

From Abundant Generation to Scarce Attention: Attentional Bottlenecks and Cognitive Load in Visual Human-AI Co-Creation

Zhang Ziyang

School of Journalism and Communication, Communication University of Zhejiang

January 2026

Abstract

Visual generative AI creates an abundance of possible images, yet the human capacity to evaluate, select, and refine them remains limited. This paper examines attentional bottlenecks and cognitive load in visual human-AI co-creation, tracing the move from linear prompt interfaces to spatial canvas workspaces. It synthesizes research on visual saliency, choice architecture, articulation barriers, and human agency, then proposes a mixed-methods study of how interaction design shapes creative attention.

Keywords: *visual generative AI; attention funnel; cognitive load; interaction design; spatial canvas*

Literature Review

The Evolutionary Trajectory: From Deterministic Execution to Agentic Partnership

The interaction paradigm between humans and generative technology is undergoing a fundamental transition from AI as a passive "executor" to an active "co-creative partner" (Gómez et al., 2025; Haase & Pokutta, 2024). Traditional "executor models," exemplified by earlier iterations of productivity suites like Microsoft 365 Copilot, operated primarily through deterministic automation where AI fulfilled discrete, linear transactions similar to a search engine (Salma et al., 2025). However, the advent of Agentic AI—systems capable of a "Perceive-Plan-Act" loop—requires a shift toward synergistic collaboration where AI contributes interpretively throughout the creative lifecycle (Haase & Pokutta, 2024; Salma et al., 2025).

From a product management perspective, this evolution represents a move beyond simple prompt-output transactions toward a collaborative partnership where humans and AI operate as a cohesive system (Salma et al., 2025). Haase and Pokutta (2024) categorize this trajectory into four levels: the Digital Pen, the AI Task Specialist, the AI Assistant, and finally, the AI Co-Creator, which demonstrates autonomous creativity in producing novel outcomes. While Western scholarship emphasizes the efficiency of intent-based specification and the expansion of the "Latent Design Space" (Nielsen, 2025), a critical counter-narrative warns of a "mechanized convergence" (Tankelevitch et al., 2025). In this regime, creative labor is systematically reduced to "parameterized selection," potentially displacing the human creator from an active maker to a reactive adjudicator of algorithmic outputs (Salma et al., 2025; Tankelevitch et al., 2025).

The Attention Paradox and Computational Saliency

Visual Generative AI (GenAI) has precipitated a structural "Attention Paradox": while algorithmic generative capacity is near-infinite, human cognitive resources for filtering, processing, and decision-making remain strictly bound by physiological thresholds (Salma et al., 2025; Tkouat, 2024). This conflict is best modeled through the

"Attention Funnel," where information flow must pass through layers of Awareness, Selection, and Action (Tkiouat, 2024).

Product efficacy at the Awareness stage is increasingly governed by computational saliency. According to the Itti-Koch Saliency Model, visual attention capture is a feature decomposition process where inputs are integrated into a Saliency Map (S) based on Intensity (I), Color (C), and Orientation (O):

$$S = I + C + O$$

(Tkiouat, 2024). This logic reveals a foundational design bottleneck: if AI-generated content lacks clear "Visual Anchors," it is filtered out within the first second of interaction, leading to cognitive fragmentation (Bilic, 2025; Tkiouat, 2024). Consequently, the "abundance" of generation becomes a liability if the system fails to establish a visual hierarchy that respects human cognitive limits in the "micro-windows" of modern interaction (Bilic, 2025; Tkiouat, 2024).

Cognitive Load Matrix and the "Articulation Barrier"

The effectiveness of human-AI collaboration depends on the dynamic management of the Cognitive Load Theory (CLT) matrix, which distinguishes between intrinsic, extraneous, and germane loads (Gkintoni et al., 2025; Tkiouat, 2024). In visual GenAI scenarios, these loads manifest as follows:

Table 1. Cognitive Load Matrix in Visual GenAI Scenarios

Load Type	Product Manifestation	Design Optimization Goal
Intrinsic	Deconstructing complex styles or spatial relations.	Scaffolding: Progressively introducing complexity (Gkintoni et al., 2025).
Extraneous	Articulation Barrier: The strain of prompt engineering.	Hybrid UI: Using spatial tokens and direct manipulation (Shi et al., 2025).
Germane	Synthesizing AI variants into final decisions.	Feedback Loops: Providing verifiable data support (Salma et al., 2025).

