From Answer Machines to Ignorant Co-Learners: Designing AI to Augment Rather than Replace Human Thinking

Manni Cheung
man.cheung.24@ucl.ac.uk
UCL Interaction Centre
London, UK

ABSTRACT

Generative AI (GenAI) is increasingly used in education and knowledge work, often framed as an authoritative "answer machine" delivering seamless responses. This paradigm fosters passive reliance and limits critical reflection. In response, we propose the concept of "artificial ignorance," AI designed not to instruct, but to prompt inquiry, ambiguity, and epistemic friction. Drawing on feminist theory, critical pedagogy, and humanistic HCI, we argue that ignorance should not be seen as a deficit but as a generative stance that resists intuitive, frictionless design. We introduce design principles that position users as reflective co-creators rather than passive recipients. Through speculative interventions, such as meta-prompts and conflicting outputs, artificial ignorance systems aim to slow down interactions and invite users to question both AI and self. Our contribution is a reframing of AI literacy and educational AI design that resists solutionism, advocates for pluralism, and opens space for critical, collaborative sense-making between humans and machines.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; **HCI theory, concepts and models**; • **Social and professional topics** → *Computing literacy*; • **Computing methodologies** → *Artificial intelligence*.

KEYWORDS

AI Literacy, Counter-Intuitive AI, Critical Thinking, Epistemic Friction, Generative AI, Human-AI Interaction, Humanistic HCI

ACM Reference Format:

Manni Cheung. 2025. From Answer Machines to Ignorant Co-Learners: Designing AI to Augment Rather than Replace Human Thinking. In *CHI '25: Workshop on Tools for Thought, April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 4 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION: MOVING BEYOND THE "ANSWER MACHINE" PARADIGM

Artificial intelligence (AI) in education is increasingly framed as an "answer machine," a seamless, authoritative system delivering instant, polished responses. This intuitive model privileges smooth interaction and convenience over critical engagement, encouraging overreliance and obscuring the sociopolitical structures embedded in AI outputs [3, 9]. By positioning AI as a source of final answers, this paradigm naturalizes assumptions about what counts as legitimate knowledge—and who gets to define it [1, 7].

Emerging critiques describe this trend as the "automation of common sense" [9], which reinforces dominant ideologies and limits epistemic diversity. Rather than fostering intellectual agency, intuitive AI streamlines decision-making in ways that discourage deeper reflection, particularly in educational and knowledge-building contexts. As critical HCI and educational researchers have pointed out, such systems are less about enhancing thought than replacing it [5, 12].

This paper proposes a counter-framing: artificial ignorance. Instead of designing AI to simulate intelligence or expertise, we argue for systems that prompt reflection, ambiguity, and inquiry. Drawing from feminist epistemologies and critical pedagogy, we reconceive ignorance not as deficiency, but as a generative stance—one that embraces partiality, resists premature closure, and makes space for interpretive labor [1, 6, 7, 14].

We are inspired in particular by Jacques Rancière's figure of the "ignorant schoolmaster," who disrupts hierarchical teaching by refusing to explicate knowledge. In this model, teaching becomes a shared investigation rather than a transmission of answers. Translating this pedagogical ethic into AI design, we envision systems that relinquish claims to authority and instead foster epistemic friction—moments of uncertainty, dissonance, or pause that compel users to think critically and reflexively [10].

Throughout this paper, we explore how artificial ignorance can be realized through specific design strategies: presenting conflicting outputs, surfacing system limitations, and embedding meta-prompts that ask users to articulate assumptions or question dominant framings. These interventions invite users to act as co-creators of meaning rather than passive consumers of algorithmic knowledge.

Our goal is not to perfect AI's explanatory power but to reconfigure its pedagogical role. Rather than automating common sense, artificial ignorance invites uncommon questioning. It positions AI not as an omniscient tutor but as a deliberately limited, open-ended co-learner, one that enables deeper, more pluralistic engagement with the complexities of knowing and not-knowing.

