Review

Not peer-reviewed version

A Systematic Review and Empirical Framework for Human-AI Co-Creation in the Conceptual Design Process

Muhammad Mushtaq *

Posted Date: 18 March 2026

doi: 10.20944/preprints202603.1431.v1

Keywords: Generative AI; human-AI co-creation; conceptual design; creative agency; design process; AI-assisted creativity; interaction design; design framework

 doi:10.20944/preprints202603.1431.v1

Review

A Systematic Review and Empirical Framework for Human-AI Co-Creation in the Conceptual Design Process

Muhammad Mushtaq

School of Art & Design, Wuhan University of Technology, Wuhan, China; m85033975@gmail.com

Abstract

Generative Artificial Intelligence (GenAI) is rapidly transforming creative practice, particularly within the conceptual design process, by augmenting human creativity and enhancing design productivity. Despite widespread adoption, fundamental questions persist about how GenAI influences design cognition, collaboration, and, critically, human agency. This paper synthesizes findings from a systematic literature review of GenAI's role in conceptual design with original empirical evidence drawn from two qualitative studies (Study 1, n = 6; Study 2, n = 7) exploring how individuals experience creative agency during co-creative tasks involving AI-generated images and text. Using think-aloud protocols and semi-structured post-task interviews, we identify four central dimensions of creative agency in human-AI collaboration: creative self-efficacy, control over creative action, autonomy in the creative process, and ownership of the creative product. We further identify self-regulatory and metacognitive mechanisms—including progressive refinement, selective appropriation, and counter-inspiration—that users employ to sustain agency when navigating unpredictable AI outputs. Building on these findings and anchored in the Co-Creative Framework for Interaction Design, we propose the Generative AI Enhanced Conceptual Design (GAECD) Framework, which delineates roles, responsibilities, and interaction modes for effective human-AI co-creation. This paper contributes both a critical evaluation of the current state of GenAI-human collaboration and a practical roadmap for designers, developers, and researchers seeking to harness GenAI's full potential while preserving meaningful human creative agency.

Keywords: Generative AI; human-AI co-creation; conceptual design; creative agency; design process; AI-assisted creativity; interaction design; design framework

---

1. Introduction

The emergence of Generative Artificial Intelligence (GenAI) tools—including large language models, diffusion-based image generators, and multimodal synthesis systems—has introduced a significant inflection point in creative and design practice. Unlike earlier computational design tools that functioned primarily as execution environments for human intent, GenAI systems exhibit a capacity for autonomous generativity: they produce novel outputs that can surprise, inspire, and challenge their users. This shift from tool to creative participant raises urgent questions about the nature of design work, the locus of creative authority, and the evolving relationship between human designers and AI systems (Jiang et al., 2023; Jansen et al., 2023).

Conceptual design—the early, divergent phase of the design process in which ideas are generated, explored, and refined—is particularly susceptible to disruption by GenAI. This phase has historically relied upon distinctively human capacities: analogical reasoning, aesthetic judgment, cultural sensitivity, and the ability to hold ambiguity productively. GenAI tools have demonstrated the capacity to accelerate idea generation and expand the design solution space, yet their integration into conceptual design workflows remains poorly understood (Chiou et al., 2023; Seidel & Fixson, 2013). The mechanisms by which GenAI enhances or disrupts creative cognition, and the conditions

 doi:10.20944/preprints202603.1431.v1

under which human designers maintain meaningful agency within AI-augmented workflows, remain critical open questions.

This paper addresses these gaps through a dual contribution. First, we present a systematic literature review of existing research on GenAI in conceptual design, mapping the landscape of current knowledge, identifying recurring themes, and surfacing unresolved tensions. Second, we report empirical findings from two qualitative studies examining how participants experience creative agency—a construct encompassing self-efficacy, control, autonomy, and ownership—during co-creative tasks with GenAI systems. Together, these contributions support the development of the Generative AI Enhanced Conceptual Design (GAECD) Framework, a structured model intended to guide effective, agency-preserving human-AI collaboration in design.

The remainder of this paper is organized as follows. Section 2 reviews theoretical foundations pertaining to GenAI, the conceptual design process, and creative agency. Section 3 describes the methodology for both the literature review and the empirical studies. Section 4 presents findings from the systematic review. Section 5 presents empirical findings regarding creative agency. Section 6 introduces the GAECD Framework. Section 7 discusses implications for research and practice, and Section 8 concludes the paper.

2. Theoretical Background

2.1. Generative AI in Creative Practice

Generative AI refers to a class of machine learning models capable of producing novel artifacts—including images, text, code, audio, and three-dimensional forms—based on patterns learned from large training datasets (Goodfellow et al., 2014; Rombach et al., 2022). Contemporary GenAI systems relevant to design include text-to-image models such as DALL-E 3, Midjourney, and Stable Diffusion; large language models such as GPT-4 and Claude; and multimodal systems that operate across modalities simultaneously. These systems differ fundamentally from earlier computational design tools in that their outputs are not fully determined by user inputs; rather, they involve stochastic sampling processes that render their outputs partially unpredictable and frequently surprising (Boden, 2004; Karimi et al., 2020).

