

Stop Writing for Me: Generative Refusal in AI Tools for Thought

Sora Kang
sorakang@snu.ac.kr
Seoul National University
Seoul, Republic of Korea

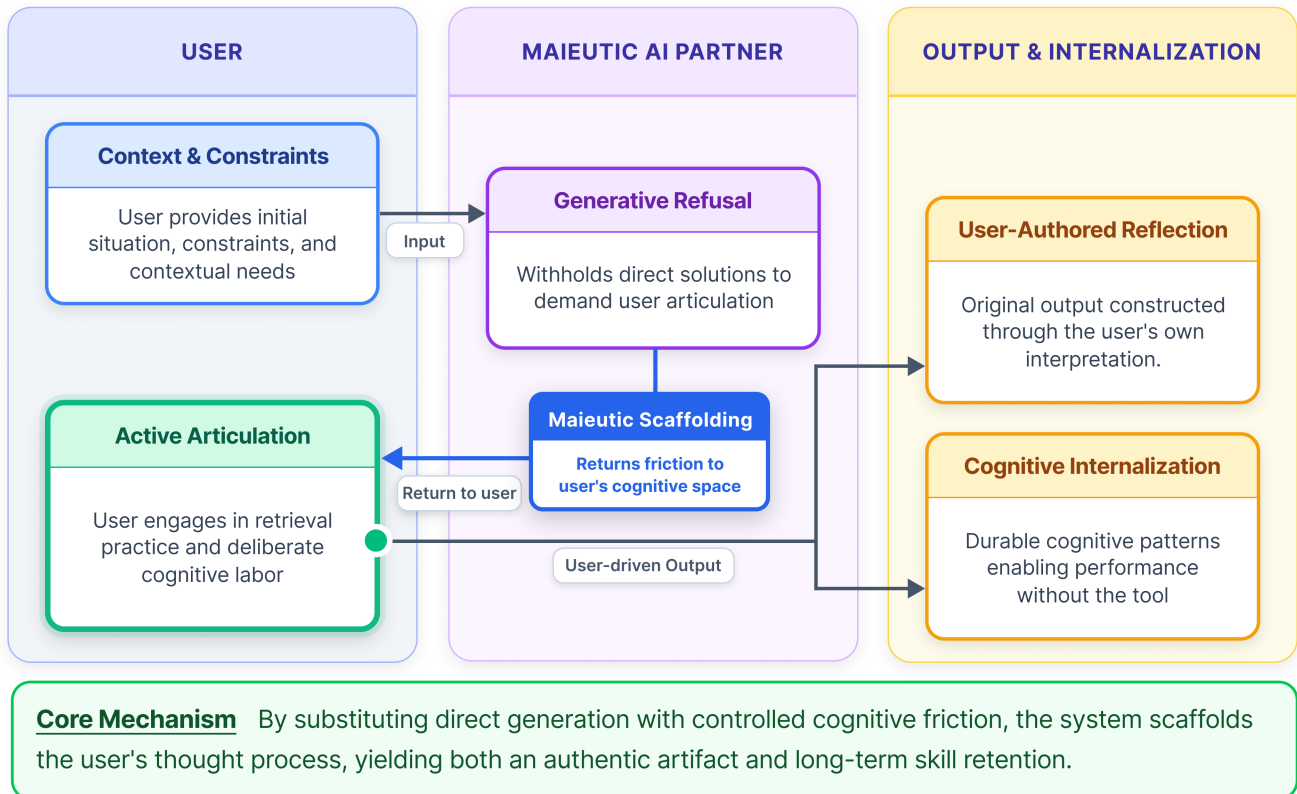


Figure 1: The Maieutic Interaction Framework. Instead of bypassing cognition, the system uses Generative Refusal to return cognitive friction to the user, shifting the interaction from delegation to active articulation.

Abstract

In creative domains where the labor of articulation is central to the craft, how should we design Tools for Thought that enhance rather than bypass human cognition? Current GenAI paradigms often prioritize "cognitive offloading"—writing on behalf of users—risking the erosion of the constructive thought process essential to artistic training. In this position paper, we explore AI as a Maieutic Partner through "Generative Refusal"—strategically withholding text generation to demand user articulation. We discuss Actor's Note, a journaling tool that generates context-aware questions instead of draft text. Our field study suggests that this constraint significantly reduced cognitive burden while fostering a residual effect of internalized questioning habits. We use these findings to discuss broader design implications for protecting human cognition against the tendency of generative efficiency.

