Proceedings of CONF-MLA 2026 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2026.LD37911

From Prompting to Direct Manipulation: Interaction Patterns and Designer Experience in Generative AI-Supported Visual Workflows

Yucheng Wang

Department of Interaction Design, ArtCenter College of Design, Pasadena, USA
ywang116@inside.artcenter.edu

Abstract. Generative artificial intelligence (AI) is now employed in many visual design programs to create images rapidly, but these images may not meet the designer's expectations at first glance. This paper studies mismatch via text prompting, parameters and presets, reference-based input, and in-canvas manipulation. The fourteen academic sources in the review include thirteen peer-reviewed publications from 2024-2026 and Shneiderman's original work on direct manipulation. The workflow stage, participant group, interaction pattern, and the five dimensions of efficiency, perceived control, cognitive load, creative agency, and output predictability were used for evidence coding. Based on this evidence, the demand for refinement reduces the apparent advantage of rapid exploration. Effort saved in generation may be recovered by reducing prompt modifications, filtering output, or re-retrieving a suitable image. More local interactions help express and observe the decisions, but they do not eliminate the uncertainty in the generated results. Evidence for the parameters and presets is relatively scarce, so all four patterns cannot be evaluated equally. This framing clarifies how interface design, rather than model capability alone, affects professional agency and supports more informed tool selection in creative practice.

Keywords: generative AI, human–AI interaction, creative workflow, designer experience, creative control

1. Introduction

1.1. Research background

Since 2022, when systems such as DALL·E 2, Midjourney, and Stable Diffusion became publicly accessible, generative artificial intelligence (AI) image systems have shifted from experimental demonstrations to tools used in visual communication, concept development, and design education. This shift marked a transition from closed research prototypes to mass-market applications, fundamentally changing the availability and design integration of generative AI. Once these products entered the commercial workflow, the problem was no longer whether a model could generate an image. Designers also had to decide how to combine generation with sketching, comparison, editing, and final production [1-4].

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.LD37911

An interface expresses the user's intent to a probabilistic model. Prompts, sliders, reference files, and masks each present a different part of the design, and the other parts are open to interpretation. A device that can only show the initial idea conveniently will be of little use for correcting the resulting image. This can help to explain the different roles of the same system: it may serve as a quick production tool for one task but be an inconvenient collaborator in another [5, 6].

1.2. Research problem and questions

Reports of faster ideation and a lighter workload are grouped with accounts of heightened anxiety, reduced editability, and weaker stylistic control [7-9]. Based on the problems that have been studied so far, most research has focused on a single prototype, profession, or course. Time saved in generation does not necessarily include the work of revising prompts, filtering outputs, or adjusting a selected direction. Differences in experience also make it harder to compare them; novice and expert designers use the help differently [10-13].

Without a unified analytical framework, practitioners cannot reliably predict which interaction mechanisms will support their specific workflow stage or design task. This study therefore addresses three research questions:

RQ1. What are the typical interaction modes of generative AI tools for visual design?

RQ2. How do the above patterns affect efficiency, perceived control, cognitive load, creative agency, and output predictability?

RQ3. Which interaction patterns are better supported by existing evidence during early ideation, and which are better supported during iterative refinement?

1.3. Research objective and significance

The four modes of interaction serve as the foundation for the current review: text prompting, parameters and presets, reference-based input, and in-canvas manipulation. The comparison is of visual and interaction designers, design students, and other creative practitioners, and covers ideation, generation, selection, and refinement. Model benchmarks, writing systems, programming tools, and copyright-only arguments were excluded. Measured task performance is not the same as participants' perceptions of speed or creativity in this scope, as these forms of evidence answer different questions.

In this way, the studies can link interface mechanisms with the actual work performed during their application. It also has a deficiency that fast image production cannot solve: although a designer may have achieved a quick output, they will be unable to anticipate changes or modify the result precisely to meet expectations. Research on creativity and professional teams selects tools according to workflow fit rather than simply claiming that generative AI boosts efficiency or creativity generally.

2. Literature review

2.1. Generative AI in visual design: an overview

Task-based accounts include concept visualization, image variation, layout support, and sketch transformation. Li and others have collected 33 studies of graphic-design paradigms [1], and Furtado and others have classified capabilities according to the design tasks they support [2]. These accounts

DOI: 10.54254/2755-2721/2026.LD37911

constitute one stream of post-2022 scholarship and view image generation as having different aims rather than being a single type of intervention.

Questions on Design Thinking, education and professional identity also form another group. Rapid prototyping and various visual stimuli have appeared alongside unresolved problems of bias, accountability and human judgment [3]. Yu's view of AI as a co-creator and design material links these problems to practice: designers face limitations in the medium and must choose which possibilities to use [6].