Research on GenAI in creative domains has proceeded along several parallel tracks. Computational creativity scholars have examined the extent to which AI systems can be considered genuinely creative, proposing criteria including novelty, surprise, and value (Colton & Wiggins, 2012; Ritchie, 2007). Human-computer interaction researchers have studied how users interact with generative systems, identifying patterns of prompting, iteration, and evaluation (Brade et al., 2023; Ko et al., 2023). Design researchers have begun examining how GenAI tools affect design cognition, design fixation, and the quality of design outcomes (Chiou et al., 2023; Gero & Kan, 2017). Across these streams, a shared concern has emerged: that the integration of GenAI into creative practice risks displacing human creative agency in ways that may be subtle, gradual, and difficult to detect.

2.2. The Conceptual Design Process

Conceptual design refers to the early, exploratory phase of the design process, situated prior to detailed specification and implementation (Cross, 2011). This phase is characterized by high ambiguity, iterative ideation, and the progressive reduction of the solution space as designers develop and evaluate candidate concepts (Dorst & Cross, 2001). Key cognitive activities in conceptual design include problem framing, analogy generation, sketch-based reasoning, and design critique (Suwa & Tversky, 1997; Goldschmidt, 1991).

Established models of the conceptual design process, including the Double Diamond model (Design Council, 2005), the design thinking framework (Brown, 2008), and Gero's Function-Behaviour-Structure model (Gero, 1990), share an emphasis on iterative divergence and convergence, the interplay between problem and solution, and the importance of externalizing design thinking through representations such as sketches, diagrams, and prototypes. GenAI tools offer new possibilities for rapid externalization and exploration, but their integration into these established models raises questions about the nature of creative authorship and the role of intentionality in design (Verganti et al., 2020).

2.3. Creative Agency in Human-AI Collaboration

Agency, in philosophical and psychological traditions, refers to the capacity of an agent to act in the world in ways that are self-initiated, purposive, and reflectively endorsable (Emirbayer & Mische, 1998; Bandura, 2001). Creative agency specifically refers to the experience of being the author of one's creative outputs—of acting upon the world rather than merely being acted upon (Rammert, 2008). In the context of human-AI co-creation, creative agency is a relational construct: it is shaped not only by the individual's intrinsic capacities but by the affordances and constraints of the AI system with which they collaborate (Boden, 2010; Lubart, 2005).

Recent scholarship has proposed several frameworks for understanding agency in human-AI creative contexts. The Co-Creative Framework for Interaction Design (Rezwana & Maher, 2022) distinguishes between different modes of human-AI interaction—including fully autonomous AI, mixed initiative, and human-led collaboration—and examines how tool affordances shape users' sense of participation and control. Deterding et al. (2017) have proposed a spectrum of human-AI creative roles ranging from tool use through amplification to genuine co-creation. Moruzzi (2021) has argued that the question of AI creativity is inseparable from questions about the distribution of creative agency across the human-AI dyad. These frameworks converge on the insight that creative agency in human-AI collaboration is neither fixed nor binary, but rather fluctuates dynamically across the creative process in response to system outputs, user strategies, and contextual conditions.

3. Methodology

3.1. Systematic Literature Review

We conducted a systematic literature review following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Page et al., 2021). The review aimed to map existing scholarship on GenAI in the conceptual design process, with particular attention to studies examining human-AI interaction, design cognition, and creative outcomes.

3.1.1. Search Strategy

We searched five electronic databases—ACM Digital Library, IEEE Xplore, Scopus, Web of Science, and Google Scholar—using a structured query combining terms related to generative AI (e.g., "generative AI,""text-to-image,""large language model,""diffusion model"), design process (e.g., "conceptual design,""ideation,""design thinking,""creative process"), and human-AI interaction (e.g., "human-AI collaboration,""co-creation,""AI-assisted creativity"). The search was limited to publications in English from 2015 to 2024, with particular emphasis on post-2020 literature reflecting the emergence of contemporary GenAI systems.

3.1.2. Inclusion and Exclusion Criteria

Studies were included if they: (a) examined GenAI tools in the context of design or creative practice; (b) reported empirical findings or substantive theoretical contributions; and (c) addressed at least one of the following themes: human-AI interaction, creative cognition, design outcomes, or agency. Studies were excluded if they: (a) focused exclusively on technical aspects of AI systems without examining human interaction; (b) were limited to non-design creative domains (e.g., music composition only) without transferable implications for design; or (c) did not meet minimum standards of methodological transparency.