CCS Concepts

• **Human-centered computing** → **Human computer interaction (HCI)**; • **Applied computing** → *Performing arts*.

Keywords

Human-AI Interaction, Tool for Thoughts, Theater, Journaling, Creativity Support Tools, Large Language Models, Co-creativity

ACM Reference Format:

Sora Kang. 2026. Stop Writing for Me: Generative Refusal in AI Tools for Thought. In *Proceedings of CHI 2026 Workshop on Tools for Thought*. ACM, New York, NY, USA, 4 pages.

1 Introduction

The concept of Tools for Thought (TfT) aims to extend human cognitive capabilities [12, 14, 31]. In the domain of actor training,

character journaling—a pedagogical practice where actors write first-person narratives from their character’s perspective to internalize inner motivations and emotional histories—serves as a foundational TfT for internalizing roles [11, 16, 19, 22, 29]. However, despite its established benefits, this practice is notoriously difficult to sustain [9, 25, 32]. Actors frequently abandon journaling not due to a lack of material, but because the high cognitive load required to initiate reflection after exhaustive rehearsals creates a significant barrier [6, 30], often manifesting as the “blank page” problem [23, 26].

While Generative AI offers a potential solution to this sustainability gap [7, 10, 17, 18], current design paradigms prioritize efficiency through cognitive offloading—generating text on the user’s behalf [13, 15]. Although this approach rapidly produces artifacts, it effectively bypasses the cognitive process of articulation. In domains like artistic training, where the mental effort of interpretation is central to learning [21, 27, 28], such “write-for-me” features could deprive users of the opportunity for cognitive appropriation and internalization.

In this position paper, we argue that for GenAI to serve as a true TfT, it must shift from a “Co-author” role to that of a “Maieutic Partner” that facilitates inquiry. This perspective resonates with recent HCI literature exploring “Socratic agents” and cognitive scaffolding [3]. Recent studies demonstrate that intentionally employing a “refusal to generate” strategy—such as question-driven prewriting [5] or metacognitive prompting [1, 4]—effectively promotes critical thinking over cognitive offloading. Furthermore, Socratic AI has shown promise in mitigating evaluation anxiety in high-stakes training environments [2]. However, existing work has primarily applied this paradigm to objective academic tasks or structured prewriting.

To demonstrate how this Socratic approach can be extended into the arts—specifically, the subjective, embodied, and emotionally demanding domain of artistic training—we discuss Actor’s Note, a work to be presented at CHI 2026 [24]. This system implements a constraint-based design that strictly refuses to generate text, instead providing context-aware questions tailored to the actor’s rehearsal stage. By employing “good friction” [20], the system lowers the cognitive burden of initiation while augmenting the depth of thought. Based on empirical findings from this implementation, we aim to discuss broader design implications for future TfTs, each accompanied by a workshop discussion question: (1) employing Generative Refusal to demand user articulation rather than filling the gap with synthetic text; (2) implementing Adaptive Scaffolding that adjusts friction based on the user’s workflow phase; (3) leveraging AI as a Non-judgmental Space to mitigate social evaluation threats; and (4) adopting Internalization—performance without the tool—as a primary success metric.

2 Case Study: Actor’s Note

We developed Actor’s Note to operationalize the vision of Maieutic AI and validated it through a rigorous field deployment.

2.1 System Overview

Actor’s Note is a web-based journaling tool powered by a Large Language Model (GPT-4o). Its core design principle is the refusal to

generate text on behalf of the user. Instead, the system analyzes the user’s uploaded script, role, and performance date to generate three context-specific questions daily. Crucially, the system utilizes the actor’s current rehearsal stage as a prompt parameter—focusing, for example, on backstory details during table work versus emotional continuity during run-throughs. These questions are framed in the second person to elicit first-person reflection, ensuring the cognitive labor of interpretation remains with the user.

2.2 Study Overview

To evaluate the system, we conducted a 14-day in-the-wild study with 29 professional and student actors using a randomized crossover design. Following a baseline period to establish journaling habits, participants were counterbalanced across two conditions: an AI-assisted phase using Actor’s Note and an unassisted freewriting phase. We collected interaction logs to analyze writing latency and linguistic diversity, alongside daily surveys measuring cognitive burden, intrinsic motivation, and acting confidence. Semi-structured interviews were conducted post-study to examine qualitative shifts in the actors’ reflective processes.