The "Articulation Barrier" is arguably the most significant hurdle; users often possess a clear but "vague" visual vision and struggle to translate this intent into precise textual instructions (Subramonyam et al., 2024; Tahira et al., 2025). Nielsen (2025) characterizes the current obsession with "prompt engineering" as an "absurd" compensation for failed system usability. When creators are forced to behave like "command coders," their attention is diverted from high-level strategic orchestration to low-level syntax (Salma et al., 2025; Tahira et al., 2025).

Interaction Paradigms: From Linear Chat to Spatial Canvas

To resolve these bottlenecks, product design is shifting from the Linear Chat UI to the Spatial Canvas paradigm. Chat interfaces, used by tools like Midjourney, are criticized for being ephemeral, forcing users into sequential recall that causes visual memory decay (Bilic, 2025; Tkiouat, 2024). Users are jumping straight to outputs rather than building mastery through exploration (Bilic, 2025).

Actual Case: The Brickify System Next-generation tools like Gemini Canvas, Box Canvas, and the Brickify paradigm offer persistent visual workspaces (Bilic, 2025; Box, 2025; Gemini, 2025). Brickify demonstrates a visual-centric paradigm by reifying visual elements—style, subject, and color—into interactive "Design Tokens" (Shi et al., 2025). By allowing users to directly manipulate these tokens (dragging, resizing, linking), the system enables a "Text-to-Draft, Mouse-to-Final" workflow (Shi et al., 2025). This leverages the human capacity for parallel visual search, allowing side-by-side comparison of multiple variations without the cognitive cost of sequential recall (Bilic, 2025; Shi et al., 2025; Tkiouat, 2024). Similarly, the AIdeation system for concept designers uses flexible reference

searching and recombination to allow users to explore more creative branches within a single attentional window (Wang et al., 2025).

Choice Architecture and the "AI Sandwich" Workflow

Generative abundance frequently leads to "Choice Overload" and "Decision Paralysis" (Salma et al., 2025; Tkouat, 2024). Effective choice architecture seeks to move from "data volume" to "structured information" using strategies such as:

- Intelligent Clustering: Automatically grouping variants by semantic similarity (Nielsen, 2025; Tkouat, 2024).
- Tiered Disclosure: Presenting primary options in a main view while hiding technical variants in a secondary view to protect the user's "flow" (Nielsen, 2025).

This aligns with the "AI Sandwich" workflow pattern (Nielsen, 2025):

1. Top Slice (Human): Strategic spark, context, and constraints.
2. Filling (AI): Divergent volume, variations, and outlines.
3. Bottom Slice (Human): Convergent curation, polishing, and final "taste" check.

Targeted Friction: Preserving Critical Thinking

While product designers typically aim for "frictionless" experiences, uncritical adoption of AI outputs can lead to "Attentional Deskilling" (Salma et al., 2025; Stackpole, 2024). Humans often adopt a "cognitive miserly" strategy, blindly accepting AI outputs through fast "System 1" thinking (Wu et al., 2025).

Actual Case: MIT Sloan and Accenture "Speed Bumps" Experimental research by MIT Sloan and Accenture suggests that introducing "Targeted Friction" (beneficial speed bumps) can improve output accuracy (Stackpole, 2024). By using visual cues—such as orange highlighting for potentially untrue statements or purple for confirmed data—the system nudges users into deliberate "System 2" thinking (Stackpole, 2024). This moderate friction allows users to catch errors and omissions without significantly dragging on overall task efficiency (Stackpole, 2024).

Psychological Costs: Motivation and the "Shadow AI" Phenomenon

Despite immediate performance gains, human-AI collaboration introduces a "Psychological Deprivation Effect" (Wu et al., 2025). In studies involving 3,562 participants, Wu et al. (2025) found that transitioning from AI-assisted work to solo tasks led to a significant decrease in intrinsic motivation and an increase in feelings of boredom. While AI enhances productivity, it may take over the "fun" or creative parts of a task, reducing the fulfillments associated with critical analysis and craftsmanship (Wu et al., 2025).

Furthermore, a global phenomenon of "Shadow AI use" has emerged, where 69% of workers (70% among creatives) hide their use of AI to avoid stigma or the devaluation of their work (Nielsen, 2025). This secret usage creates an extraneous step in the workflow—manually "humanizing" or retyping AI-generated work to hide its origin—which further drains the professional's scarce attentional resources (Nielsen, 2025).

