Generative UI–Enabled Experience Design: Learnings from early practitioners

From shifting team vision and creative direction to mastering the prompt-to-polish loop, discover how practitioners are redefining roles, processes, and psychological safety in enterprise-scale HCI.

Bradley McGraw

Experience Strategy and Design (XSD) Studio at Microsoft

bradley.mcgraw@microsoft.com

Guillermo Barquero

Experience Strategy and Design (XSD) Studio at Microsoft

gbarquero@microsoft.com

Samuel Martin

Experience Strategy and Design (XSD) Studio at Microsoft

samuelmartin@microsoft.com

ABSTRACT

Generative UI (GenUI) promises to compress the distance between intent and interface, yet its implications for HCI practice remain underspecified. Drawing on early, in-situ work designing and validating GenUI enabled experiences in a large-scale enterprise setting, we synthesize five practitioner facing learnings that surface how human centered methods must adapt when UI is produced through human and AI collaboration. We first describe an applied framework that couples integrated (not substitutive) AI tooling with foundational user research, shared design briefs, rapid GenUI iteration, and subsequent handoff into conventional design and engineering workflows. We then report key learnings: (1) GenUI can shift effort from tool syntax toward product vision when teams share a structured source of truth; (2) GenUI enables rapid experimentation in experience, elevating the need to assess interaction feel, flow, and personality with sufficiently realistic prototypes; (3) designers' roles shift toward creative direction and stewardship of responsible use, supported by "play first, critique later" rituals; (4) psychological safety becomes a prerequisite for exploration to avoid regressing to safe, conventional ideas under perceived automation threat; and (5) prompting alone does not achieve production polish, motivating a repeatable "prompt to polish" loop that reintroduces human precision for usability and quality. Together, these learnings outline actionable adjustments to processes, roles, and evaluation practices that can help HCI practitioners leverage GenUI while preserving human agency, rigor, and trust in emergent AI mediated interfaces.

1 INTRODUCTION

As Generative UI (GenUI) augments research, design, and delivery practices, HCI methods must evolve to account for UI produced through human and AI collaboration. Based on our early practitioner work designing and validating GenUI-enabled experiences in large-scale enterprise environments, we identify five considerations that reshape HCI practice, spanning process adjustments, role shifts, team dynamics, and evaluation expectations.

These learnings emerged from in-situ work addressing production-level constraints such as scale, system behavior under load, and global user diversity within the complexities of one of the world's largest cloud and enterprise software providers. We further refined them by translating internal early-adopter practices to external enterprise contexts. Together, they serve as foundational considerations for adapting HCI practice as GenUI becomes integrated into design workflows.

2 LIMITATIONS AND SCOPE

These learnings reflect early practitioner work in large-scale enterprise contexts and prioritize process, role, and evaluation implications over claims of universal generality. They should be interpreted as formative guidance for GenUI-enabled experience design, and future work should validate how these practices transfer across different team sizes, domains, and organizational constraints.

3 APPLIED FRAMEWORK

Our GenUI design approach is grounded in three pillars: human-centered research, iterative collaboration, and integration of AI tooling, rather than a reliance on it. We begin with foundational research to identify user needs and business objectives, synthesizing prior findings and conducting targeted studies to address gaps. We formalized these insights into a shared design brief that defines objectives, users, scope, success metrics, and design principles, providing a stable source of truth across designers and AI systems.

We used GenUI for rapid iteration, curating and refining outputs against research-backed intent. Once interaction direction stabilized, work transitioned into conventional design and engineering workflows, returning to GenUI as new needs emerged.

This process aligns with Lee's [2025] characterization of GenUI as iterative co-creation, while extending it through continuous research integration and user feedback. Compared to traditional UX processes, GenUI's pace and output plausibility shift where rigor is required, forcing earlier alignment through briefs as infrastructure, delaying critique to preserve exploration, and necessitating explicit evaluation gates.

4 KEY LEARNINGS

Across our use of GenUI, we observed a consistent pattern: these tools reshape where design effort is spent, how exploration unfolds, and when evaluation must occur. While GenUI lowers the cost of producing plausible artifacts, it simultaneously raises the importance of shared intent, social conditions, and explicit evaluation practices. The following learnings describe how GenUI alters established HCI dynamics and where practitioners must deliberately intervene to preserve design quality.

4.1 GenUI tools free teams to focus on product vision over syntax and tooling

GenUI reduces time spent on low-level construction, enabling designers to focus on higher-order concerns such as intent, behavior, and experience. Prior HCI research on creativity support tools shows that reducing construction effort can amplify human judgment—while also introducing automation bias risks if not scaffolded [2, 3].

However, without a shared source of truth, in our experience, GenUI outputs can optimize for local plausibility rather than coherent product vision. Fragmented prompts or constraints produce divergent but superficially convincing directions that are difficult to reconcile. Treating structured design briefs as first-class infrastructure ensures both AI systems and human collaborators operate from the same semantic grounding, preserving vision as the primary driver of design quality.

4.2 GenUI processes allow for rapid experimentation in experience feel

By lowering the cost of generating interaction variants, GenUI enables rapid exploration of experience feel, flow, and personality. Because users form impressions of an experience within milliseconds, experiential qualities strongly shape perceived usability and acceptance, making "feel" a first-order design variable [4]. A recurring failure mode occurs when GenUI artifacts are evaluated using representations that cannot convey timing, interaction dynamics, or embodied flow. While low-fidelity prototypes support reasoning about structure, they can obscure or miscommunicate interaction feel and experiential qualities, particularly when teams do not make explicit what aspects of the experience a

prototype is intended to represent [5]. To counter this, HCI practitioners should introduce explicit experience-focused review criteria and require demonstrations or tests with sufficient realism for critical interactions before progressing work.