2.3 Key Findings

The immediate effect of the system was a simultaneous reduction in entry barriers and an improvement in the quality of reflection. Despite maintaining the friction of writing, results showed a significant decrease in Cognitive Burden ($\eta_p^2 = .381, p < .001$) and an increase in Intrinsic Motivation ($\eta_p^2 = .201, p = .017$). Acting Confidence also increased significantly ($\eta_p^2 = .352, p < .001$), suggesting that the AI-generated questions concretized character interpretation. Beyond subjective metrics, log analysis confirmed a deepening of inquiry; entries written with AI assistance exhibited significantly higher Lexical Diversity ($q = .025$) and increased usage of both negative ($q = .001$) and positive ($q = .025$) emotion words compared to unassisted entries, indicating that the questions prompted actors to articulate a broader and more complex range of emotions.

Crucially, the tool’s impact varied by timing and fostered lasting cognitive habits. For the group introducing AI early, the tool functioned as a momentum starter that rapidly reduced cognitive burden, whereas for the group introducing AI later, it functioned as a deepener, leading to a significant increase in Narrative Transportation ($p = .0128$). This suggests friction plays a different role depending on the user’s workflow stage. Furthermore, post-study data indicates that the questioning framework was internalized; after AI assistance was removed, participants reported a high tendency to recall the AI’s questioning style or self-generate similar questions ($M = 4.87$ on a 7-point scale), demonstrating that the intervention fostered a lasting cognitive habit rather than a temporary dependency.

3 Discussion

The deployment of Actor’s Note demonstrates that constraining generative capabilities can, counter-intuitively, enhance user engagement and cognitive depth. We offer four implications for designing GenAI tools that aim to support, rather than replace, human thought processes.

3.1 Generative Refusal

The prevailing design logic in Generative AI prioritizes the reduction of friction, aiming to produce the final artifact with minimal user effort. However, our findings suggest that in educational and training contexts, this efficiency is detrimental to learning. By implementing a "Refusal to Generate," Actor's Note forced users to engage in retrieval practice and articulation—cognitive processes essential for memory consolidation and understanding. We argue that "Generative Refusal" should be treated not as a lack of functionality, but as a core mechanic for Tools for Thought. Designers should utilize GenAI to create a structured gap that demands user input, rather than filling that gap with synthetic text. This shifts the interaction model from "delegation" (AI does it for me) to "activation" (AI prompts me to do it).

This concept of Generative Refusal aligns closely with Bjork's notion of "desirable difficulties" [8], which posits that introducing controlled cognitive friction into the learning process may slow down immediate performance but significantly enhances long-term retention and the transfer of skills. By intentionally refusing to bypass the "blank page" for the user, Actor's Note introduces this desirable difficulty. Furthermore, the AI-generated questions serve as cognitive scaffolding within the user's Zone of Proximal Development (ZPD) [33], guiding the actors to construct their own interpretations rather than passively consuming synthetic text.

Workshop Question: *Where lies the "optimal friction point" that balances the immediate desire for productivity with the long-term necessity of cognitive depth, and how can we design interfaces that help users negotiate this trade-off?*

3.2 Adaptive Scaffolding

Our study revealed that the user's cognitive needs shift significantly as their work progresses. In the early stages of rehearsal, users struggled with inertia and the "blank page," requiring low-friction prompts to initiate writing. In later stages, however, users struggled with fixation on established interpretations, requiring high-friction, challenging questions to deepen their understanding. A static tool cannot address both needs effectively. Future TfT systems must be context-aware, adjusting their scaffolding strategy dynamically based on the user's workflow phase. Detecting this phase presents a key design challenge, and we envision several potential approaches to explore. For example, a system might employ (1) **System-driven inference**, estimating progress based on script metadata or temporal proximity to the performance date; alternatively, it could rely on (2) **User-driven selection**, allowing actors to self-report their current rehearsal stage (e.g., 'table work' vs. 'run-through') prior to journaling; or even (3) **Leader-driven synchronization**, where a director manually updates the global phase for the entire cast. Depending on the detected phase, the system could then act as a pacemaker in the beginning to build momentum, and as a challenger or devil's advocate in later stages to prevent cognitive stagnation.