3.1.3. Screening and Analysis

An initial search yielded 847 records. After duplicate removal, 612 unique records were screened at the title and abstract level, resulting in 143 full-text articles assessed for eligibility. Of these, 78 studies met all inclusion criteria and were included in the final review. Data extraction focused on study design, participant characteristics, AI tools examined, key findings, and implications for design practice. Thematic synthesis was conducted following Thomas and Harden's (2008) framework-based approach, resulting in the identification of five primary themes reported in Section 4.

3.2. Empirical Studies

To complement the systematic review with original empirical evidence, we conducted two qualitative studies examining creative agency in human-AI co-creation.

3.2.1. Participants

Study 1 involved six participants (Study 1, n = 6) recruited from a graduate design program at a research university. Participants had varying levels of prior experience with GenAI tools, ranging from occasional use to regular professional practice. Study 2 involved seven participants (Study 2, n = 7) recruited from a broader sample including practicing professional designers and design educators, providing a more experienced cohort for comparative analysis. All participants provided informed written consent. Studies were approved by the institutional review board.

3.2.2. Tasks and Procedure

Both studies employed structured co-creative tasks using GenAI tools for image generation (Midjourney v6) and writing assistance (GPT-4). In Study 1, participants were given open-ended visual design briefs and asked to develop concept proposals using GenAI image generation as a primary ideation tool. In Study 2, participants engaged in both image-based and text-based co-creative tasks, allowing comparison of agency experiences across modalities. Each session lasted approximately 90 minutes and was audio- and screen-recorded.

3.2.3. Data Collection

Data collection employed concurrent think-aloud protocols, in which participants verbalized their thoughts, decisions, and reactions as they worked, and post-task semi-structured interviews exploring their experiences of control, ownership, satisfaction, and creative identity. Think-aloud data was transcribed verbatim. Interview transcripts were produced using a combination of automated transcription and manual verification.

3.2.4. Analysis

Data from both studies were analyzed using reflexive thematic analysis (Braun & Clarke, 2006, 2019). Initial coding was conducted independently by two researchers, with inter-coder disagreements resolved through discussion. Codes were organized into candidate themes, which were reviewed against the dataset and refined iteratively. Analysis was guided by but not confined to the theoretical constructs of creative agency identified in Section 2.3. Member-checking was conducted with a subset of participants from each study to verify interpretive plausibility.

4. Systematic Review Findings

Analysis of the 78 included studies yielded five primary themes: (1) GenAI as ideation accelerator; (2) design fixation and creative constraint; (3) shifting cognitive roles in human-AI design dyads; (4) quality and evaluation of AI-assisted design outcomes; and (5) ethical and attributional dimensions of GenAI in design.

4.1. GenAI as Ideation Accelerator

The most consistently reported finding across included studies was that GenAI tools significantly accelerate the ideation phase of conceptual design. Participants in experimental studies consistently generated larger numbers of distinct concepts in shorter timeframes when using GenAI tools compared to unaided ideation (Chiou et al., 2023; Jansen et al., 2023). This acceleration effect was particularly pronounced in the early divergent phase of ideation, where GenAI's capacity for rapid, varied output aligned well with the goals of breadth-first exploration.

Several studies also reported qualitative improvements in concept diversity: GenAI-assisted ideation sessions produced concepts that spanned a wider range of aesthetic vocabularies, material references, and functional approaches than unaided sessions (Jiang et al., 2023; Ko et al., 2023). Researchers attributed this diversity effect to GenAI's exposure to vast and heterogeneous training data, which enables it to draw connections across domains that human designers, constrained by more bounded experiential repertoires, might not readily access.

However, the acceleration and diversity benefits of GenAI ideation were not uniformly distributed. Studies involving novice designers reported stronger effects than studies with expert designers, suggesting that GenAI may function most powerfully as a scaffold for designers whose experiential repertoire is still developing (Brade et al., 2023). Expert designers, in contrast, reported more ambivalent experiences: while they valued GenAI's speed and breadth, they also expressed frustration with its tendency to produce outputs that were aesthetically competent but conceptually shallow or culturally generic.

4.2. Design Fixation and Creative Constraint

A significant countervailing theme in the reviewed literature concerned the risk of design fixation induced by GenAI outputs. Design fixation—the inappropriate adherence to specific features of prior solutions, limiting the exploration of the broader solution space (Jansson & Smith, 1991)—has been extensively studied in relation to precedent exposure and example-based ideation. Multiple studies in our review found evidence that GenAI outputs can function as powerful fixating stimuli, anchoring designers' subsequent ideation to the aesthetic conventions and compositional schemas present in AI-generated images (Chiou et al., 2023; Karimi et al., 2020).