2 THEORETICAL FOUNDATIONS

To design AI that supports reflection rather than passive consumption, we build on critical traditions that reconceptualize ignorance as a generative force. Across epistemology of ignorance, feminist

CHI '25, April 26–May 01, 2025, Yokohama, Japan

ACM ISBN 978-1-4503-XXXX-X/2018/06
https://doi.org/XXXXXXX.XXXXXXX

theory, critical pedagogy, and humanistic HCI, we find a shared call to resist the illusion of complete knowledge. These perspectives help reframe educational AI from a tool of seamless explanation to one that invites friction, interpretation, and co-learning.

2.1 Ignorance as a Design Ethic

Ignorance is commonly treated as a gap to be filled by information. In contrast, scholars such as Nancy Tuana and José Medina argue that ignorance can also be actively constructed, maintained, or resisted [7, 14]. It is not merely absence, but a political condition, shaped by what is silenced, overlooked, or deemed irrelevant. In AI design, this typically manifests as user ignorance being positioned as the problem to be solved, while the system's authority remains unquestioned.

This framing is particularly evident in "answer machine" models that treat AI outputs as complete and neutral. The user is cast as epistemically lacking, while AI appears omniscient. Medina describes how this logic suppresses "resistant ignorance," the decision not to accept received knowledge uncritically, but to question its authority and omissions [7]. Designing for resistant ignorance, then, is not about concealing truth but about exposing complexity and prompting reflective doubt.

Tuana's concept of agnotology—the study of ignorance—underscores how omissions in data or uncertainty in claims are not accidental, but structured [14]. When AI systems obscure their limitations, they foreclose opportunities for users to engage critically. Reframing ignorance as productive opens a space for AI to reveal its blind spots rather than conceal them, inviting the user to become an investigator rather than a recipient.

2.2 Challenging Hierarchies through the "Ignorant Co-Learner"

Jacques Rancière's figure of the "ignorant schoolmaster" challenges the idea that teaching depends on expert explanation [10]. Instead, the teacher withholds knowledge and affirms the learner's capacity to think independently. This pedagogical inversion resonates strongly with educational AI design. Most systems are built on the assumption that AI must provide "correct" answers to fill user gaps. In contrast, a system that refrains from definitive instruction can prompt users to evaluate, compare, and construct meaning for themselves.

This shift is not about feigning ignorance or making systems deliberately unhelpful. Rather, it is about flattening epistemic hierarchies by designing AI to step back, acknowledging its limitations, surfacing contradictions, and deferring to human judgment. Such systems would prompt rather than resolve, offering multiple pathways rather than single conclusions. Users remain cognitively and ethically "in the loop," sustaining responsibility over how knowledge is interpreted and applied.

By disrupting expectations of mastery, artificial ignorance counters AI overreliance. Cognitive forcing—designing moments that slow users down to prompt reflection—can be embedded into AI interfaces to encourage this shift [3]. Whether through conflicting responses or overt uncertainty, the goal is not confusion but invitation: to make the user think, not just accept.

2.3 Friction, Situatedness, and the Feminist Critique

Feminist epistemologists such as Donna Haraway and Sara Ahmed have long challenged the myth of universal, context-free knowledge [1, 6]. Haraway's "situated knowledges" reminds us that all claims are partial, emerging from specific perspectives [6]. For AI to acknowledge its situatedness is to resist the fiction of objectivity, and, instead, expose the values and limitations embedded in its training and logic.

Ahmed extends this thinking to discomfort. Disorientation, she argues, can be politically generative [1]. Friction, whether in interface design, dialogue, or experience, can spark reflection by resisting smooth navigation. In AI, this might mean highlighting contradictions between sources, presenting multiple readings, or prompting users to ask what is missing. Rather than treating friction as a design flaw, it becomes a pedagogical tool.

Designing for epistemic friction aligns with feminist commitments to plurality, critique, and interruption. "Safe" AI interactions often reproduce dominant assumptions by smoothing over conflict. Artificial ignorance, in contrast, creates space for difference and complexity to surface—even when this feels less efficient or comfortable.

2.4 Humanistic HCI and Anti-Solutionism

Humanistic HCI similarly resists instrumental views of technology, emphasizing interpretation, ambiguity, and cultural entanglement [2, 4, 13]. In educational contexts, humanistic HCI provides a counterpoint to solutionist approaches that treat learning as a problem of delivery speed or content accuracy.