Workshop Question: *How might this "Maieutic" form of questioning be operationalized in other creative domains like coding or academic writing—where might "refusal" look like a syntax error, and where might it look like a helpful prompt?*

3.3 Mitigating Social Evaluation Threat

Deep reflection requires vulnerability, yet human feedback loops often introduce a social evaluation threat that leads to self-censorship. Our participants explicitly noted that they shared raw, unpolished thoughts with the AI that they would hide from human directors or peers to maintain their professional image. This indicates that AI's lack of social agency is a functional asset. Designers should explicitly frame AI-based reflection tools as "non-judgmental spaces" where the cost of error is zero. By removing the fear of social judgment, AI can facilitate a level of honest introspection and experimentation that is often inhibited in human-to-human training environments. However, we must also acknowledge an inherent tension here: the very absence of social stakes that makes AI a safe space might also reduce the extrinsic motivation, adrenaline, and performative energy typically derived from a human audience or director. Therefore, AI-based reflection tools should be positioned not as replacements for the rehearsal floor, but as preparatory sandboxes. Future research should explore how the private introspection facilitated by AI can be effectively bridged with the necessary, high-stakes environment of live performance, ensuring these tools complement rather than isolate actors from human feedback loops.

Workshop Question: *Could the "Character Journaling" method—adopting a persona to bypass self-censorship—be generalized to other high-stakes professional contexts (e.g., executive coaching, therapy), and does the non-human nature of AI make it the ideal mediator for this specific technique?*

3.4 Internalization as the Success Metric

The current industry standard for evaluating GenAI focuses on performance with the tool—measuring how much faster or better a user can produce an output while using the AI. However, this metric fails to distinguish between genuine cognitive augmentation and dependency. Our study measured "performance without the tool" (residual effects) and found that users internalized the questioning patterns, applying them independently after the intervention. We propose that the HCI community must adopt "Internalization" as a primary success metric for TfTs. If a user cannot perform the task better after the tool is removed, the system has merely served as a prosthetic for production, not a scaffold for thought.

Workshop Question: *If "Post-Tool Performance" is the true measure of a TfT, what specific longitudinal metrics or "unassisted" evaluation protocols should the HCI community adopt to standardize the measurement of cognitive internalization?*

4 Conclusion

We presented Actor's Note to demonstrate that "Generative Refusal" is a viable design strategy for protecting cognitive engagement. To serve as a potential blueprint for future Tools for Thought, we propose four design shifts: (1) treating refusal as an active feature to demand user articulation, (2) implementing adaptive scaffolding that shifts from support to challenge based on workflow, (3) leveraging AI as a non-judgmental space to mitigate social evaluation threat, and (4) adopting internalization (post-tool performance) as a primary success metric. We hope the workshop discussion will encourage the HCI community to move beyond prosthetic AI toward systems that support independent human capabilities.

