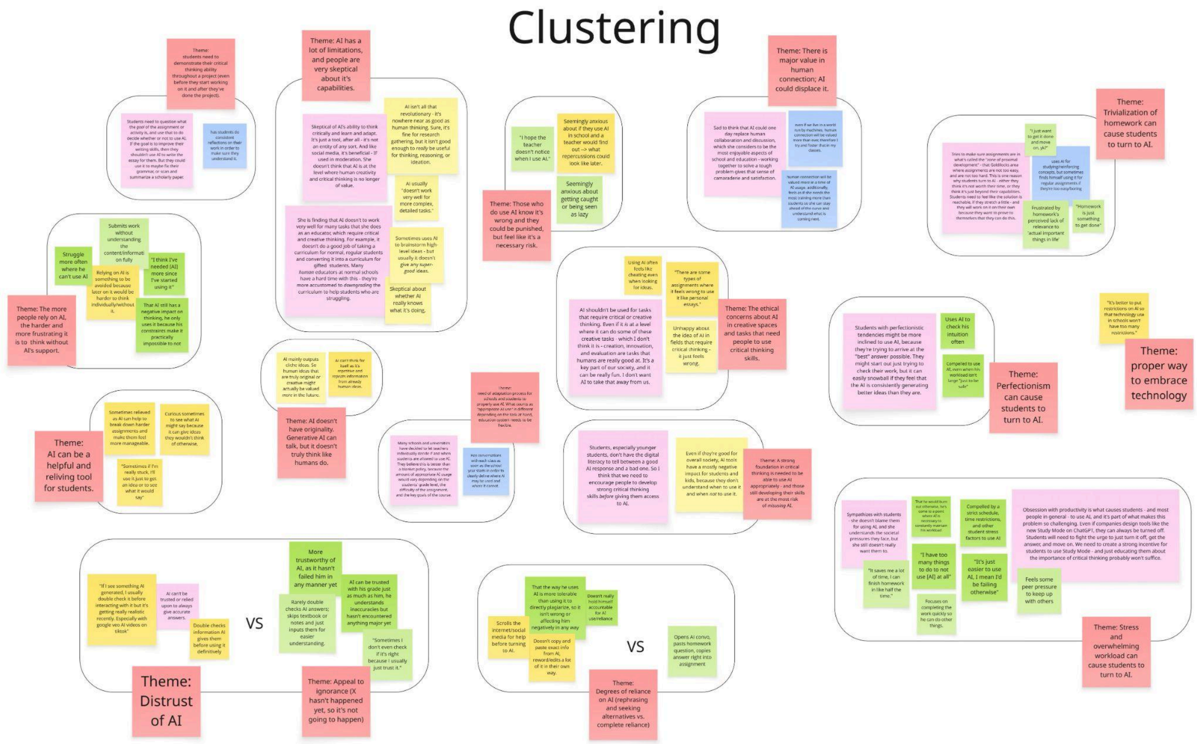

Fig 8: Clustered ideas and themes from all three consolidated PPP empathy maps

Teachers, Administrators, and Educators Need …	**High School Students Who Rely on AI Heavily to Complete Work Need …**	**High School Students Who Don't Rely on AI to Complete Work Need …**
AI Literacy & Education – Teachers need to understand AI and stay updated about AI's current capabilities to tweak current tasks and assignments to make sure only students can complete them.	An incentive to care about the quality of their work, and actually struggle through their assignments on their own, rather than only being motivated by the grade and the amount of work being done	AI Literacy & Education – to teach these students that AI can be very beneficial if used appropriately, and help them learn how to properly interact with AI in their daily life.
Pedagogical strategies and frameworks to combat educators' potential biases against AI, and allow teachers to help their students learn to interact with AI.	Mental tutoring/ counseling to relieve them from perfectionism and stress An academic support system to handle burnout/overload	Encouragement from educators and other school officials to use AI when appropriate (maybe through in-class activities practicing AI use)
Flexibility and personalizability in schools' AI regulations – to allow teachers to decide which tasks AI may or may not be used for, based on the learning goals of each course and/or assignment.	Clear education and awareness of AI's limitations Reminders and guidelines regarding ethical AI use so there can be a reduction of "risk taking" behaviors and misuse.	

