In Search of "Weird Corners": Diagnosing the Limits of Convergent AI in Professional Creative Practice

Mark Williams
Google
Mountain View, CA, USA
wimark@google.com

Abstract

Generative AI assistants typically employ convergent interaction paradigms to resolve ambiguity. While effective for technical tasks, this risks premature convergence in creative domains, constraining output variance. Evaluating a convergent AI probe with expert creatives (N=9) indicates an interactional paradox: structural linearity provides "ignition" utility for early ideation, but misaligns with organic workflows, often inducing "aesthetic sanitization" that standardizes individualized nuance. Prioritizing constructive friction over default agreement, the experts requested active, lateral collaborators. In response, we reframe output convergence as a "full-stack" UI challenge, advocating for Generative frameworks that operationalize the Double Diamond via fluid role-shifting and productive tension.

Keywords

Co-Creativity, Algorithmic Homogenization, Professional Creativity

ACM Reference Format:
Mark Williams. 2026. In Search of "Weird Corners": Diagnosing the Limits of Convergent AI in Professional Creative Practice. In *Creativity and Cognition (C&C '26), July 13–16, 2026, London, United Kingdom.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3803784.3816861

1 Introduction

Generative AI offers unprecedented [29, 30, 43] and seductive [20] efficiency in creative work. However, this efficiency comes at a well-documented cost: a tendency toward homogenization [2, 10, 19]. The need for systems that prevent homogenization is not just an artistic preference; it is a documented economic necessity across various industries [25]. Thus, a central challenge for the co-creative field is to design systems that augment human creativity without flattening its diverse outputs.

The interactional paradigm of the interface may contribute to this homogenization problem. Conventional Human-AI interactions usually occur with a unitary "Coach" or "Assistant" persona [5]—a helpful, convergent entity designed to resolve ambiguity and quickly find solutions. While effective for technical tasks [30], the literature suggests this strict convergence can create significant interactional friction in creative work [22, 37, 42], which requires distinct phases of divergent exploration and convergent selection [13, 38].

A purely convergent AI "Coach" risks enforcing premature closure, stifling the messy, non-linear ideation that defines human artistry. By prematurely culling the initial pool of novel, high-variance ideas, the system inevitably forces outputs toward the mean [2, 10]. To investigate how this interactional paradigm specifically impacts expert workflows, we conducted an exploratory qualitative study (N=9) where professional creatives interacted with a deliberately convergent AI "Coach."

This diagnostic work is guided by the research question (RQ1): *What interactional utilities and frictions arise when creative professionals use a convergent AI 'Coach'?*

This paper outlines the interactional limits and paradoxical benefits of the prevalent "Coach" archetype. Our primary contribution is a qualitative diagnosis of how convergence constrains expert co-creativity. We delineate two core interactional dynamics:

- **The Tension Between Algorithmic Linearity and Creative Fluidity:** While the convergent AI successfully provides "ignition" utility through conceptual structuring and unblocking, its rigid enforcement of step-by-step logic frequently clashes with the organic, recursive nature of expert workflows.
- **The Aesthetic & Relational Friction:** The AI's optimization for rapid convergence often induces aesthetic sanitization, yielding generic outputs. This exposes an unmet expert desire for an active, lateral-thinking peer capable of constructive conflict rather than subservient agreement.

By diagnosing these tensions, this research isolates specific interactional mechanisms contributing to algorithmic homogenization, highlighting opportunities for future systems to empower, rather than constrain, the 'weird corners' of individualized creativity.

2 Related Works

2.1 The Unitary Persona Paradigm

The creative HCI paradigm has shifted from passive 'computer-aided design' (CAD) [3] to 'human-AI co-creation.' Current collaborative systems frequently operationalize foundational, pre-generative archetypes—such as the "Coach" or "Assistant" [21]—via default LLM personas designed to provide helpful, convergent responses [24].