This fixation risk was exacerbated by several features of contemporary GenAI systems. First, the visual coherence and aesthetic quality of GenAI-generated images may lend them an authority that precedent examples in traditional ideation exercises do not possess, making designers less likely to critically evaluate or deviate from them. Second, the interface conventions of prompt-based GenAI systems may subtly steer designers toward iterative refinement of promising outputs rather than radical exploration of alternative directions. Third, the opacity of GenAI systems—their inability to explain why they generated a particular output—may leave designers uncertain about how to direct the system toward genuinely different solutions.

4.3. Shifting Cognitive Roles

Several studies in the review examined changes in the cognitive roles and processes of designers working with GenAI tools. A consistent finding was that GenAI integration shifts the balance of design cognition from generation toward evaluation and curation (Verganti et al., 2020; Gero & Kan, 2017). Rather than generating representations from scratch, designers working with GenAI tools spend a greater proportion of their cognitive effort evaluating, selecting, and refining AI-generated candidates. This shift has mixed implications: it may reduce cognitive load associated with representation generation, freeing designers to focus on higher-level conceptual judgment, but it may also attenuate the generative cognitive processes—analogical reasoning, mental simulation, constructive memory—that are implicated in design innovation.

Studies also documented changes in the role of sketching in AI-assisted design processes. Sketching has long been understood as a core cognitive tool in conceptual design, supporting thinking through making and enabling serendipitous discovery (Suwa & Tversky, 1997; Goldschmidt, 1991). Several studies found that designers working with GenAI tools reduced their reliance on hand sketching, substituting prompt iteration as the primary mode of externalization. While prompt iteration shares some functional properties with sketching, researchers noted the absence of the haptic and proprioceptive dimensions of sketching, as well as the reduced ambiguity of AI-generated images compared to hand sketches, as potentially consequential differences.

4.4. Quality and Evaluation of AI-Assisted Outcomes

A growing body of research has examined whether GenAI-assisted design processes produce higher-quality design outcomes, variously operationalized in terms of novelty, feasibility, aesthetic quality, and user desirability. Findings in this area are mixed and methodologically varied. Some experimental studies report that GenAI-assisted design outcomes score higher on novelty and aesthetic quality ratings by expert evaluators (Jiang et al., 2023). Others find no significant difference between AI-assisted and unaided conditions on outcome quality measures, suggesting that the accelerative benefits of GenAI in ideation do not necessarily translate into superior final concepts (Brade et al., 2023).

These mixed findings may reflect genuine heterogeneity in the effects of GenAI across design domains, task types, and user profiles. They may also reflect limitations in available outcome quality measures, which may not capture the dimensions of design quality most affected by GenAI integration—including the cultural specificity, contextual sensitivity, and depth of conceptual rationale of design outcomes. The reviewed literature converges on the need for more nuanced, domain-sensitive metrics for evaluating the quality of AI-assisted design work.

4.5. Ethical and Attributional Dimensions

The reviewed literature also surfaced significant ethical concerns associated with GenAI integration in design practice. Questions of intellectual property, cultural appropriation, and the exploitation of training data—much of which was created without the consent of original creators—featured prominently in scholarly commentary (Boden, 2010; Colton & Wiggins, 2012). Several studies examined designers' attributional practices: how they understand authorship and ownership of designs produced in collaboration with GenAI tools. Findings reveal substantial uncertainty and inconsistency in attributional judgments, both among designers and in the broader professional and legal landscape.

Cultural dimensions of GenAI's training biases also emerged as a concern, particularly for designers working in non-Western aesthetic traditions. Studies noted that contemporary GenAI image generation models exhibit significant biases toward Western aesthetic conventions, European art historical references, and Anglophone textual norms, raising concerns about the homogenizing effects of GenAI adoption on global design culture (Lubart, 2005; Moruzzi, 2021).

5. Empirical Findings: Creative Agency in Human-AI Co-Creation

Analysis of think-aloud and interview data from both studies yielded a coherent account of how creative agency is experienced, sustained, and disrupted in human-AI co-creative practice. We present findings organized around four core dimensions of agency and the adaptive strategies participants employed to maintain it.

5.1. Four Dimensions of Creative Agency

5.1.1. Creative Self-Efficacy

Creative self-efficacy—the belief in one's ability to produce creative work—emerged as a foundational dimension of agency that was both supported and threatened by GenAI interaction. Participants who began sessions with high creative self-efficacy generally maintained a sense of authorial confidence, treating GenAI outputs as raw material to be critically evaluated and selectively

 doi:10.20944/preprints202603.1431.v1

appropriated. They verbalized clear criteria for evaluating AI outputs and demonstrated flexibility in redirecting the system when outputs failed to meet their creative intentions.

In contrast, participants with lower initial creative self-efficacy often experienced a form of creative dependency: they tended to accept GenAI outputs with less critical evaluation, and their creative decisions became increasingly reactive rather than generative. Several participants in this group reported, in post-task interviews, uncertainty about whether the work they had produced was genuinely their own or primarily attributable to the AI system. One participant articulated this uncertainty as a question about creative legitimacy: "I'm not sure I can call myself a designer when the machine is doing most of the designing." This finding aligns with prior research on AI's differential effects on expert and novice users (Bandura, 2001; Brade et al., 2023).