The above directions have been covered in a structured search for literature from 2024-2026. Fourteen sources were selected: thirteen recent peer-reviewed papers on designers' experience with generative AI tools and one foundational human–computer interaction paper that established the theoretical basis for direct manipulation [14]. Based on interaction pattern, workflow stage, participant group, and the five dimensions of experience, these selected studies were coded.

2.2. Human–AI interaction in creative workflows

Palani and Ramos have observed changes in the roles of AI for creative works during the process of idea generation and realization [10]. Wang and others have added a division based on experience: novices gained more from the initial generation, while experienced designers used assistance selectively during development and optimization [11]. Together, these accounts place human–AI co-creation within a sequence of decisions rather than a single exchange ending in an image.

Interaction studies offer specific cases of friction. Sixteen automotive designers wanted visual and multimodal information rather than only verbal descriptions [8]. Oppenlaender's three-month study generated 885 modified images through multiple attempts with modifiers and weights to show that prompt control has emerged [4]. Think-aloud work by Rafner et al. found agency shifting as participants delegated decisions and later reclaimed them through selection or editing [12]. Student research adds a related point: prompt literacy supports problem-solving when interpreting and deliberately modifying generated materials [13]. Shneiderman's four principles of direct manipulation can be employed to assess the canvas-based alternatives [14].

2.3. Gaps in current research

Jang's within-subject study shows that a favourable result in one measure cannot represent the entire experience; a lower perceived workload and higher perceived task efficiency were associated with lower ratings for editability, personal style, satisfaction, and goal achievement [9]. Jiang describes generative fill for invention and local revision in a classroom poster project [5]. Different participants, tasks, and measures restrict what each account can achieve in the environment.

The present study addresses this gap by conducting a cross-pattern, stage-sensitive analysis of thirteen recent peer-reviewed sources, supported by one foundational human–computer interaction article. The analysis makes three contributions. First, it organizes scattered evidence into four interaction patterns and five shared experience dimensions. Second, it distinguishes ideation from refinement and clarifies which patterns are better supported during each stage. Third, it provides actionable guidance for tool developers and design educators seeking to integrate generative AI into professional creative workflows.

DOI: 10.54254/2755-2721/2026.LD37911

3. Research method

3.1. Search strategy

Searches were conducted in August 2026. Google Scholar was used for broad discovery, and ACM Digital Library, ScienceDirect, SpringerLink, Taylor & Francis Online, and IEEE Xplore were employed to verify publication information. Query combinations joined "generative AI" or "text-to-image" with terms such as "visual design", "creative workflow", "designer experience", "human–AI co-creation", "prompting", "reference input", and "direct manipulation". Midjourney, DALL·E, Adobe Firefly, and Figma AI were added when product-specific results were required.

3.2. Inclusion and exclusion criteria

Eligible sources must be visual or multimodal creative works that include designers, design students, or other practitioners. Empirical studies were preferred, although rigorous syntheses and clearly specified frameworks were also eligible when they contributed categories or theory. A peer-reviewed record and DOI were needed. Model benchmarks, writing and programming studies, copyright-only discussions, marketing materials, and unsupported commentary were all deleted. Thirteen contemporary cases met these conditions. Shneiderman's 1983 paper was the only pre-2022 reference that defined the direct-manipulation concepts used in the analysis [14].

3.3. Evidence coding

The coding sheet recorded the participant group of each source, experience level, task, tool or modality, method, workflow stage, interaction pattern, and stated limitations. The results are efficiency, perceived control, cognitive load, creative agency, and output predictability. The obtained results had positive, negative, mixed, or unsupported labels, and objective indicators were recorded separately from people's perceptions. Coding also separated ideation, generation, selection, and refinement, so an advantage reported at one stage was not automatically applied to the next.

3.4. Analytical procedure

The evidence was grouped by interaction pattern and then compared across workflow stages and experience levels. Disagreement led to a revision of the task, sample, and measurement method, and thus the relationship was not retained. Direct evidence from users was given more weight than conceptual statements, and a pattern could not be concluded from a single source. The various studies were heterogeneous and thus could not be quantitatively ranked by tool.

4. Findings: interaction patterns and designer experience

4.1. Text prompting: fast exploration but indirect control

4.1.1. Efficiency in early ideation

Simple exploration does not necessarily produce more ideas. Jang's within-subject study of 34 third-year undergraduate fashion design students explicitly shows this difference: a generative AI-assisted workflow in Adobe Illustrator reduced perceived difficulty and overall workload without increasing the number of ideas [9]. Hwang and Lee found that non-professional students used generative AI to

DOI: 10.54254/2755-2721/2026.LD37911

quickly generate ideas [13], and this perceived benefit should be treated separately from an actual increase in creativity. Before a direction is selected, rapid text-based generation can supply reference materials that inform manual alternatives. Furtado and others have identified this use in exploratory concept work [2].