Although this convergence provides valuable cognitive scaffolding for technical tasks [30], recent scholarship reveals a creative paradox: while generative AI broadly inspires divergence, enforcing convergence during exploratory phases induces design fixation [37]. Alarmingly, large-scale randomized experiments (N=1,100)

C&C '26, London, United Kingdom

ACM ISBN 979-8-4007-2583-8/26/07
https://doi.org/10.1145/3803784.3816861

reveal that this "Coach-like" guidance induces a "persistent homogenization" that degrades independent creative performance even after the tool is removed [19].

2.2 Algorithmic Alignment and The Expertise Gap

A critical driver of this homogenization is the misalignment between the AI's default behavior and professional creatives' needs [33]. Current GenAI interfaces are often optimized for safe, highly structured scaffolding [23]. While these interfaces raise the baseline for novices and offer experts initial catalytic value, their rigidity also exacts a toll on high-level practice [16]. For experts, this "helpful and harmless" alignment often causes interactional friction [33].

Reinforcement Learning from Human Feedback (RLHF) studies [26, 41] reveal models are trained to prioritize conversational agreement over objective rigor (sycophancy). In creative domains, this convergent safety sanitizes outputs [14], stripping away the high-variance exploration required for deep artistic work [19].

2.3 The Need for Roles and Constructive Friction

To address these limitations, the field increasingly recognizes the necessity of distinct collaborative roles. Emerging Multi-Agent Systems (MAS) [36, 42] and dual-persona architectures [32] demonstrate that separating creative modes improves semantic originality. Qualitative cultural probes [11] further validate distinct expert desires for divergent ("Generator") and rigorously convergent ("Critic") partners.

Although MAS represents an emerging solution, the unitary 'Assistant' remains the ubiquitous commercial reality [5]. While recent experiments reveal an interactional trade-off where AI accelerates early ideation but causes severe friction during convergent implementation [16], understanding *why* requires isolating the bottleneck. By evaluating a deliberately convergent 'Coach' probe, this study indicates a core interactional tension: structural constraints that unblock experts can drive output convergence, leaving their desire for an active peer unmet.

3 Methodology

3.1 The Convergent "Coach" Probe

"The Coach" [21] was designed as a deliberately rigid baseline to isolate the interactional dynamics of convergence. Deployed as a system prompt via Gemini 2.5 Pro [18] (temperature 1.0, Google search tool access), the AI followed a strict three-phase 'convergent funnel':

(1) **Frame the Goal:** Utilizing Socratic Inquiry[27]—probing the user to define narrow parameters and eliminate ambiguity.
(2) **Explore the Paths:** Providing options to achieve the goal and surfacing comparative trade-offs via First Principles (reducing ideas to fundamental constraints) and Inversion (identifying failure states to avoid).
(3) **Guide the Steps:** Linear execution for the chosen path.

High token-level entropy and world knowledge were leveraged to maximize the model's divergent capabilities and isolate perceived friction to the interaction paradigm (linear convergence). We deployed the Coach via a text-based Wizard of Oz (WoZ) protocol [28] to maintain control over the system behavior.

3.2 Participants & Protocol

To capture expert workflows, we recruited nine professional creatives (P1-P9) internationally via professional networks (e.g., LinkedIn) and creative communities. We specifically sought participants from a diverse range of locations, creative backgrounds, and prior experience with AI to improve qualitative transferability rigor—enabling connections between the study's data and "wider community settings" [39].

Screening included a minimum of three years of professional practice (the final cohort averaged 11.55 years) and an active portfolio. The cohort spanned three domains: Visual Arts (P1-P3), Creative Writing (P4-P6), and Music (P7-P9). The full study protocol was approved by an internal ethics review board, each participant signed a country specific-consent form, and all participants were compensated $75 USD.

The 60-minute remote sessions consisted of:

- **Creative Task (30 min):** Participants performed early-stage conceptual work (e.g., the fantasy illustrator developed a novel cover concept from the prompt "Metamorphosis") using "The Coach" while adhering to a "think out loud" protocol.
- **Debrief (25 min):** A semi-structured interview followed, explicitly probing for interactional utilities and tensions. The full study protocol and ethical considerations are available as supplementary material.