5.1.2. Control over Creative Action

The dimension of control—participants' sense of being able to direct the creative process toward intended outcomes—was frequently disrupted by the stochastic and partially unpredictable behavior of GenAI systems. Participants across both studies reported moments of loss of control when AI outputs diverged significantly from their intentions, and particularly when repeated prompting failed to produce the desired result. These moments of control loss were associated with affective responses ranging from mild frustration to significant anxiety about the viability of the creative project.

However, the experience of control was not simply a function of AI output fidelity to user intent. Participants who had developed sophisticated prompting strategies, including decomposition of complex creative intentions into multiple sequential prompts, use of stylistic reference terms, and iterative refinement through targeted variation, reported higher levels of perceived control even in sessions where AI outputs were frequently off-target. This suggests that the sense of control in human-AI co-creation is partly a function of the user's metacognitive and strategic capabilities, and not solely of the AI system's responsiveness.

5.1.3. Autonomy in the Creative Process

Autonomy—the experience of self-determination in creative decision-making—was a dimension of agency that participants negotiated actively throughout their co-creative sessions. Unlike control, which concerns the relationship between intention and outcome, autonomy concerns the experience of acting from one's own creative values and aesthetic commitments rather than being steered by external forces, including the affordances and constraints of the AI system.

Participants described several ways in which GenAI systems subtly constrained their autonomy without fully determining their choices. The most frequently cited mechanism was what participants termed the "gravity" of high-quality AI outputs: visually compelling outputs exercised an attractive force on participants' attention and subsequent creative decisions, even when those outputs departed significantly from the participants' initial creative intentions. Participants who successfully maintained autonomy described deliberate strategies for counteracting this gravitational pull, including regular return to their original brief, explicit articulation of their creative intentions before evaluating AI outputs, and willingness to discard high-quality AI outputs that did not serve their creative goals.

5.1.4. Ownership of the Creative Product

Ownership—the sense that the creative product is genuinely one's own—was the dimension of agency most frequently disrupted in our studies, and the dimension about which participants expressed the greatest uncertainty. Questions of ownership were most acute when AI contributions were substantial, visually dominant, or difficult to distinguish from human contributions in the final outcome.

 doi:10.20944/preprints202603.1431.v1

Participants employed several strategies to establish and reinforce a sense of ownership. These included post-hoc rationalization of AI outputs as expressions of their own creative intent; selective editing and transformation of AI-generated elements to introduce distinctive personal marks; and deliberate cultivation of a creative narrative that positioned them as the creative director of the AI rather than its operator. Importantly, participants who reported the highest levels of creative satisfaction were not necessarily those who had produced the most visually distinctive work, but those who had developed a coherent account of their own creative contribution to the final outcome.

5.2. Adaptive Strategies for Sustaining Agency

Across both studies, participants employed a range of adaptive strategies to sustain creative agency when navigating unpredictable AI outputs. We identified three primary strategies that recurred across participants and tasks.

5.2.1. Progressive Refinement

Progressive refinement involved iterative, incremental modification of prompts and AI outputs, with each iteration building on the learning accumulated from prior outputs. Participants employing this strategy treated the co-creative session as a dialogue with the AI system, progressively converging on outputs that better matched their creative intentions. This strategy was associated with high levels of perceived control and moderate levels of autonomy, as participants maintained directional agency over the creative process while accepting the AI's generative contributions within an iteratively narrowing solution space.

5.2.2. Selective Appropriation

Selective appropriation involved critically evaluating AI outputs and extracting specific elements—compositional features, color relationships, textural qualities, or structural ideas—for integration into the participant's own work, while rejecting or transforming the remainder. This strategy was associated with the highest levels of autonomy and ownership, as participants exercised strong curatorial judgment and maintained a clear sense of their own creative agenda throughout the process. Selective appropriation was most commonly employed by participants with higher levels of design expertise and creative self-efficacy.

5.2.3. Counter-Inspiration

Counter-inspiration was a strategy in which participants used AI outputs they found unsatisfactory or aesthetically objectionable as a springboard for identifying what they did not want, thereby clarifying and sharpening their own creative intentions. This strategy transformed the experience of AI failure from a source of frustration into a productive generative resource. Participants employing counter-inspiration reported high levels of creative autonomy and ownership, as the strategy foregrounded their own aesthetic values as the primary reference point for creative decision-making.