Scholarly Gaps and Future Research Directions

Despite the depth of current research, two major gaps persist:

- Attention as a Foreground Mechanism: Most models focus on interaction patterns or performance outcomes without formalizing attention as a structural constraint in multimodal visual tasks⁹¹³.

- Long-term Cognitive Impact: There is a lack of longitudinal studies investigating how prolonged AI collaboration reshapes human intentionality and independent reasoning abilities⁴⁶⁴⁷.

Existing research effectively identifies interaction pitfalls but lacks a comprehensive formalization of attention as a structural production constraint (Tkouat, 2024). There is a critical need to transition toward "Context-First" design (Tkouat, 2024). Future systems must evolve toward "Full-Duplex Communication," where AI acts as a proactive partner in shared cognitive spaces, building on fragments of ideas alongside humans (Tankelevitch et al., 2025). Longitudinal studies are required to determine if sustained AI use fundamentally alters the development of expert schemas and long-term critical thinking skills (Tankelevitch et al., 2025).

Research Proposal

Introduction and Background

The current technological landscape is defined by an "Attention Paradox": while algorithmic generative capacity has become near-infinite, the human cognitive resource for filtering and decision-making remains strictly bound by physiological thresholds (Tkouat, 2024). Visual Generative AI (GenAI) has transitioned from a passive automation tool to an agentic co-creative partner, capable of participating in a "Perceive-Plan-Act" loop (Baird & Maruping, 2021; Salma et al., 2025). However, a significant mismatch persists: most commercial tools still operate as linear "executors," which forces users into sequential, turn-based loops that fundamentally clash with the non-linear, iterative nature of human creativity (Salma et al., 2025).

This research focuses on identifying the specific "clogs" within the attention funnel that occur when users interact with these executor models. By analyzing cognitive load and the potential erosion of creative agency, the researcher aims to provide evidence-based insights into optimizing interaction designs—moving from "Linear Chat" to "Spatial Canvas" paradigms (Bilic, 2025; Shi et al., 2025).

Research Questions and Objectives

The primary objective of this study is to characterize the formation of attentional bottlenecks within the human-AI co-creative cycle and identify product-level interventions that mitigate cognitive overload.

- RQ1 (The Action Bottleneck): How does the "Articulation Barrier"—the strain of translating ambiguous visual intent into precise text prompts—function as an extraneous cognitive load that redirects attention from strategic creativity to "syntactic engineering" (Subramonyam et al., 2024; Tahira et al., 2025)?
- RQ2 (The Selection Bottleneck): How does the "Choice Overload" in linear chat interfaces (e.g., Midjourney) trigger decision paralysis and visual memory decay during the "Selection" phase of the attention funnel (Salma et al., 2025; Tkouat, 2024)?
- RQ3 (The Agency Gap): To what extent does the transition from "Creator" to "Adjudicator" (adopting the "AI Sandwich" workflow) impact the user's "Sense of Agency" and lead to the phenomenon of "Shadow AI use" (Nielsen, 2025; Wu et al., 2025)?

Brief Literature Review

Current scholarship identifies a paradigm shift from AI as a passive "Digital Pen" to an active "AI Co-Creator" (Haase & Pokutta, 2024). However, this evolution faces a structural bottleneck theorized as the "Attention Funnel" (Tkouat, 2024). According to Cognitive Load Theory (CLT), poor interface design increases "extraneous load," while the lack of feedback loops hinders "germane load" or deep reflection (Gkintoni et al., 2025; Tkouat, 2024).

Existing research has cataloged interaction pitfalls—such as the "Agony of Choice" and "Invisible AI Boundaries" (Salma et al., 2025). Furthermore, Western research emphasizes the efficiency of navigating the "Latent Design Space," while critical counter-narratives warn of a "mechanized convergence" where human creativity is reduced to "parameterized selection" (Nielsen, 2025; Tankelevitch et al., 2025). This study aims to fill the gap by examining these theoretical bottlenecks within naturalistic professional workflows.

Methodology

This study adopts a convergent mixed-methods approach, prioritizing naturalistic behavioral data and depth-focused narratives.