Fig. 10: Needs Table for each of the three PPP groups

Results: Themes from Inspiration Activities

After conducting listening sessions with three stakeholder groups and reviewing secondary research, we identified key findings that shaped both our updated CLD and our final solution designs.

Findings by Stakeholder Group

Teachers, Administrators, and Educators

The most consistent finding from educator interviews was a sense of institutional helplessness. Educators understand the goal of using AI to augment rather than automate student thinking, but struggle to put it into practice as AI capabilities rapidly expand. Jamie Shushan (Harvard) described constant class discussions about appropriate AI use and in-class reflections to assess understanding without AI. Dr. Jessica Potts (Davidson Academy) noted that schools increasingly leave AI policy to individual teachers, creating inconsistency across classrooms.

Three findings from this group shaped our research. First, the "Goldilocks Zone" of assignments students can complete without AI is shrinking as AI capabilities grow (E1). Second, both participants showed a personal aversion to AI that limited their thinking about AI-enhanced teaching strategies, suggesting teacher training is as important as student training (E2). Third, in-class reflection is an effective verification tool, but too time-intensive to scale easily (E3).

Students Who Rely Heavily on AI

This group's central finding challenges the common assumption that AI overuse stems from laziness. Our participant turned to AI due to structural time pressure: a boarding school requiring year-round athletic and extracurricular participation created frequent overload events where tests, projects, and assignments coincided. AI became a risk-management tool to prevent one bad week from derailing his academic balance.

Two additional findings emerged. Social normalization removes the stigma from AI use. Once peers are using it, the question shifts from "should I?" to "why shouldn't I?", ultimately creating a coordination problem that individual willpower cannot solve (S2). Students using AI heavily also rarely verify outputs, trusting AI because it "hasn't embarrassed me yet," a pattern consistent with automation bias (S3).

Students Who Avoid AI

Students who avoid AI face a different challenge: understanding when AI use is actually appropriate. Several expressed strong distrust of AI, treating any use as a form of cheating even in contexts where it would be legitimate, such as background research or grammar checking. These students also experience indirect harm, as peer AI use shifts grading curves and inflates output volume, making independent work feel increasingly costly (A2).

Cross-Cutting Themes

Five themes emerged consistently across all three groups. First, AI is simultaneously helpful and harmful, but context determines which, and no group supported a blanket ban. Second, overreliance stems from structural pressures like stress, workload, and peer competition rather than individual failure. Third, effective AI use requires explicit instruction that none of our participants had received. Fourth, educators need pedagogical frameworks, not just policies. Fifth, any solution must offer students a tangible incentive. It must be stronger than the pull of productivity, because intrinsic motivation alone is insufficient against competitive academic pressure.

The Needs Table above translates these themes into specific unmet needs for each stakeholder group.

How Findings Informed the Updated CLD and Solution Design

Finding S1 led us to mark "competition for success and individuality" as a root node in the updated CLD because it is a structural driver with no upstream variables, meaning solutions must work around it rather than try to eliminate it. Findings E2 and S2 introduced two new CLD variables: educator confidence in AI and social normalization of AI use, both of which feed the reinforcing loops that amplify overuse. The Lee et al. (2025) finding on confidence dynamics, that users who trust AI think less critically, while users who trust their own reasoning think more critically when comparing it against AI, directly inspired ScholarSync's core activity of having students evaluate AI outputs against their own reasoning. Finally, Theme 5 became a hard design constraint: every proposed solution was evaluated on whether it creates a concrete incentive for engagement, not just access to information or tools.