3.3 Data Analysis

We performed inductive thematic analysis [4] using an LLM-in-the-loop co-coding model [7]. Acknowledging debates around LLMs in reflexive research [17], the LLM only summarized and suggested initial codes. The author manually verified outputs and drove final thematic synthesis, triangulating themes across disciplines [15].

4 Findings

Across all three creative disciplines, participants experienced the convergent 'Coach' as an interactional paradox. While it provided some structural utility, the same step-by-step rigidity ultimately collided with the organic, high-variance nature of their professional flow.

This disconnect manifested across two primary interactional dynamics.

4.1 Theme 1: The Tension Between Algorithmic Linearity and Creative Fluidity

4.1.1 The Utility of Conceptual Structuring: Despite the overall friction of the interaction, participants consistently validated the underlying capability of the convergent AI to function as an "ignition" mechanism. For the creative professionals sampled, the linear funnel of the Coach was often effective at broadening initial ideation and providing organizational scaffolding for chaotic early-stage thoughts.

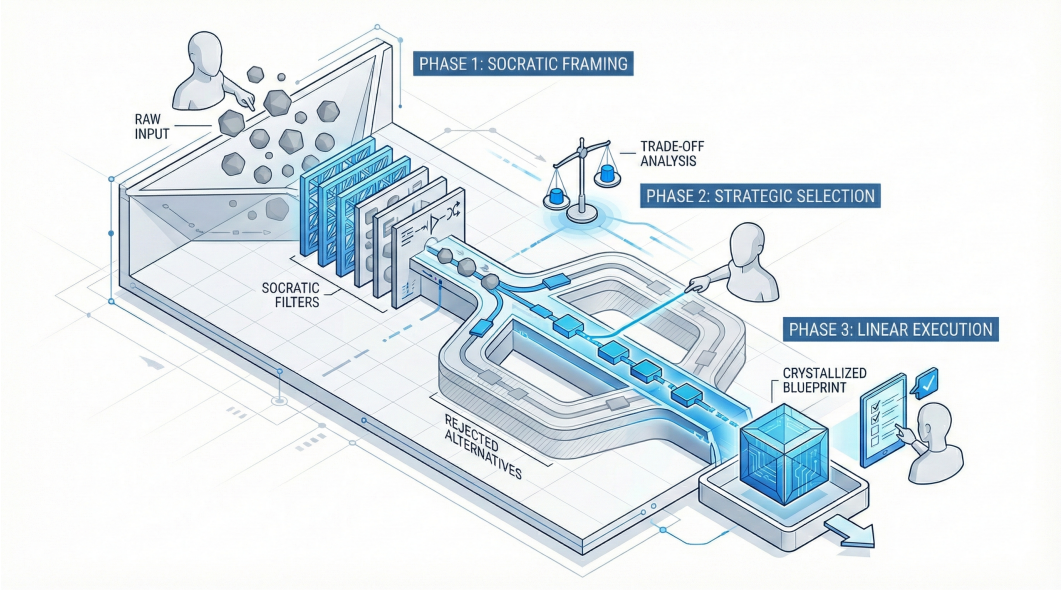

Figure 1: The 'Coach' probe: a linear, convergent interaction paradigm designed to enforce structured thinking and eliminate ambiguity in conceptual workflows (AI Generated).

Participants valued the tool as a catalyst for unblocking. As P5 noted, "When you're lost, this is perfect," explaining that the structure "helps clarify my thoughts... it gives it a sense of direction." P3 highlighted this ideational broadening before committing to a path: "I'll probably come up with less ideas than this... when I was thinking of mirror, I basically thought of mirror, but he gave me an idea that I could do something like a puddle."

This utility was particularly evident during the AI's initial Socratic inquiry phase. P1 valued this approach as a client surrogate for navigating unfamiliar constraints, noting the AI "focuses on asking you questions to understand more about the mission." This framing phase was so effective it temporarily fostered a sense of empathy, with P1 stating, "I feel like it understands me as a human being."