References

- [1] A. Author et al. 2024. Enhancing Critical Thinking in Education by means of a Socratic Chatbot. *arXiv preprint arXiv:2409.05511* (2024).
- [2] B. Author et al. 2024. Using GenAI for Socratic Questioning: An Approach to Higher-Order Thinking for Nursing Education. *PMC 12557457* (2024).
- [3] C. Author et al. 2025. Exploring the Potential of Metacognitive Support Agents for Human-AI Co-Creation. *arXiv preprint arXiv:2506.12879* (2025).
- [4] D. Author et al. 2025. Investigating the effects of an LLM-based Socratic conversational agent on students' academic performance and reflective thinking in higher education. *Computers & Education* (2025).
- [5] E. Author et al. 2026. When AI only asks: how question-driven dialogue shapes prewriting in the classroom. *Frontiers in Education* (2026).
- [6] J.S. Bascomb. 2019. *Performing Arts and Performance Anxiety*. Ph. D. Dissertation. Marshall University. Theses, Dissertations and Capstones. 1184.
- [7] N. Begus. 2024. Experimental Narratives: A Comparison of Human Crowdsourced Storytelling and AI Storytelling. *Humanities and Social Sciences Communications* 11, 1 (2024), 1392. doi:10.1057/s41599-024-03868-8
- [8] Robert A Bjork. 1994. Memory and metamemory considerations in the training of human beings. In *Metacognition: Knowing about knowing*, MIT Press, 185–205.
- [9] S.B. Blix. 2015. Professional Emotion Management as a Rehearsal Process. *Professions and Professionalism* 5, 2 (2015). doi:10.7577/pp.1322
- [10] B. Branch, P. Mirowski, and K.W. Mathewson. 2021. Collaborative Storytelling with Human Actors and AI Narrators. *arXiv preprint* (2021).
- [11] M. Bruder, L.M. Cohn, M. Olnek, N. Pollack, R. Previto, and S. Zigler. 2012. *A Practical Handbook for the Actor*. Knopf Doubleday Publishing Group.
- [12] Vannevar Bush. 1945. As We May Think. *The Atlantic Monthly* 176, 1 (1945), 101–108.
- [13] T. Chakrabarty, V. Padmakumar, F. Brahman, and S. Muresan. 2024. Creativity Support in the Age of Large Language Models: An Empirical Study Involving Emerging Writers. *arXiv preprint* (2024).
- [14] Andy Clark and David Chalmers. 1998. The Extended Mind. *Analysis* 58, 1 (1998), 7–19.
- [15] A. Coenen, L. Davis, D. Ippolito, E. Reif, and A. Yuan. 2021. Wordcraft: a Human-AI Collaborative Editor for Story Writing. *arXiv preprint* (2021).
- [16] T. Cremin, K. Gooch, L. Blakemore, E. Goff, and R. Macdonald. 2006. Connecting drama and writing: seizing the moment to write. *Research in Drama Education: The Journal of Applied Theatre and Performance* 11, 3 (2006), 273–291. doi:10.1080/13569780600900636
- [17] F. Dayo, A.A. Memon, and N. Dharejo. 2023. Scriptwriting in the Age of AI: Revolutionizing Storytelling with Artificial Intelligence. *Journal of Media & Communication* 4, 1 (2023), 24–38.
- [18] R. Dharaniya, J. Indumathi, and V. Kaliraj. 2023. A design of movie script generation based on natural language processing by optimized ensemble deep learning with heuristic algorithm. *Data & Knowledge Engineering* 146 (2023), 102150. doi:10.1016/j.datak.2023.102150
- [19] Drama-Based Pedagogy. n.d. Writing in Role. <https://dbp.theatredance.utexas.edu/teaching-strategies/writing-role> Accessed: 2025-08-20.
- [20] R. Fleck and G. Fitzpatrick. 2010. Reflecting on reflection: framing a design landscape. In *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction*. 216–223.
- [21] U. Hagen. 1991. *Challenge For The Actor*. Simon and Schuster.
- [22] M.R. Hancock. 1993. Character Journals: Initiating Involvement and Identification through Literature. *The Journal of Reading* 36 (1993).
- [23] C.K. Joyce. 2009. *The Blank Page: Effects of Constraint on Creativity*. Ph. D. Dissertation. University of California, Berkeley.
- [24] Sora Kang. 2026. Actor's Note: Examining the Role of AI-Generated Questions in Character Journaling for Actor Training. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI 26)* (Barcelona, Spain). ACM, New York, NY, USA. doi:10.1145/3772318.3790370
- [25] S. Norrthon and A. Schmidt. 2023. Knowledge Accumulation in Theatre Rehearsals: The Emergence of a Gesture as a Solution for Embodying a Certain Aesthetic Concept. *Human Studies* 46, 2 (2023), 337–369. doi:10.1007/s10746-022-09654-2
- [26] L.M. Schultz and M. Rose. 1985. Writer's Block: The Cognitive Dimension. *College Composition and Communication* 36, 4 (1985), 497. doi:10.2307/357873
- [27] K. Stanislavski. 2009. *An Actor's Work on a Role*. Routledge.
- [28] K.S. Stanislavskij. 1986. *An actor prepares*. Methuen.
- [29] T. Stanley and P. Strandberg-Long. 2022. *An Actor's Research: Investigating Choices for Practice and Performance*. Routledge.
- [30] K.L. Sørensen, N. Hald, R.M. Olsen, and K.B. Ørjasæter. 2024. "Always be relevant": a phenomenological study of the actor's workday. *Arts & Health* (2024), 1–14. doi:10.1080/17533015.2024.2420822
- [31] Lev Tankelevitch, Advait Sarkar, Sean Rintel, Richard Banks, Mina Lee, et al. 2025. Understanding, Protecting, and Augmenting Human Cognition with Generative AI: A Synthesis of the CHI 2025 Tools for Thought Workshop. *arXiv preprint arXiv:2508.21036* (2025). <https://arxiv.org/abs/2508.21036>
- [32] S.L. Taylor. 2016. *Actor training and emotions: Finding a balance*. Ph. D. Dissertation. Edith Cowan University.
- [33] Lev S Vygotsky. 1978. *Mind in society: The development of higher psychological processes*. Harvard University Press.