Method 1: Interaction Log Analysis (Quantifying the "Language Trap")

- **Methodological Justification:** To understand the "Articulation Barrier," we must see how users actually "talk" to the AI and where they get stuck in loops.
- **Sample:** 200 interaction sequences from Midjourney galleries (representing both professional design tasks and casual ideation).
- **Specific Operation:**
 1. **Coding "Building Blocks":** Each turn in the log is categorized using the "Interaction Building Blocks" framework: Predict-Outcome (the initial prompt), Modify-Information (tweaking text like adding "lighting" or "4k"), and Collect-Inputs (uploading reference images).
 2. **The "Refinement Ratio":** I will calculate how many times a user had to "Modify" a prompt before arriving at a final image. A high ratio (e.g., 10 prompts for 1 image) signals a bottleneck in the Action stage, showing that the user is trapped in "syntactic engineering" rather than creative flow.
 3. **Qualitative Sifting:** I will identify "Frustrated Iterations"—moments where the user changes the prompt slightly but the AI output is almost identical, illustrating "prompt instability".

Method 2: "Think-Aloud" Task Observation (Tracking the "Selection Crisis")

- **Methodological Justification:** To capture where the "Attention Funnel" leaks, we need to hear the user's internal monologue in real-time as they sift through AI variants.
- **Participants:** 12 university design students.
- **Specific Operation:**
 1. **The Comparison Task:** Half the participants use the Linear Chat UI (Midjourney), and the other half use a Spatial Canvas UI (e.g., Gemini Canvas or Box Canvas) to perform the same task: "Designing a logo for an eco-friendly brand".
 2. **Observation Metric:** While users work, they will narrate their thoughts. I will record the frequency of "Scroll-Backs" (how many times they go back to look at an old result). In Chat UIs, this indicates "visual memory decay".
 3. **Saliency Qualitative Check:** Following the Itti-Koch logic, I will ask participants: "Which image grabbed you first?" and "Why?" I will analyze if their attention was driven by visual anchors (e.g., high contrast) or semantic relevance. This reveals if the AI output is actually surviving the "Awareness" layer of the funnel.

Method 3: Semi-Structured Depth Interviews (Exposing the "Secret Cyborg")

- **Methodological Justification:** The psychological costs of AI—boredom and stigma—are best explored through personal narratives.
- **Participants:** 15 creative professionals who use GenAI in their daily work.
- **Specific Operation:**
 1. The "AI Sandwich" Probe: I will ask participants to describe their workflow using the "AI Sandwich" model. Do they feel they are the "Bottom Slice" (just a janitor polishing AI work) or the "Top Slice" (the strategic spark)?
 2. "Shadow Use" Inquiry: I will ask: "Who do you tell that you use AI?" and "Do you find yourself 'humanizing' AI outputs to hide its scent?"
 3. Motivation Mapping: Using Thematic Analysis, I will identify moments of "Psychological Deprivation"—when a task that used to be "fun" becomes a chore because the AI does the divergent part.

Method 4: Case-Based "Direct Manipulation" Evaluation (Testing the "Canvas" Solution)

- **Methodological Justification:** This method tests the "product solution" proposed in the literature: moving from "Text-to-Image" to "Text-to-Draft, Mouse-to-Final".
- **Object:** A case study of the Brickify system logic.
- **Specific Operation:**
 1. "Design Token" Walkthrough: I will simulate a task using the "Design Token" approach—reifying subjects (e.g., "an owl"), styles (e.g., "Cyberpunk"), and colors into interactive blocks that can be dragged and resized.
 2. Intent Continuity Check: I will document whether the ability to "Link" or "Group" these tokens allows for better context retention compared to writing a whole new prompt from scratch.
 3. Synthesis: I will analyze if this "Visual-Centric" paradigm actually bridges the "Gulf of Envisioning" by letting the mouse do the work that language fails to do.

Data Integration and Analysis: The Attentional Cost Matrix

The researcher will synthesize the data through an "Attentional Cost Matrix." Behavioral data from logs and observations will provide the where of bottlenecks (Tkouat, 2024), while narrative data from interviews and case studies will provide the why—explaining how these cogs reshape the creator's professional identity and motivational sustained effort (Wu et al., 2025).

Analogy for Understanding: The creative process with AI is currently like a Water Treatment Plant. Generative AI provides a massive flood of "raw water" (endless generated images), but the human brain serves as the Fine-Mesh Filter. If the filter is clogged by a narrow "Articulation Barrier" (a tiny input pipe) or overwhelmed by a "Selection Bottleneck" (sediment arriving too fast), the plant fails despite the abundance of water (Tkouat, 2024). This study seeks to optimize the structure of that filter.