THE UPDATED CONSOLIDATED GROUP MAP

After our listening sessions and secondary research, we updated the consolidated CLD with five additions: a new loop representing educator confidence and its effect on pedagogical AI integration; a branch at "AI Sophistication" separating automation-driven use from augmentation use; a new confidence dynamic variable linking AI trust and self-trust to critical thinking effort; a team highlight on "competition for success and individuality" to mark it as a root; and explicit leverage-point annotations identifying where interventions can most effectively change system behavior.

The diagram reinforces our core research finding: the system has a natural momentum toward increasing AI use and decreasing critical thinking. Yet this momentum can be redirected if interventions target the right leverage points, primarily AI literacy education, assignment redesign, and existing yet flawed incentive structures.

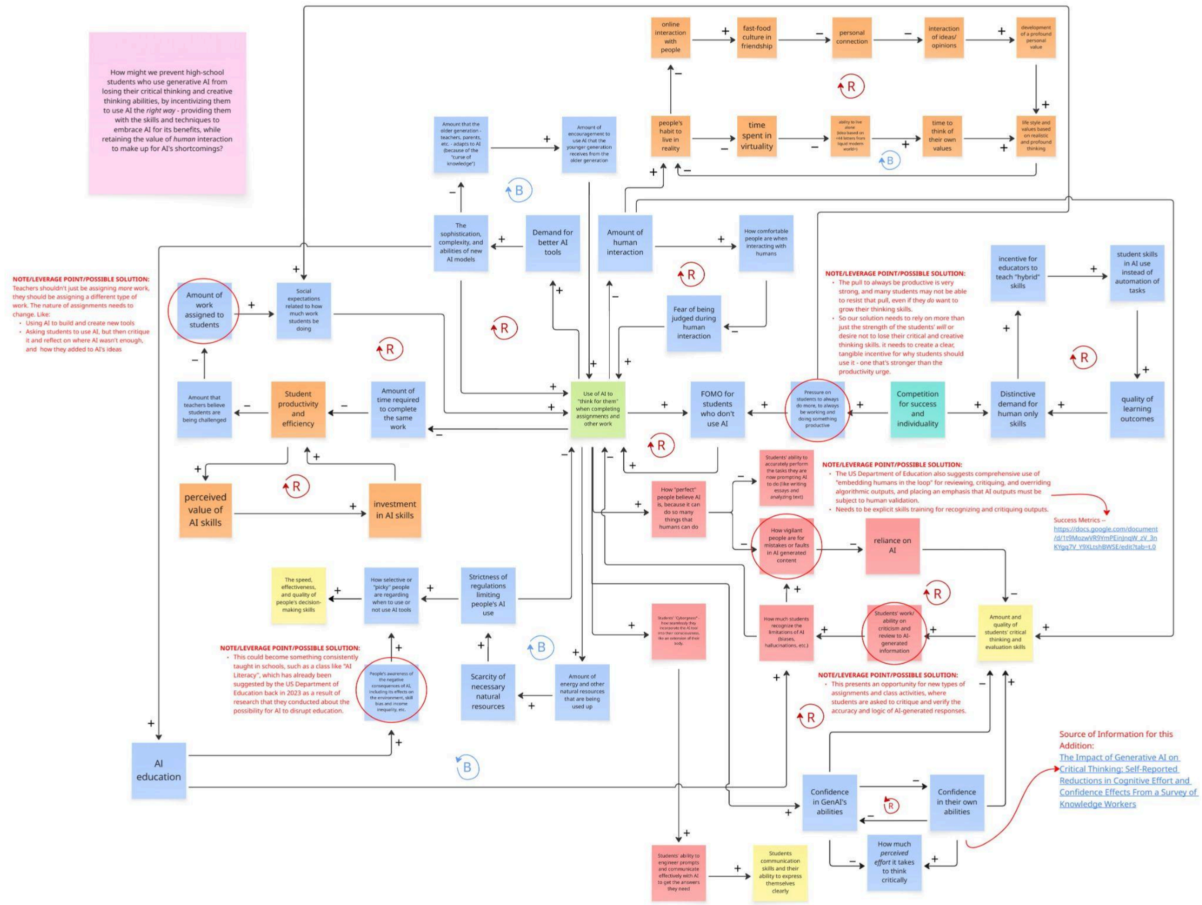

Figure 10: Our updated, final consolidated causal loop diagram

UPDATED PROBLEM (HOW MIGHT WE?) STATEMENT

Evolution of Problem Statements:

Our problem statement evolved through four key iterations as we refined our scope and integrated insights from PPP interviews:

- **Phase 1: Identifying the Shortcut.** We initially asked how AI limits critical thinking when used as a "shortcut." However, we realized "people" was too broad and our tone was overly negative.
- **Phase 2: Balancing Benefits.** We shifted focus to "people connected to AI" and aimed to balance critical thinking with the benefits of embracing new tech.
- **Phase 3: The Feedback Loop.** We explored the interdependence of humans and AI, framing the relationship as a "positive feedback loop" where both can adapt and grow together.
- **Phase 4: Final Focus.** We landed on our final statement: **"How might we provide students who use generative AI with the skills and techniques to embrace AI for its benefits, while still retaining the value of human interaction and critical thinking to make up for AI's shortcomings?"**

Why this works:

- **Target Audience:** We narrowed "people" to **students**, as they are most vulnerable to skill displacement and are at the heart of the changing job market.
- **Specificity:** We pivoted from general AI to **Generative AI**, which is most disruptive to the education system.
- **Stakeholders:** This focus allows us to engage a clear ecosystem of students, educators, parents, and policymakers.