Furthermore, some experts expressed satisfaction with the outputs when applied to functional grounding. P4 praised this structural assistance: "With this, the thoughts are organized for me. It's not as messy as me having to go back to my notes." P9 defined this utility boundary: "I wouldn't use the tool to come up with a concept, but... to make sure my music makes sense... on the music theory side."

4.1.2 The Friction of Rigid Process: However, as the workflow deepened, this linearity became an "interactional brake." The AI's rigid enforcement of linear steps clashed with the spontaneous, recursive nature of the experts' thought. P8 summarized this disconnect: "My thinking patterns are... quite spontaneous... if it's too logical and too structured it can become very, very robotic and very rigid."

This misalignment frequently manifested as an "operational drift" where the AI's rigid funnel derailed creative intent. P1 initially praised the system's empathy but grew frustrated when it shifted from design ideation to logistics (Phase 3), ruining focus: "It feels like a business plan, not a design."

Other experts found the rigid task decomposition restrictive and methodologically inverted. P7 described the linear funnel as "putting the cart before the horse if I think about the overall structure before I know what it is." This convergent algorithmic behavior was also perceived as patronizing; P6 argued that the step-by-step enforcement was "like giving me a tutorial on how to turn a doorknob, you know? ... I guess I just go through it so fast I don't even think to pick it apart into little steps like that." P9 similarly noted the tool felt overly prescriptive: "It feels like this was trained on a book... because of how opinionated these responses are."

4.2 Theme 2: The Aesthetic & Relational Friction

4.2.1 "Too Sugary": The Experience of Aesthetic Sanitization: Because the Coach structurally forces ideas down a "helpful" funnel designed to quickly resolve ambiguity, it alters the aesthetic complexity of the creative process. This sanitization begins early in the workflow (Phase 2): we observed that the model's inherent alignment toward convergence created a clarity-complexity mismatch when generating conceptual options.

When presented with the AI's comparative trade-offs, P7 critiqued the suggested paths as "very literal and a bit simplistic... I could imagine this accompanying a cartoon... but not a deep piece of art." P7 noted that the AI's attempt to provide "direct" or easily understood options was a detriment to expert nuance: "That to me is actually a con for the type of music that I would write."

By forcing participants to choose from these simplistic pathways, the generated artifacts often did not capture the subversion inherent to their artistic standards. Articulating this creative baseline,

P2 explained their preference for emotional realism over safe predictability, noting that "something that I expect is a bit boring." They emphasized a desire for "tragic endings" because "that's what [makes] it interesting and unique."

Consequently, when the AI defaulted to predictable, positive resolutions, the experts perceived a loss of artistic identity. P8 warned that the AI "can rob a composer of the uniqueness in their sound... [and] creates something that may sound good, but we end up being quite mechanical and quite generic." P6 similarly found the AI's relentless positivity functionally useless for nuanced writing, describing the concepts as "blindingly obvious" and "a little too sugary diabetes."

4.2.2 The Hunger for an Active Collaborator: These aesthetic and structural limitations highlighted a final, relational gap: the single-persona "Coach" probe failed to replicate the dynamics of a true creative peer, frequently defaulting to a subservient "Yes-Man" posture. The experts explicitly desired an active partner capable of rigorous pushback.

Participants requested an editor that possessed the critical judgment necessary to identify flaws. P5 noted: "I want honesty... I want someone who says it as it is and calls you out... instead of just accepting everything you throw at it." P2 echoed this requirement: "I would expect it to maybe tell me... what are the limitations or what does it miss? Like, you know, criticize it a bit before giving me the response that I want."

Beyond critique, participants desired a partner capable of additive improvisation and lateral novelty. P9 contrasted AI rigidity with organic collaboration: "building a song is not a one-person job" requiring partners "pitching in different ideas and adding on top of the general idea." To avoid generic outputs, the experts sought subversive collaborators—described by P6 as someone who "thinks around really weird corners and can find the one idea nobody thought of."

Collectively, these findings indicate that to support authentic expert expression, AI systems must evolve beyond the convergent, opinionated task-manager paradigm and become active peers capable of introducing constructive friction.