Evgeny Morozov critiques this "solutionism" as a reductive mindset that seeks technical fixes for deeply social challenges [8]. In educational AI, this often appears as systems designed to instantly resolve learning gaps, undermining the slow, recursive nature of genuine understanding. Artificial ignorance pushes back against this tendency by embedding constructive tension into interactions. Rather than optimizing for correctness, it makes space for ambiguity, doubt, and ongoing inquiry.

In practical terms, this could involve presenting divergent viewpoints, withholding full explanations, or asking users to interpret conflicting evidence. These strategies activate metacognitive processes—thinking about thinking—that are essential to critical education [5, 11]. The goal is not to frustrate but to decelerate, giving users room to reflect on how they know what they know.

Importantly, these interventions are not neutral. Designing for friction and partiality is a political act, one that redistributes interpretive authority and invites users into epistemic responsibility. As both Haraway and Ahmed suggest, discomfort can be ethically necessary [1, 6]. Artificial ignorance builds this into its logic, treating learning not as resolution, but as an unfolding, participatory process.

3 TOWARD ARTIFICIAL IGNORANCE IN PRACTICE

Translating artificial ignorance into practice requires design strategies that embed reflection, uncertainty, and pluralism into the structure of AI interaction. These interventions aim not to withhold

assistance arbitrarily, but to create deliberate friction that prompts users to pause, question, and co-construct meaning.

3.1 Designing for Deliberate Friction and Inquiry

A first strategy involves introducing epistemic tension through ambiguity or partial responses. Rather than providing conclusive outputs, the system might offer multiple, conflicting perspectives or highlight areas of uncertainty. For instance, an AI summarizing a news article could present contrasting interpretations and ask the user which aligns best with available evidence, or suggest what data might be missing. These designed gaps turn the user into an active sense-maker, reducing the likelihood of blind deference to algorithmic authority.

Complementing this are meta-prompts that foreground the interpretive process. After an AI-generated recommendation, the system might ask: "What assumptions underlie this suggestion?" or "What alternative explanations could apply?" These prompts do not add content, but structure attention, encouraging reflection-in-action. This design echoes Paulo Freire's "problem-posing education," which values inquiry over explanation, and aligns with José Medina's call for epistemic humility [5, 7].

To support this, interfaces can include multi-modal uncertainty displays, i.e., visual elements that represent ambiguity not as error, but as context. Rather than a confidence score, users might see a spectrum of interpretations, a branching logic, or highlighted areas where training data was sparse or contested. This makes the AI's partiality visible, reinforcing the idea that knowledge is always situated and subject to interpretation.

3.2 Educational Integration and Institutional Considerations

Realizing artificial ignorance also demands alignment with educational environments. Systems that embed friction may initially appear inefficient or unintuitive, particularly in settings driven by standardized testing or productivity metrics. To be viable, these tools must be clearly linked to pedagogical goals such as critical thinking, metacognition, and epistemic agency.

Educators play a central role in this alignment. Teacher-facing versions of artificial ignorance tools can provide scaffolds for classroom use—suggesting discussion questions, writing prompts, or reflection tasks that respond to the system's ambiguity. Professional development can support teachers in interpreting contradictory outputs not as errors but as invitations for deeper conversation.

Just as importantly, friction must be adjustable. Not all learners will benefit equally from the same degree or kind of uncertainty. Systems might offer modes or toggles that modulate the level of ambiguity, giving educators and students control over how much friction to introduce—and when.

Finally, artificial ignorance benefits from being part of a broader ecosystem. Complementary resources, such as ethical guidelines, open data practices, interoperable platforms, can reinforce its values. Design frameworks might prioritize transparency over polish, explicitly surface system limitations, and support diverse forms of sense-making. This scaffolding helps position friction not as failure, but as a core value of AI design in education.