BENCHMARKING EXISTING SOLUTIONS, THEIR STRENGTHS AND LIMITATIONS

Existing Solution #1: ChatGPT Study Mode:

OpenAI's Study mode (released July 2025) is a Socratic chatbot interface guiding users towards answers through questions rather than providing them directly. It preserves AI's core benefits while encouraging active engagement. However, students can disable it at any time. Dr. Potts confirmed this in our listening session, which means that it fails the incentive criterion. When under time pressure, a factor that we identified is one of the primary drivers of AI overuse, students will simply switch to the standard mode. Study Mode addresses the form of AI interaction without addressing the structural pressure that makes shortcuts attractive. There would need to be significant changes to the product to make sure it actually forces students to learn.

Source: https://openai.com/index/chatgpt-study-mode/

Existing Solution #2: Quillbot and School blocking AI on the wifi

Beginning in 2022, many U.S. schools blocked AI tools on school networks. This is a pure restriction approach, the least effective category. Our findings confirm and extend this, students report bypassing blocks via VPNs or home internet (Finding S2), making the restriction largely cosmetic. More fundamentally, as Lipman and Distler (2023) argue, this approach mistakes the symptom (AI use) for the cause (structural pressure, lack of literacy, etc..). Our CLD reveals an additional harm, by cutting off AI interaction entirely, schools also sever the reinforcing loop through which guided AI use could build literacy. Restriction does not simply fail to help, it may actively prevent the development of skills students will need.

Existing Solution #3: In-class recommendations and guidance on which AI tools can/cannot be used

Based on our listening session with Quiniene Gao (PhD student, UNC Chap Hill Education), this approach involves teachers explicitly guiding students on *how* to use Ai, demonstrating biased outputs, assigning Ai-error-correction tasks, and recommending specific tools. This is the closest existing approach to pedagogical integration and shows the most promise. Gao noted it has produced positive results in pilot classrooms. Its limitation is that it is **teacher-dependent**. It requires educators who are both technically confident in AI and pedagogically willing to experiment, two conditions that finding E2 suggests are quite rare. Scaling this approach would require substantial professional development investment.