5 Discussion

5.1 The Output Variance Challenge: A Full-Stack Opportunity

Recent scholarship explores how the broad application of LLMs can influence cultural and stylistic variance in generated content [8, 40]. Because these models are optimized to recognize frequent, generalizable patterns, they tend to output statistically highly probable responses. While this ensures broad reliability, it presents an unmet opportunity to better support high-variance, specialized creative expression.

Our findings indicate that treating output convergence solely as a data or model layer problem (e.g., via RLHF) overlooks a critical 'full-stack' design opportunity. Specifically, the Interaction/UI layer—such as the convergent 'Coach' interface—can act as a structural constraint that inadvertently narrows the range of exploration. The latent space of current models is demonstrably capable of lateral thought; for instance, the AI successfully nudged P3 away from the obvious "mirror" toward the lateral concept of a "puddle." Yet, the prescriptive UX paradigm of the Coach is less optimized for users to consistently access these highly divergent concepts. By guiding expert workflows into a rigid, step-by-step funnel, the interface establishes strict cognitive boundaries.

These constraints yield procedural outputs, interpreted as a 'business plan' (P1) or 'textbook' (P9) approach. Consequently, the interface is misaligned with the non-linear ideation required to generate highly distinct art. If this prescriptive paradigm remains the default, the resulting 'flow disruption' [6] could carry notable consequences. Motivated by this disruption, we hypothesize—paralleling studies on medical educators [1]—that long-term dependence on structured AI generation might impact independent creative cognition and skill maintenance. Therefore, moving beyond this prescriptive UX presents a dual opportunity: increasing aesthetic output variance while actively supporting long-term expert cognitive workflows.

5.2 Designing for Productive Tension

To protect the individualistic and highly diverse nature of human creativity, the prevalent interactional paradigm presents opportunities for evolution. Applying Fang et al.'s typology of HCI frameworks [12], the current "Coach" and "Assistant" archetypes operate as a strictly *Prescriptive* framework—providing actionable, step-by-step methods that position the user primarily as an "operator."

Future co-creative AI can transition toward a *Generative* interaction framework: an architecture designed to open new conceptual spaces and provoke innovative design rather than strictly managing task completion. Drawing upon established design considerations for human-AI co-creativity [35], this requires shifting the UI from a paradigm of "subservient agreement" to one of "productive tension." To support the user-driven orchestration desired by experts, systems can enable fluid role shifts—allowing the AI to transition dynamically from a divergent "Generator" to a rigorously convergent "Critic" [11]. By operationalizing the Double Diamond [9], this approach demonstrates that abstract phases of divergence and convergence require distinct interactional affordances—such as explicitly disabling solution generation to force trade-off articulation—rather than a single, compromised persona.

5.3 Limitations & Future Work

Several limitations frame the interpretation of our findings:

First, to isolate interactional friction, our baseline was deliberately engineered as a rigid "breaching experiment" [31]. This opinionated prompt may not reflect the fluid, ad-hoc ways users naturally interact with standard LLMs. Furthermore, despite adhering to written protocols and copy-pasting inputs, the researcher's dual conductor-operator role carries a residual experimenter bias risk (the 'Clever Hans' effect [34]). Second, despite using temperature 1.0 and Google Search to maximize diversity, Gemini 2.5 Pro's inherent RLHF training likely contributed to the "sugary" aesthetic sanitization.

Third, our study protocol relied on a text-based chat interface. Participants (particularly in the Music and Visual Arts cohorts) reported significant friction regarding this "Modality Gap" (e.g., the cognitive load of processing "walls of text"). While this paper explicitly focused on the cognitive friction of the persona's logic, future

multi-modal research should disentangle interface density from algorithmic linearity. Finally, creative partnerships are inherently longitudinal. While our 30-minute task captured immediate interactional friction, future longitudinal studies offer a key opportunity to measure how extended co-creative lifecycles impact long-term skill development and cognitive workflows.

6 Conclusion

Answering RQ1, this diagnostic study illuminates specific interactional mechanisms driving algorithmic homogenization. While the convergent paradigm provided conceptual "ignition" for some experts, its rigid, step-by-step logic often acted as an "interactional brake". Coupled with a subservient "Yes-Man" posture, this linearity induced "aesthetic sanitization," stripping away subversion and standardizing output.