4 DISCUSSION AND FUTURE WORK

4.1 Rethinking AI in Education: From Intelligence to Ignorance

The dominant model of educational AI imagines systems that dispense correct answers efficiently and authoritatively. Artificial ignorance offers a conceptual shift: away from automation of knowledge and toward the cultivation of inquiry. By reframing ignorance as a generative design principle, rather than a deficiency to correct—this approach resists the logic of frictionless learning, inviting deeper forms of reflection, interpretation, and agency.

Rooted in Rancière's principle of the ignorant schoolmaster, feminist theories of situated knowledge, and Ahmed's politics of disorientation, artificial ignorance aims to slow down interpretation and expose the partiality of both human and machine understanding [1, 6, 10]. Rather than resolving ambiguity, these systems hold it open—creating space for users to grapple with complexity, contradiction, and the provisional nature of knowledge.

This reframing has implications beyond interface design. It calls for a pedagogical and ethical shift: educators become facilitators of epistemic engagement, and users become co-participants in meaning-making. Success is not measured by fluency or satisfaction but by whether the system supports intellectual struggle and critical growth.

4.2 Empirical Validation and Design Research

While this paper develops the conceptual foundations of artificial ignorance, future work must examine its effects in practice. Key questions include: Does deliberate friction increase critical thinking or metacognitive awareness? How do users respond emotionally to AI that foregrounds uncertainty? Can epistemic humility be taught—or designed?

Empirical studies could compare friction-based systems with conventional AI in educational settings, assessing their impact on student engagement, reasoning quality, and learning outcomes. Longitudinal work might explore how user attitudes evolve over time, especially across diverse learner populations and disciplinary contexts. Researchers should also consider potential drawbacks—such as user frustration, accessibility barriers, or cognitive overload—and how these might be mitigated through scaffolding or personalization.

One promising context for early evaluation is undergraduate seminar courses focused on media literacy, ethics, or writing—domains where students already engage with GenAI tools for brainstorming or composition. A between-subjects design could compare conventional "answer-oriented" interfaces with a friction-rich "ignorant co-learner" prototype that integrates meta-prompts and uncertainty visualizations. Outcome measures might include argument diversity, source evaluation skill, perceived cognitive effort, and qualitative feedback on trust, engagement, and perceived value. Such a study would offer early insight into whether epistemic friction enhances reflective learning or introduces new burdens—and under what conditions.

4.3 Extending to Other Domains

The concept of artificial ignorance may also apply beyond education, particularly in domains where AI is used to inform decisions: civic platforms, journalism, policy development, or design. In these spaces, overreliance on machine-generated conclusions can reinforce biases or obscure contestable issues. Introducing ambiguity, alternatives, or dialogic prompts could help sustain epistemic plurality and invite more democratic forms of participation.

Tailoring friction to different domains, e.g., legal reasoning vs. creative collaboration, requires attention to context, risk, and user expectation. Future research might explore how domain-specific norms shape what kinds of ignorance are acceptable, generative, or threatening.

4.4 Policy, Ethics, and Public Imagination

Finally, a broader uptake of artificial ignorance will require cultural and institutional support. Commercial systems are typically optimized for clarity and confidence. Making room for ambiguity may involve redefining metrics of success, incentivizing slower interactions, and reimagining what we want from "intelligent" systems.

Policy frameworks and ethics guidelines could explicitly endorse reflection-oriented design, particularly in education and public-interest technology. Initiatives that promote algorithmic transparency and open-ended learning, rather than seamless automation, will be key to making artificial ignorance a viable and visible design direction.

4.5 Toward Frictional AI Design: Summary Principles

To support the development of artificial ignorance in practice, we summarize several design principles that synthesize the preceding discussion. These are not prescriptive heuristics but provocations—intended to guide experimentation, curriculum integration, and platform-level shifts toward epistemically responsible AI.

- Epistemic Friction by Design: Prioritize ambiguity, delay, or conflict over seamlessness in AI feedback.
- Meta-Cognition Prompts: Use prompts that explicitly ask learners to reflect on underlying assumptions or alternative interpretations.
- Situated Partiality: Make system limitations visible through uncertainty visualizations and data transparency.
- Adjustable Friction: Allow teachers and learners to modulate the level of ambiguity based on context and goals.
- Plurality over Precision: Present divergent perspectives or reasoning paths rather than one "correct" output.