4.3 Designers shift their roles toward creative direction, emphasizing play

As GenUI expands the speed and volume of iterations, designers increasingly act as creative directors, shaping the design space through prompts and constraints. Early GenUI outputs often function as strong example solutions, increasing the risk of design fixation and premature convergence [6]. When designers introduce critique too early, they amplify fixation and convergence risks. Early critique can legitimize initial examples, suppress divergence, and discourage designers from challenging or reframing AI-generated directions.

Research on evaluation apprehension shows that early evaluation suppresses divergence in human–computer collaboration [7], while creativity benefits from selective, deep critique rather than constant feedback [8]. We therefore treat play as a bounded, inquiry-driven phase followed by an explicit transition into structured critique ("play first, critique later"), separating time-boxed divergence from convergence with clear evaluative criteria.

4.4 Psychological safety enables deeper exploration

Play-first only works when the social conditions protect disagreement with the AI output. When early GenUI outputs function as example solutions, teams risk design fixation by anchoring subsequent exploration on their structure, which narrows the design space and reduces originality [6]. In our GenUI work, teams that treated psychological safety as an explicit design condition sustained deeper exploration, with more creative, original ideas.

This was operationalized through rituals that separated exploratory sessions from evaluative reviews, creating protected space to challenge AI outputs without reputational risk. Without explicit norms separating exploration from evaluation, early judgment increases evaluation apprehension, reducing designers' willingness to propose unconventional alternatives in collaborative settings [7].

4.5 Prompting alone fails to deliver the level of polish expected in GenUI experiences

GenUI enables rapid production of artifacts that appear polished before usability and interaction quality are adequately validated. However, surface fluency is not a reliable proxy for experiential effectiveness, which still requires explicit evaluation methods in HCI practice [9].

Polished GenUI outputs also increase the risk of premature acceptance. Research on automation bias shows that people can over-rely on confident automated outputs and reduce verification effort in human–AI collaboration [3]. To counter this, GenUI workflows should separate polishing from validation and introduce explicit evaluation gates, such as heuristic evaluation and structured walkthroughs of critical interaction flows [9]. Each gate should surface a small, explicit set of risks and decisions, such as what interaction assumptions remain unproven, where breakdowns are most likely, and what must change before further commitment.

HCI practitioners should develop a repeatable "prompt to polish" process that makes human judgment explicit after each meaningful prompt round. These checkpoints reintroduce deliberate verification and align rapid GenUI iteration with established HCI evaluation practice.

5 CONCLUSION

GenUI is transforming how design is conceived, iterated, and delivered by compressing cycles from intent to interface while expanding the space of possible experiences. Based on our work as practitioners in enterprise-scale GenUI programs, this paper contributes (1) an applied framework that anchors GenUI iteration in human-centered research, shared design briefs, and integrated AI tooling, and (2) five actionable learnings that describe how practice shifts when UI becomes a co-created artifact. Collectively, our findings suggest that GenUI's value is not simply faster production but a reallocation of effort toward product vision and experience feel, alongside new responsibilities for creative direction, responsible use, and team psychological safety. At the same time, the work highlights practical constraints that GenUI outputs still require human precision to reach usability and quality expectations and prompting is better understood as one step within a repeatable "prompt to polish" loop that includes critique, refinement, and evaluation with sufficiently realistic prototypes.

These learnings are formative insights from early enterprise adoption and future research should validate their applicability, create evaluation methods for AI interface feel and trust, and enhance tools for accountability and collaboration in design handoffs. By articulating concrete adjustments to roles, rituals, and evaluation practices, we aim to help HCI practitioners harness GenUI while preserving human agency, rigor, and trust in emerging human and AI partnerships.

6 REFERENCES

[1] Kyungho Lee. 2025. *Towards a Working Definition of Designing Generative User Interfaces.* In Companion Publication of the 2025 ACM Designing Interactive Systems Conference (DIS '25). ACM, New York, NY, USA. https://doi.org/10.1145/3715668.3736365

[2] Remy, C., MacDonald Vermeulen, L., Frich, J., Biskjaer, M. M., and Dalsgaard, P. 2020. Evaluating Creativity Support Tools in HCI Research. In Proceedings of DIS '20. ACM. https://doi.org/10.1145/3357236.3395474

[3] Romeo, G., and Conti, D. 2026. Exploring automation bias in human–AI collaboration: A review and implications for explainable AI. AI & Society 41, 259–278. https://doi.org/10.1007/s00146-025-02422-7

[4] Gitte Lindgaard, Gary Fernandes, Cathy Dudek, and J. Brown. 2006. Attention web designers: You have 50 milliseconds to make a good first impression! Behaviour & Information Technology 25, 2 (2006), 115–126.

[5] Stephanie Houde and Charles Hill. 1997. What do prototypes prototype? In Handbook of Human-Computer Interaction (2nd ed.), Martin G. Helander, Thomas K. Landauer, and Prasad V. Prabhu (Eds.). Elsevier Science B.V., Amsterdam

[6] David G. Jansson and Steven M. Smith. 1991. Design fixation. Design Studies 12, 1 (January 1991), 3–11. https://doi.org/10.1016/0142-694X(91)90003-F

[7] Dominik Siemon. 2023. Let the computer evaluate your idea: evaluation apprehension in human-computer collaboration. Behaviour & Information Technology 42, 5 (2023), 459–477. https://doi.org/10.1080/0144929X.2021.2023638

[8] Carter Gibson and Michael D. Mumford. 2013. Evaluation, criticism, and creativity: Criticism content and effects on creative problem solving. Psychology of Aesthetics, Creativity, and the Arts 7, 4 (2013), 314–331. https://doi.org/10.1037/a0032616

[9] Jakob Nielsen. 1994. Usability Engineering. Academic Press, Boston, MA, USA.