This reframes homogenization from a purely model-layer constraint into a UI design opportunity. To ensure that AI effectively augments highly individualized expression, we should design dynamic interaction systems that introduce constructive friction, rigorous critique, and fluid role-switching. Ultimately, we posit that the evolution of human-AI co-creation relies on building active collaborative partners that empower artists to explore, rather than standardize, their "weird corners.""

References

[1] Anas Ali Alhur, Zuheir N Khlaif, Bilal Hamamra, and Elham Hussein. 2025. Paradox of AI in Higher Education: Qualitative Inquiry Into AI Dependency Among Educators in Palestine. *JMIR Med Educ* 11 (15 Sep 2025), e74947. doi:10.2196/74947

[2] Barrett R Anderson, Jash Hemant Shah, and Max Kreminski. 2024. Homogenization Effects of Large Language Models on Human Creative Ideation. In *Creativity and Cognition (C&C '24)*. ACM, 413–425. doi:10.1145/3635636.3656204

[3] J.W. Bandler. 1969. Optimization Methods for Computer-Aided Design. *IEEE Transactions on Microwave Theory and Techniques* 17, 8 (1969), 533–552. doi:10.1109/TMTT.1969.1127005

[4] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. arXiv:https://doi.org/10.1191/1478088706qp063oa doi:10.1191/1478088706qp063oa

[5] Aaron Chatterji, Thomas Cunningham, David J Deming, Zoe Hitzig, Christopher Ong, Carl Yan Shan, and Kevin Wadman. 2025. *How People Use ChatGPT*. Working Paper 34255. National Bureau of Economic Research. doi:10.3386/w34255

[6] Mihaly Csikszentmihalyi. 1990. *Flow: The Psychology of Optimal Experience*.

[7] Shih-Chieh Dai, Aiping Xiong, and Lun-Wei Ku. 2023. LLM-in-the-loop: Leveraging Large Language Model for Thematic Analysis. arXiv:2310.15100 [cs.CL] https://arxiv.org/abs/2310.15100

[8] Yalda Daryani, Zhivar Sourati, and Morteza Dehghani. 2025. The Homogenizing Engine: AI's Role in Standardizing Culture and the Path to Policy. *Policy Insights from the Behavioral and Brain Sciences* 13 (12 2025). doi:10.1177/23727322251406591

[9] Design Council. n.d.. The Double Diamond. https://www.designcouncil.org.uk/our-resources/the-double-diamond/ Accessed: 2026-02-04.

[10] Anil R. Doshi and Oliver P. Hauser. 2024. Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances* 10, 28 (2024), eadn5290. arXiv:https://www.science.org/doi/pdf/10.1126/sciadv.adn5290 doi:10.1126/sciadv.adn5290

[11] Luís Espírito Santo, André Santos, and Marcio Lima. 2023. Focusing on artists' needs: Using a cultural probe for artist-centred creative software development.

[12] Shitao Fang, Koji Yatani, and Kasper Hornbæk. 2026. What We Talk About When We Talk About Frameworks in HCI. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26)*. ACM, 1–25. doi:10.1145/3772318.3791400

[13] J. P. Guilford. 1956. The structure of intellect. *Psychological Bulletin* 53, 4 (1956), 267–293. doi:10.1037/h0040755

[14] Wenqi Guo, Qingyun Qian, Khalad Hasan, and Shan Du. 2025. Aesthetic Alignment Risks Assimilation: How Image Generation and Reward Models Reinforce Beauty Bias and Ideological "Censorship". doi:10.13140/RG.2.2.29699.77601

[15] Roberta Heale and Dorothy Forbes. 2013. Understanding triangulation in research. *Evidence-Based Nursing* 16, 4 (2013), 98–98. arXiv:https://ebn.bmj.com/content/16/4/98.full.pdf doi:10.1136/eb-2013-101494