AI mentioned:
https://www.kimi.com/
https://getliner.com/
https://slidestutor.com/

ALTERNATIVES MATRIX	**Possibility to implement/Feasibility**	**Incentives for students**	**How well it meets the need of improving their interaction with AI**	**Invasiveness/Potential to take things too far and automate entirely rather than augment -- overly pushing people (infringing upon people's critical thinking)**	**How well it can encourage students (Can it turn AI into an actual overall beneficial tool for students on future development)**	**How well it educates people about AI's strengths *and* limitations**	**Ease of use & Accessibility**
Solution 1: ChatGPT's study mode (socratic questions// reasoning)	2	2	3	3	3	2	3
Solution 2: internet block in school network	3	2	2	2	1	1	3
Solution 3: in-campus courses and recommendation by teachers	3	2	3	3	2	3	2

KEY	**Possibility to implement/Feasibility**	**Incentives for students**	**How well it meets the need of improving their interaction with AI**	**Invasiveness/Potential to take things too far and automate entirely rather than augment -- overly pushing people (infringing upon people's critical thinking)**	**How well it can encourage students (Can it turn AI into an actual overall beneficial tool for students on future development?)**	**How well it educates people about AI's strengths *and* limitations**	**Ease of use & Accessibility**
1	Low feasibility/difficult to implement	Provides little to no incentive for students to use it	Does not help students interact with AI in a responsible, appropriate way.	Can overtake an assignment and automate it entirely	it can hardly encourage students' creativity and may limit students thought	Little to no education on feasibility or limitations	Low accessibility
2	Medium feasibility/not too difficult to implement	Has somewhat of an incentive for students to use it	Somewhat helps students interact with AI in a responsible, appropriate way.	Somewhat overtakes an assignment and automates it, but can be used without this.	it may or may not encourage students' creativity and a promotion of thought (do not have a major impact)	Somewhat educates people about AI's strengths and limitations	Medium accessibility
3	High feasibility/easy to implement	Provides large incentives for students to use it	Provides a lot of support to teach students to interact with AI in a responsible, appropriate way.	Does not overtake assignments and automate, instead lets students augment their work with it.	it will encourage students thought, giving students chance and encouraging to think further,	Allows people to assess AI's strengths and limitations themselves	High acessibility

Figure 11: Alternatives Matrix benchmarking existing solutions

BRAINSTORMED IDEAS

Figure 11: The clustered set of ideas from our group brainstorm, along with the corresponding themes for each.

Our brainstorming process had three stages: independent rapid-fire idea generation, clustering of similar ideas, and labeling of clusters. We generated ideas across seven themes:

Theme 1: Encouraging genuine student effort and reducing academic stress as structural root causes of AI overuse.

Theme 2: Creating clearer AI regulations that give teachers more control and consistency.

Theme 3: Building AI literacy courses and class activities that teach ethical, effective AI interaction.

Theme 4: Helping students recognize and value human interaction, collaboration, and original thought.

Theme 5: Encouraging the creative and generative use of AI — using it to expand, not replace, human ideas.

Theme 6: Training students to detect bias, hallucinations, and errors in AI outputs and to critique before accepting.

Theme 7: Developing a new AI interaction paradigm — instructional and guiding rather than answer-dispensing.

Themes 1, 2, and 7 informed StudyBuddy AI. Themes 3, 4, 5, and 6 informed ScholarSync. Themes 1 and 3 together informed AI BrainQuest.

SOLUTION APPROACHES (MOST IMPACTFUL, MOST FEASIBLE, AND WILDEST) AND HOW THEY ARE DIFFERENT FROM EXISTING SOLUTIONS

1. Most Impactful Idea: **StudyBuddy AI**

StudyBuddy AI is a proposed AI tutoring tool integrated directly into learning management systems so it can receive assignment context and instructions automatically. Unlike ChatGPT Study mode, which students can disable with a click of one button, StudyBuddy Ai would be the default AI tool available within the school's assignment platform. This guides students through stumbling blocks using Socratic questioning (similar to the study mode aforementioned) without producing complete answers and would additionally function as an executive functioning coach. It would help students build study schedules, manage deadlines, and decompress workload.