5 CONCLUSION

This paper has introduced artificial ignorance as a critical and constructive alternative to dominant AI paradigms in education. In contrast to systems that prioritize speed, certainty, and seamless delivery, we argue for designs that foreground ambiguity, friction, and partiality. Drawing from feminist epistemology, ignorance studies, critical pedagogy, and humanistic HCI, we have shown how ignorance can be reframed as a generative stance—one that fosters epistemic humility and sustained interpretive engagement.

Rather than correcting user deficits or optimizing for cognitive ease, artificial ignorance supports users in questioning, comparing, and co-constructing meaning. We outlined concrete strategies—meta-prompts, conflicting outputs, and uncertainty visualizations—that instantiate this approach and reposition AI not as an omniscient tutor but as a partial, reflexive co-learner.

We also emphasized the importance of broader institutional scaffolding, from educator agency and curricular alignment to ethics guidelines and platform governance. The summary design principles in Section 4.5 provide a practical synthesis: they reimagine educational AI as frictional, adjustable, dialogic, and ethically accountable.

Ultimately, artificial ignorance is not a bug to be fixed but a feature to be embraced. It shifts the goal of AI in education from answer provision to inquiry provocation—from automation to co-thinking. We invite HCI researchers, educators, designers, and policymakers to take up this framing, adapt it across contexts, and explore how not-knowing can catalyze deeper learning and more just forms of knowledge engagement.

ACKNOWLEDGMENTS

To Chaeyeon Lim, Margo Spiyrdonava, Shih-Han (May) Wang, and Umut Yucharoglu—the *Aye-Eye* team from UCL Interaction Centre—thank you for your invaluable contributions to the development of Aye-Eye and for shaping the ideas explored in this work.

REFERENCES

[1] Sara Ahmed. 2010. *The Promise of Happiness.* Duke University Press.
[2] Jeffrey Bardzell and Shaowen Bardzell. 2015. Humanistic HCI. *Interactions* 22, 4 (2015), 20–29. https://doi.org/10.1145/2888576
[3] Zana Buçinca, Maja Barbara Malaya, and Krzysztof Z. Gajos. 2021. To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-assisted Decision-making. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 188 (April 2021), 21 pages. https://doi.org/10.1145/3449287
[4] Paul Dourish. 2004. *Where the Action Is: The Foundations of Embodied Interaction.* MIT Press.
[5] Paulo Freire. 1970. *Pedagogy of the Oppressed.* Continuum.
[6] Donna Haraway. 1988. Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies* 14, 3 (1988), 575–599.
[7] José Medina. 2013. *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and Resistant Imaginations.* Oxford University Press.
[8] Evgeny Morozov. 2013. *To Save Everything, Click Here: The Folly of Technological Solutionism.* PublicAffairs.
[9] Carolyn Pedwell. 2024. The Intuitive and the Counter-Intuitive: AI and the A ffective Ideologies of Common Sense. *New Formations* 112, 112 (2024), 70–93. https://doi.org/doi:10.3898/NewF:112.04.2024
[10] Jacques Rancière. 1991. *The Ignorant Schoolmaster: Five Lessons in Intellectual Emancipation.* Stanford University Press.
[11] Donald A. Schön. 1983. *The Reflective Practitioner: How Professionals Think in Action.* Basic Books.
[12] Neil Selwyn. 2022. The future of AI and education: Some cautionary notes. *European Journal of Education* 57, 4 (2022), 620–631. https://doi.org/10.1111/ejed.12532 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/ejed.12532
[13] Phoebe Sengers, Kirsten Boehner, Shay David, and Joseph "Jofish" Kaye. 2006. Reflective Design. In *Proceedings of the 4th conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques.* ACM, 49–58. https://doi.org/10.1145/1142405.1142422
[14] Nancy Tuana. 2006. The Speculum of Ignorance: The Women's Health Movement and Epistemologies of Ignorance. *Hypatia* 21, 3 (2006), 1–19. http://www.jstor.org/stable/3810948

Received 20 February 2025; revised 11 April 2025