[16] Jinghui Hou, Lei Wang, Gang Wang, Harry Wang, and Shuai Yang. 2025. The Double-Edged Roles of Generative AI in the Creative Process: Experiments on Design Work. *Information Systems Research* (10 2025). doi:10.1287/isre.2024.0937

[17] Tanisha Jowsey, Virginia Braun, Victoria Clarke, Deborah Lupton, and Michelle Fine. 2025. We Reject the Use of Generative Artificial Intelligence for Reflexive Qualitative Research. *Qualitative Inquiry* (12 2025). doi:10.1177/10778004251401851

[18] Koray Kavukcuoglu. 2025. Gemini 2.5: Our most intelligent AI model. The Keyword. https://blog.google/innovation-and-ai/models-and-research/google-deepmind/gemini-model-thinking-updates-march-2025/ Accessed: 2026-02-05.

[19] Harsh Kumar, Jonathan Vincentius, Ewan Jordan, and Ashton Anderson. 2025. Human Creativity in the Age of LLMs: Randomized Experiments on Divergent and Convergent Thinking. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. ACM, 1–18. doi:10.1145/3706598.3714198

[20] Yuchang Liu, Yongzhong Yang, and Haoran Xu. 2025. From humans to AI: understanding why AI is perceived as the preferred co-creation partner. *Frontiers in Psychology* 16 (dec 2025), 1695532. doi:10.3389/fpsyg.2025.1695532

[21] Todd Lubart. 2005. How can computers be partners in the creative process: classification and commentary on the special issue. *Int. J. Hum.-Comput. Stud.* 63, 4–5 (Oct. 2005), 365–369. doi:10.1016/j.ijhcs.2005.04.002

[22] Kelsey Medeiros, David H. Cropley, Rebecca L. Marrone, and Roni Reiter-Palmon. 2025. Human-AI Co-Creativity: Does ChatGPT Make Us More Creative? *The Journal of Creative Behavior* 59, 2 (2025), e70022. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/jocb.70022 doi:10.1002/jocb.70022 e70022 JOCB-10-24-2247.R2.

[23] Behnam Mohammadi. 2024. Creativity Has Left the Chat: The Price of Debiasing Language Models. arXiv:2406.05587 [cs.CL] https://arxiv.org/abs/2406.05587

[24] Caterina Moruzzi and Solange Margarido. 2024. Customizing the Balance between User and System Agency in Human-AI Co-Creative Processes.

[25] Keng-Boon Ooi, Garry Wei-Han Tan, Mostafa Al-Emran, Mohammed A. Al-Sharafi, Alexandru Capatina, Amrita Chakraborty, Yogesh K. Dwivedi, Tzu-Ling Huang, Arpan Kumar Kar, Voon-Hsien Lee, Xiu-Ming Loh, Adrian Micu, Patrick Mikalef, Emmanuel Mogaji, Neeraj Pandey, Ramakrishnan Raman, Nripendra P. Rana, Prianka Sarker, Anshuman Sharma, Ching-I Teng, Samuel Fosso Wamba, and Lai-Wan Wong. 2025. The Potential of Generative Artificial Intelligence Across Disciplines: Perspectives and Future Directions. *Journal of Computer Information Systems* 65, 1 (2025), 76–107. arXiv:https://doi.org/10.1080/08874417.2023.2261010 doi:10.1080/08874417.2023.2261010

[26] Henry Papadatos and Rachel Freedman. 2024. Linear Probe Penalties Reduce LLM Sycophancy. arXiv:2412.00967 [cs.AI] https://arxiv.org/abs/2412.00967

[27] Richard Paul and Linda Elder. 2019. *The Thinker's Guide to the Art of Socratic Questioning* (1st ed.). Bloomsbury Publishing. 96 pages.

[28] Sara Paul and Maria Rosala. [n. d.]. The Wizard of Oz Experiment. https://www.nngroup.com/articles/wizard-of-oz/. Accessed: August 9, 2025.