Key differences: Assignment context-awareness (more personalized); removal of "opt-out problem"; integration of stress management features; a competitive advantage of external Ai tools due to better assignment alignment.

2. Most Feasible Idea: **ScholarSync Educational Approach**

ScholarSync is a structured classroom protocol. Here, students would complete initial assignments outside class using AI, then bring their work to class for two activities. First, there would be a teacher-facilitated discussion of how they used AI, with feedback on more effective approaches. Second, and open, technology-free discussion that transforms that assignment into project-based learning through human collaboration. This directly operationalizes the augmentation model. Here, AI is used for productivity and initial thinking, while class time is devoted to human-centered evaluative and creative skills AI cannot replicate.

Key Differences: Unlike in-class teacher recommendations (Solution 3 in our benchmarking), ScholarSync does not require teachers to be AI experts, rather it requires them to be good discussion facilitators. This is a skill most already have, so it is feasible without developing new technology investment and therefore making it more scalable.

3. Wildest Idea: **AI BrainQuest**

AI Brainquest would place students inside an augmented/virtual reality simulation showing the neurological effects of different AI use platforms. Rather than telling students about AI's effects on the brain, BrainQuest would let them experience them through a first-person simulation. It would integrate medical, social development, and critical thinking impact models. This directly would address one of our themes about incentive creation, as seeing your own brain cells atrophy is a more viscerally motivating incentive than abstract warnings would be.

Limitation: BrainQuest would require substantial AR/VR development, and also would cost a lot of money to implement. This development and investment is not feasible right now, which is why it is included in the "wildest idea" section.

ALTERNATIVES MATRIX - OUR SOLUTIONS	**Possibility to implement/Feasibility**	**Incentives for students**	**How well it meets the need of improving their interaction with AI**	**Invasiveness/Potential to take things too far and automate entirely rather than augment -- overly pushing people (infringing upon people's critical thinking)**	**How well it can encourage students (Can it turn AI into an actual overall beneficial tool for students on future development)**	**How well it educates people about AI's strengths *and* limitations**	**Ease of use & Accessibility**
Solution 1: StudyBuddy AI	3	3	3	3	3	1	3
Solution 2: ScholarSync	3	2	3	2	3	3	2
Solution 3: AI BrainQuest	1	3	3	3	3	3	2

Figure 12: Alternatives matrix evaluating our own solutions

STRENGTHS AND LIMITATIONS

Our project's main strength is its systemic perspective. Our usage of CLDs rather than single-variable analysis allows us to identify root causes that are invisible to conventional studies. Our findings align with and extend recent empirical work (Lee et al.,2025; Pew Research, 2025), and our methodology adds a new tool to the educational AI research toolkit and equips teachers with the best tools to help their students learn.

We also benefit from proximity to the problem. As high-school students ourselves, our access to the student participants and our insider understanding of academic pressure dynamics allowed us to surface findings that an external researcher might have missed.

Our most significant limitation is sample size and diversity. We interviewed a small number of participants, all people connected through personal networks. This could create some bias and they may not be representative of the broader population. Future research should probably expand to a more diverse sample.

We also lacked quantitative validation. Our CLD identifies feedback structures and leverage points, but it does not provide empirical estimates of effect sizes. All of the hypotheses we propose should be tested through controlled studies before solutions are implemented at scale.

LESSONS LEARNED AND NEXT STEPS

Our most feasible solution is ScholarSync AI, a controlled method of AI usage in classrooms to help students learn more effectively both in and out of the classroom. The concept of ScholarSync AI was largely inspired by our labelled cluster from Section 7, labelled "[Creating] new rules and regulations to give teachers more control over students' AI usage". We then combined this theme with our understanding of students' tendencies to use and be reliant on AI to develop this controlled approach to AI regulation.