References

-
- Baird, A., & Maruping, L. M. (2021). The next generation of research on IS use: A theoretical framework of delegation to and from agentic IS artifacts. *MIS Quarterly*, 45(1), 315–341. <https://doi.org/10.25300/MISQ/2021/15882>
- Bilic, I. (2025, October 30). Trust at first prompt: The new design challenge of AI interfaces. Honeycomb. <https://www.honeycomb.io/blog/trust-at-first-prompt-new-design-challenge-ai-interfaces>
- Box. (2025). Box Canvas: AI-powered visual collaboration for teams. <https://www.box.com/canvas>
- Gemini. (2025). Gemini Canvas: Write, code, & create in one space with AI. <https://gemini.google/overview/canvas/>

- Gkintoni, E., Antonopoulou, H., Sortwell, A., & Halkiopoulos, C. (2025). Challenging cognitive load theory: The role of educational neuroscience and artificial intelligence in redefining learning efficacy. *Brain Sciences*, 15(2), 203. <https://doi.org/10.3390/brainsci15020203>
- Gómez, C., Cho, S. M., Ke, S., Huang, C.-M., & Unberath, M. (2025). Human-AI collaboration is not very collaborative yet: A taxonomy of interaction patterns in AI-assisted decision making from a systematic review. *Frontiers in Computer Science*, 6, 1521066. <https://doi.org/10.3389/fcomp.2024.1521066>
- Haase, J., & Pokutta, S. (2024). Human-AI co-creativity: Exploring synergies across levels of creative collaboration. *arXiv*. <https://doi.org/10.48550/arXiv.2411.12527>
- Nielsen, J. (2025, December 15). UX Roundup: How people use AI | AI does user research | Founding designers. *UX Tigers*. <https://www.uxtigers.com/post/ux-roundup-20251215>
- Salma, Z., Hijón-Neira, R., & Pizarro, C. (2025). Designing co-creative systems: Five paradoxes in human-AI collaboration. *Information*, 16(10), 909. <https://doi.org/10.3390/info16100909>
- Shi, X., Wang, Y., Rossi, R., & Zhao, J. (2025). Brickify: Enabling expressive design intent specification through direct manipulation on design tokens. In *CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery. <https://doi.org/10.1145/3706598.3714087>
- Stackpole, B. (2024, November 13). To help improve the accuracy of generative AI, add speed bumps. *MIT Sloan*. <https://mitsloan.mit.edu/ideas-made-to-matter/to-help-improve-accuracy-generative-ai-add-speed-bumps>
- Subramonyam, H., Pea, R., Pondoc, C., Agrawala, M., & Seifert, C. (2024). Bridging the gulf of envisioning: Cognitive challenges in prompt based interactions with LLMs. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3613904.3642754>
- Khadija Tahira, Sarah Ilyas, Najam us Sahar, Sidra Mushtaq, & Shamin Batool. (2025). BUILDING USER-FRIENDLY FRAMEWORKS TO SUPPORT USABILITY AND LEARNING IN GENERATIVE AI. *Spectrum of Engineering Sciences*, 3(8), 1057–1065. Retrieved from <https://thesesjournal.com/index.php/1/article/view/918>
- Tankelevitch, L., Glassman, E. L., He, J., Kittur, A., Lee, M., Palani, S., Sarkar, A., Ramos, G., Rogers, Y., & Subramonyam, H. (2025). Understanding, protecting, and augmenting human cognition with generative AI: A synthesis of the CHI 2025 Tools for Thought workshop. *arXiv*. <https://arxiv.org/abs/2508.21036>
- Tkiouat, Z. (2024). The attentional cost of IT interference and delegation in highly dynamic and risky contexts: The case of digitized vehicles [Doctoral dissertation, HEC Montréal]. HEC Montréal. https://biblos.hec.ca/biblio/theses/tkiouat_zoubeir_t2024.pdf
- Wang, W.-F., Lu, C.-T., Companyà, N. P., Chen, B.-Y., & Chen, M. Y. (2025). AIdeation: Designing a human-AI collaborative ideation system for concept designers. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '25)*. <https://doi.org/10.1145/3706598.3714148>
- Webmoghuls. (2025, September 29). 10 AI UI design best practices for 2026. <https://www.webmoghuls.com/10-ai-ui-design-best-practices-2026/>
- Wu, S., Liu, Y., Ruan, M., Chen, S., & Xie, X.-Y. (2025). Human-generative AI collaboration enhances task performance but undermines human's intrinsic motivation. *Scientific Reports*, 15, 15105. <https://doi.org/10.1038/s41598-025-98385-2>