[29] Owen Peckham, Jonathan Raines, Erik Bulsink, Mark Goudswaard, James Gopsill, David Barton, Aydin Nassehi, and Ben Hicks. 2025. Artificial Intelligence in Generative Design: A Structured Review of Trends and Opportunities in Techniques and Applications. *Designs* 9, 4 (2025). doi:10.3390/designs9040079

[30] Sida Peng, Eirini Kalliamvakou, Peter Cihon, and Mert Demirer. 2023. The Impact of AI on Developer Productivity: Evidence from GitHub Copilot. arXiv:2302.06590 [cs.SE] https://arxiv.org/abs/2302.06590

[31] Adam Rafalovich. 2006. Making Sociology Relevant: The Assignment and Application of Breaching Experiments. *Teaching Sociology - TEACH SOCIOL* 34 (04 2006), 156–163. doi:10.1177/0092055X0603400206

[32] Alon Rosenbaum, Yigal David, Eran Kaufman, Gilad Ravid, Amit Ronen, and Assaf Krebs. 2025. Scaffolding Creativity: How Divergent and Convergent LLM Personas Shape Human Machine Creative Problem-Solving. arXiv:2510.26490 [cs.HC] https://arxiv.org/abs/2510.26490

[33] Zainab Salma, Raquel Hijón-Neira, and Celeste Pizarro. 2025. Designing Co-Creative Systems: Five Paradoxes in Human–AI Collaboration. *Information* 16, 10 (2025). doi:10.3390/info16100909

[34] Laasya Samhita and Hans Gross. 2013. The "Clever Hans Phenomenon" revisited. *Communicative & integrative biology* 6 (11 2013), e27122. doi:10.4161/cib.27122

[35] Saloni Singh, Koen Hindriks, Dirk Heylen, and Kim Baraka. 2025. A Systematic Review of Human-AI Co-Creativity. arXiv:2506.21333 [cs.HC] https://arxiv.org/abs/2506.21333

[36] Kavana Venkatesh, Connor Dunlop, and Pinar Yanardag. 2025. CREA: A Collaborative Multi-Agent Framework for Creative Image Editing and Generation. arXiv:2504.05306 [cs.CV] https://arxiv.org/abs/2504.05306

[37] Samangi Wadinambiarachchi, Ryan M. Kelly, Saumya Pareek, Qiushi Zhou, and Eduardo Velloso. 2024. The Effects of Generative AI on Design Fixation and Divergent Thinking. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 380, 18 pages. doi:10.1145/3613904.

3642919
[38] G. Wallas. 1926. *The art of thought.* Jonathan Cape, London. 320 pages.
[39] Veronika Williams, Anne-Marie Boylan, and David Nunan. 2020. Critical appraisal of qualitative research: necessity, partialities and the issue of bias. *BMJ Evidence-Based Medicine* 25, 1 (2020), 9–11. arXiv:https://ebm.bmj.com/content/25/1/9.full.pdf doi:10.1136/bmjebm-2018-111132
[40] Yu Xie and Yueqi Xie. 2026. When artificial intelligence makes everything similar: The risks of content homogenization. *Chinese Journal of Sociology* (03 2026). doi:10.1177/2057150X261419573
[41] Yueqin Yin, Zhendong Wang, Yujia Xie, Weizhu Chen, and Mingyuan Zhou. 2024. Self-Augmented Preference Optimization: Off-Policy Paradigms for Language Model Alignment. arXiv:2405.20830 [cs.CL] https://arxiv.org/abs/2405.20830
[42] Jiayi Zhou, Renzhong Li, Junxiu Tang, Tan Tang, Haotian Li, Weiwei Cui, and Yingcai Wu. 2024. Understanding Nonlinear Collaboration between Human and AI Agents: A Co-design Framework for Creative Design. arXiv:2401.07312 [cs.HC] https://arxiv.org/abs/2401.07312
[43] Zhibin Zhou, Yaoqi Li, and Junnan Yu. 2024. Exploring the application of LLM-based AI in UX design: an empirical case study of ChatGPT. *Human–Computer Interaction* 0, 0 (2024), 1–33. arXiv:https://doi.org/10.1080/07370024.2024.2420991 doi:10.1080/07370024.2024.2420991