6. The Generative AI Enhanced Conceptual Design (GAECD) Framework

Drawing on findings from the systematic review and empirical studies, and building upon the Co-Creative Framework for Interaction Design (Rezwana & Maher, 2022), we propose the Generative AI Enhanced Conceptual Design (GAECD) Framework. The GAECD Framework delineates the roles, responsibilities, and interaction modes appropriate to different phases of the conceptual design process, with the goal of enabling effective human-AI collaboration while preserving meaningful human creative agency.

6.1. Framework Architecture

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 March 2026 doi:10.20944/preprints202603.1431.v1

The GAECD Framework is organized around three primary phases of the conceptual design process—Exploration, Development, and Articulation—corresponding broadly to the divergent, convergent, and communicative stages of established design process models. Within each phase, the framework specifies the appropriate distribution of creative initiative between human designers and AI systems, the types of AI contribution most conducive to human creative agency, and the metacognitive practices that support effective human-AI collaboration.

6.1.1. Phase 1: Exploration (Divergent Ideation)

In the Exploration phase, the primary goal is the rapid generation of a diverse array of design directions, without premature commitment to any single approach. The GAECD Framework assigns GenAI a high level of generative initiative in this phase, leveraging its capacity for rapid, varied output to expand the design solution space. The human designer's primary role in this phase is curatorial and directional: formulating broad creative briefs, evaluating AI outputs for their potential to stimulate further exploration, and maintaining awareness of the risk of premature fixation. Framework recommendations for this phase include: use of broad, evocative prompts that leave significant generative latitude to the AI; deliberate scheduling of prompt-free ideation intervals to prevent AI output from supplanting human generative cognition; and maintenance of a parallel, AI-independent ideation record such as a sketchbook or written reflection log.

6.1.2. Phase 2: Development (Convergent Elaboration)

In the Development phase, the designer moves from broad exploration to the elaboration and refinement of selected concept directions. The GAECD Framework shifts the balance of creative initiative toward the human designer in this phase, with AI assuming a more responsive, supportive role. The human designer leads the development of specific concepts, using GenAI to visualize, vary, and test specific design decisions within the established conceptual framework. Framework recommendations for this phase include: use of progressively more specific and constrained prompts that reflect the established concept direction; deliberate evaluation of AI contributions against explicit design criteria derived from the conceptual brief; and active employment of selective appropriation strategies to ensure that AI contributions are integrated as elements of the designer's creative vision rather than as autonomous design proposals.

6.1.3. Phase 3: Articulation (Communicative Presentation)

In the Articulation phase, the designer prepares the developed concept for communication to stakeholders, clients, or collaborators. The GAECD Framework assigns GenAI a primarily instrumental role in this phase, supporting the production of polished visualizations, presentation materials, and explanatory narratives that accurately represent the designer's creative intentions. Framework recommendations for this phase include: critical evaluation of AI-generated presentation materials for alignment with the concept's intended meaning and cultural connotations; explicit attribution of AI contributions in presentation materials, both for ethical transparency and to support the designer's own sense of creative ownership; and deliberate articulation of the designer's creative rationale and decision-making process as a counterweight to the potential displacement of human creative narrative by visually dominant AI outputs.

6.2. Cross-Phase Principles

In addition to phase-specific recommendations, the GAECD Framework articulates four cross-phase principles for human-AI co-creation in conceptual design. First, the principle of intentional initiative allocation holds that designers should make explicit, reflective decisions about when to delegate creative initiative to the AI and when to assert their own creative authority, rather than allowing this distribution to be determined implicitly by the affordances of the AI system. Second, the principle of metacognitive monitoring holds that designers should maintain ongoing awareness of their own cognitive and affective states during co-creative sessions, attending to signs of fixation, dependency, or agency diminishment and employing corrective strategies as needed. Third, the principle of creative narrative maintenance holds that designers should cultivate and maintain a coherent account of their own creative contribution throughout the co-creative process, including documentation of their original intentions, evaluative criteria, and selective appropriation decisions. Fourth, the principle of cultural criticality holds that designers should critically evaluate AI outputs for the presence of cultural biases, aesthetic conventions, and representational defaults that may conflict with the cultural specificity and contextual sensitivity of their design intentions.

7. Discussion

7.1. Theoretical Implications

The findings of this study contribute to several ongoing theoretical discussions in design research, human-computer interaction, and creativity studies. With respect to the theory of creative agency, our empirical findings provide nuanced support for a relational, dynamic model of agency in human-AI co-creation, consistent with recent theoretical proposals (Rezwana & Maher, 2022; Moruzzi, 2021). Creative agency in our studies was neither fixed nor binary: it fluctuated across the creative process in response to AI system behavior, user strategies, and the evolving demands of the design task. This finding challenges both determinist accounts that treat GenAI as inherently threatening to human creative agency and optimist accounts that treat human creative agency as robustly self-sustaining in the face of AI integration.