When we benchmarked this solution against standards such as feasibility, incentive, and effect, this was the most consistently well-performing approach. In particular, per its label, we emphasized its accessibility/feasibility: both of the other solutions require a refined individual product, but ScholarSync can be used simply as a classroom integration, which is much simpler.

To physically develop ScholarSync, we would need to develop a software, say in the form of a browser extension, that monitors students' AI use and compiles their behavior in the form of a report. In addition, ScholarSync should set some regulations on students' AI use, e.g. blocking AI answers that give an overly direct solution and prompting the user to work step-by-step. Testing can be done by testing the extension on various AI algorithms, such as ChatGPT, Claude, Gemini, etc., and finding how effective ScholarSync is at facilitating proper AI conversations that also stimulate critical thinking.

To implement this idea, we could reach out to schools or school districts and propose the solution to them. The first few steps of implementation could be slower due to lack of experience and credibility, but it will become more consistent thereafter.

REFERENCES

Acemoglu, Daron, and Pascual Restrepo. "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment." *American Economic Review*, vol. 108, no. 6, 2018, pp. 1488–1542.

Autor, David H. "Why Are There Still So Many Jobs? The History and Future of Workplace Automation." *Journal of Economic Perspectives*, vol. 29, no. 3, 2015, pp. 3–30.

Braun, Virginia, and Victoria Clarke. "Using Thematic Analysis in Psychology." *Qualitative Research in Psychology*, vol. 3, no. 2, 2006, pp. 77–101.

Caliman, Lucille. "Generative AI: The Risk of Cognitive Atrophy." *Polytechnique Insights*, 4 July 2025,

www.polytechnique-insights.com/en/columns/neuroscience/generative-ai-the-risk-of-cognitive-atrophy.

Holmes, Wayne, et al. *Artificial Intelligence in Education*. UNESCO, 2022.

"Introducing Study Mode." *OpenAI*, 29 July 2025, openai.com/index/chatgpt-study-mode/.

Lee, Hao-Ping (Hank), 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." *Microsoft Research*, 19 June 2025, www.microsoft.com/en-us/research/publication/the-impact-of-generative-ai-on-critical-thinking.

Lipman, Joanne, and Rebecca Distler. "Schools Shouldn't Ban Access to ChatGPT." *TIME*, 11 Jan. 2023, time.com/6246574/schools-shouldnt-ban-access-to-chatgpt/.

Luckin, Rose. *Machine Learning and Human Intelligence: The Future of Education for the 21st Century*. UCL Institute of Education Press, 2017.

Parasuraman, Raja, and Dietrich Manzey. "Complacency and Bias in Human Use of Automation: An Attentional Integration." *Human Factors*, vol. 52, no. 3, 2010, pp. 381–410.

Sidoti, Olivia. "About a Quarter of U.S. Teens Have Used ChatGPT for Schoolwork." *Pew Research Center*, 15 Jan. 2025, www.pewresearch.org/short-reads/2025/01/15/about-a-quarter-of-us-teens-have-used-chatgpt-for-schoolwork.

Sidoti, Olivia. "About a Quarter of U.S. Teens Have Used ChatGPT for Schoolwork." *Pew Research Center,* 15 Jan. 2025, www.pewresearch.org/short-reads/2025/01/15/about-a-qrter-of-us-teens-have-used-chatgpt-for-schoolwork

Sterman, John D. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Irwin/McGraw-Hill, 2000.

Sweller, John. "Cognitive Load during Problem Solving: Effects on Learning." *Cognitive Science*, vol. 12, no. 2, 1988, pp. 257–285.

Yan, Lixiang, et al. "How Students Use Generative AI: Patterns and Implications for Assessment." *Computers & Education*, vol. 208, 2024.