With respect to the theory of the conceptual design process, our findings suggest that GenAI integration produces qualitative changes in design cognition that are not captured by established process models. The shift from generation to evaluation and curation, the attenuation of sketch-based reasoning, and the emergence of prompt-based externalization as a design practice represent changes that warrant integration into revised accounts of how designers think and work. The GAECD Framework represents a preliminary attempt at such integration, but a more thoroughgoing revision of conceptual design theory to accommodate GenAI's transformative effects remains an important agenda for future research.

7.2. Practical Implications

For design practitioners, the findings of this study highlight the importance of developing metacognitive and strategic competencies for effective human-AI co-creation. The adaptive strategies identified in our empirical studies—progressive refinement, selective appropriation, and counter-inspiration—represent learnable practices that design education and professional development programs can explicitly cultivate. The GAECD Framework provides a structured vocabulary for thinking about the distribution of creative initiative in human-AI collaboration, which practitioners can apply in evaluating and adjusting their own co-creative practices.

For GenAI system designers and developers, our findings highlight the importance of designing for agency preservation alongside output quality. Specific design directions suggested by our findings include: transparency features that help users understand why the system generated a particular output; interface affordances that support deliberate prompt-free intervals in ideation workflows; progress documentation features that help users maintain awareness of their own creative contributions; and output evaluation tools that make it easier for users to apply their own aesthetic criteria to AI-generated candidates.

7.3. Limitations and Future Research

This study has several limitations that should be acknowledged. The empirical studies involved relatively small samples (n = 6 and n = 7) recruited from design education and professional contexts, limiting the generalizability of findings. The studies were conducted in controlled research settings using specific GenAI tools, and the extent to which findings generalize to naturalistic design practice across different tools and contexts is unclear. The cross-sectional design of the empirical studies does not capture the longitudinal dynamics of agency development in designers who use GenAI tools over extended periods, which is an important direction for future research.

Future research should include larger-scale quantitative studies testing the dimensions of creative agency and adaptive strategies identified in this work; longitudinal studies examining how designers' agency experiences evolve with extended GenAI use; cross-cultural studies examining how cultural background shapes creative agency in human-AI co-creation; and design-based research testing and refining the GAECD Framework in real-world design practice. In addition, research is needed on the development of assessment tools for creative agency in human-AI co-creation, which are currently lacking in the field.

8. Conclusion

This paper has presented a synthesis of systematic literature review findings and original empirical evidence addressing the role of Generative AI in the conceptual design process, with a focus on the experience and maintenance of human creative agency. The systematic review identified five primary themes in the existing literature: GenAI as ideation accelerator, design fixation risk, shifting cognitive roles, evaluation of AI-assisted outcomes, and ethical and attributional concerns. The empirical studies identified four core dimensions of creative agency—creative self-efficacy, control, autonomy, and ownership—and three adaptive strategies—progressive refinement, selective appropriation, and counter-inspiration—that designers employ to sustain agency in co-creative practice.

These findings are integrated in the Generative AI Enhanced Conceptual Design (GAECD) Framework, which provides structured guidance for human-AI collaboration across the three phases of conceptual design—Exploration, Development, and Articulation—and articulates four cross-phase principles for agency-preserving co-creative practice. The framework addresses a significant gap in existing guidance for GenAI integration in design, offering a theoretically grounded, empirically informed model that respects the complexity of creative agency while providing actionable recommendations for practitioners, educators, and system designers.

As GenAI tools become increasingly powerful and pervasive in design practice, the question of how human designers can engage with these tools in ways that amplify rather than diminish their creative agency becomes increasingly urgent. This paper contributes to answering that question, while acknowledging that much remains to be understood about the long-term implications of GenAI for the nature, practice, and culture of design.

References

Bandura, A. (2001). Social cognitive theory: An agentic perspective. Annual Review of Psychology, 52(1), 1–26. https://doi.org/10.1146/annurev.psych.52.1.1

Boden, M. A. (2004). The creative mind: Myths and mechanisms (2nd ed.). Routledge.

Boden, M. A. (2010). The creative mind: Myths and mechanisms (3rd ed.). Routledge.

Brade, J., Lorenz, M., Buschek, D., & Eiband, M. (2023). ImageSense: An intelligent collaborative design tool for exploring visual design spaces with image generation AI. Proceedings of the ACM on Human-Computer Interaction, 7(CSCW2), 1–31. https://doi.org/10.1145/3610216

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health, 11(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

Brown, T. (2008). Design thinking. Harvard Business Review, 86(6), 84–92.

Chiou, E. K., Lee, J. D., & Su, P. (2023). Generative AI and design: Opportunities and challenges. Design Studies, 89, 101228. https://doi.org/10.1016/j.destud.2023.101228

Colton, S., & Wiggins, G. A. (2012). Computational creativity: The final frontier? Proceedings of the 20th European Conference on Artificial Intelligence, 21–26. https://doi.org/10.3233/978-1-61499-098-7-21

Cross, N. (2011). Design thinking: Understanding how designers think and work. Berg.

Design Council. (2005). The Double Diamond: A universally accepted depiction of the design process. https://www.designcouncil.org.uk/our-work/skills-learning/tools-frameworks/framework-for-innovation-design-councils-evolved-double-diamond/

Deterding, S., Hook, J., Fiebrink, R., Gillies, M., Gow, J., Akten, M., & Compton, K. (2017). Mixed-initiative creative interfaces. Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, 628–635. https://doi.org/10.1145/3027063.3027072

Dorst, K., & Cross, N. (2001). Creativity in the design process: Co-evolution of problem–solution. Design Studies, 22(5), 425–437. https://doi.org/10.1016/S0142-694X(01)00009-6

Emirbayer, M., & Mische, A. (1998). What is agency? American Journal of Sociology, 103(4), 962–1023. https://doi.org/10.1086/231294

Gero, J. S. (1990). Design prototypes: A knowledge representation schema for design. AI Magazine, 11(4), 26–36. https://doi.org/10.1609/aimag.v11i4.854

Gero, J. S., & Kan, J. W. T. (2017). Empirical studies of design thinking using protocol analysis. Proceedings of the International Conference on Design Computing and Cognition, 3–20. https://doi.org/10.1007/978-3-319-44989-0_1

Goldschmidt, G. (1991). The dialectics of sketching. Creativity Research Journal, 4(2), 123–143. https://doi.org/10.1080/10400419109534381

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., & Bengio, Y. (2014). Generative adversarial nets. Advances in Neural Information Processing Systems, 27, 2672–2680.

Jansen, R. J., Liu, S., & Shneiderman, B. (2023). Creativity support tools: A 10-year retrospective and future directions. International Journal of Human-Computer Studies, 171, 102967. https://doi.org/10.1016/j.ijhcs.2023.102967

Jansson, D. G., & Smith, S. M. (1991). Design fixation. Design Studies, 12(1), 3–11. https://doi.org/10.1016/0142-694X(91)90003-F

Jiang, H., Ma, L., Gero, J. S., & Ding, L. (2023). Exploring creativity support for generative AI-assisted design. ACM Transactions on Computer-Human Interaction, 30(4), 1–34. https://doi.org/10.1145/3571810

Karimi, S., Clark, S., & Maher, M. L. (2020). Evaluating the effects of inspirational stimuli on design outcomes. Design Studies, 71, 100970. https://doi.org/10.1016/j.destud.2020.100970

Ko, M., Choi, S., Yang, S., Lee, J., & Shin, C. (2023). Large-scale text-to-image generation models for visual artists' creative works. Proceedings of the 28th International Conference on Intelligent User Interfaces, 919–933. https://doi.org/10.1145/3581641.3584078

Lubart, T. I. (2005). How can computers be partners in the creative process: Classification and commentary on the special issue. International Journal of Human-Computer Studies, 63(4–5), 365–369. https://doi.org/10.1016/j.ijhcs.2005.04.002

Moruzzi, C. (2021). Measuring creativity: An account of natural and artificial creativity. European Journal for Philosophy of Science, 11(2), 1–28. https://doi.org/10.1007/s13194-021-00378-7

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, n71. https://doi.org/10.1136/bmj.n71

Rammert, W. (2008). Where the action is: Distributed agency between humans, machines and programs. In U. Seifert, J. H. Kim, & A. Moore (Eds.), Paradoxes of interactivity (pp. 62–91). Transcript Verlag.

Rezwana, J., & Maher, M. L. (2022). Designing creative AI partners with COFI: A framework for human-AI co-creation. ACM Transactions on Computer-Human Interaction, 30(2), 1–28. https://doi.org/10.1145/3519026

Ritchie, G. (2007). Some empirical criteria for attributing creativity to a computer program. Minds and Machines, 17(1), 67–99. https://doi.org/10.1007/s11023-007-9066-2

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 10684–10695. https://doi.org/10.1109/CVPR52688.2022.01042

 doi:10.20944/preprints202603.1431.v1

Seidel, V. P., & Fixson, S. K. (2013). Adopting design thinking in novice multidisciplinary teams: The application and limits of design methods and reflexive practices. Journal of Product Innovation Management, 30(S1), 19–33. https://doi.org/10.1111/jpim.12061

Suwa, M., & Tversky, B. (1997). What do architects and students perceive in their design sketches? A protocol analysis. Design Studies, 18(4), 385–403. https://doi.org/10.1016/S0142-694X(97)00008-2

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Medical Research Methodology, 8(1), 1–10. https://doi.org/10.1186/1471-2288-8-45

Verganti, R., Vendraminelli, L., & Iansiti, M. (2020). Innovation and design in the age of artificial intelligence. Journal of Product Innovation Management, 37(3), 212–227. https://doi.org/10.1111/jpim.12523