4.1.2. Prompt burden and loss of control

The burden shifts from manual execution to prompt maintenance rather than disappearing altogether. A designer then needs language to keep some visual choices and change others. Park et al.'s automotive designers found that verbal prompts were not suitable for spatial, emotional, and material thinking [8], and Oppenlaender documented repeated attempts with modifiers and weights to gain control [4]. These accounts help to explain why the effort moves to phrasing and screening instead of disappearing. Jang's reports of prompt anxiety and weaker personal style concern loss of direction [9], while Rafner et al.'s selective appropriation and counter-inspiration describe ways of reclaiming agency [12].

4.2. Parameters and presets: accessible but constraining

4.2.1. Reduced prompting effort

Direct evidence of reduced prompting effort is limited. Rafner and others' slider interface supported incremental adjustment, but fine-grained expression of intent was still not feasible in the small preliminary study [12]. Oppenlaender documents the positive and negative weights for combining or excluding subjects and styles [4]. Aspect ratio, style, seed, strength, and term weight can be modified without changing the whole prompt to select a model behaviour from the available options. Therefore, parameters and presets may be attractive for managing variations; however, it has not been shown that they consistently reduce the workload at all levels of design experience.

4.2.2. Platform-defined design space

A preset can fit the task closely and ease iteration, or leave the relevant visual attribute outside the available choices. Thus, the problem is how to add categories to a design brief, not whether the menus themselves limit creativity. Interviews with user experience professionals reveal how tool affordances redistribute judgment [7]. Li's and Yu's accounts of AI as a medium with distinct affordances and constraints help clarify what remains under the model's control once a preset is selected [1, 6]. Default-driven convergence is a problem to be studied, not a realised result of all preset interfaces.

4.3. Reference-based input: improved communication of intent

4.3.1. Visual and spatial communication

A sketch can specify arrangement, an image material or atmosphere, and a mask the intended location of change. They can more conveniently transmit information that would otherwise need to be expressed orally. Furtado et al. distinguish sketch-to-render transformation from text-to-image ideation [2], and Park et al.'s participants requested visual interaction suitable for professional design [8]. Jiang's classroom project shows that these purposes can be combined: students used DALL·E

DOI: 10.54254/2755-2721/2026.LD37911

for inspiration and Firefly's generative fill to extend, remove, or replace selected poster content [5]. Since the project was not a controlled experiment, it shows usage rather than multimodal superiority.

4.3.2. Interpretation and ownership risks

Jiang's students sometimes returned to conventional assets when generation failed to express their intention [5]. Returning to conventional assets removed the need for model interpretation. Image prompts can affect subject and style together, obscuring which feature influenced an output [4]. The resulting judgment will also involve authorship, as there are borrowed materials, model contributions, and the designer's choices. Related agency studies make that distinction relevant [7, 12]. Reference images help designers recall and express visual intentions; therefore, to determine whether they are suitable, it is necessary to consider what has been remembered, what has changed, and whose visual language is being employed.

4.4. In-canvas manipulation: stronger control in refinement

4.4.1. Local editing and reversibility

Shneiderman's four principles of direct manipulation—continuous object representation; physical actions or labeled buttons instead of complex syntax; rapid, incremental, reversible operations with immediately visible results; and layered learning—are good tests of in-canvas manipulation [14]. Selection, repositioning, masking, and undo can show the desired scope of a revision, allowing designers to work on one part without disrupting the whole composition. Jiang's classroom cases describe generative fill in this way [5], and Park et al. report a professional need for tools that support iteration [8]. The link to direct manipulation is the preservation and revision of accepted works, not just the placement of a generation command in a familiar canvas.

4.4.2. Probabilistic results within direct manipulation

A mask specifies the region for generation but not the pixels within it. This interaction is more direct than a text prompt but less predictable than conventional editing. Jang's ratings of editability, goal achievement, and personal style suggest that low perceived effort can coexist with an unsatisfactory outcome [9]. Rafner et al. have shown that agency changes occur through delegation, evaluation, and re-authoring [12], and a local interface must still support this evaluation. Version history, comparison, undo, and protected regions can be used to recover from a failed revision.

4.5. Cross-pattern comparison

4.5.1. Comparison by workflow stage

The comparison shows how much of the design has been approved so far. When many directions are available, text prompts can offer alternatives; references and canvas tools are more necessary if the composition needs to be modified further. Parameters are a relatively small-scale reduction of variation. Palani and Ramos's changing roles for generative AI are consistent with the division of labour by Wang et al.; novices benefited more in the initial ideation stage, and experienced designers selectively used AI for quality assurance and refinement [10, 11]. Jiang's study and Furtado et al.'s framework provide task-level examples [2, 5]. Table 1 shows these connections but does not fix an order for all projects.

DOI: 10.54254/2755-2721/2026.LD37911

Table 1. Cross-pattern comparison across workflow stages

Pattern	Primary control	Best-supported stage	Main benefit	Main limitation
Text prompting	Natural language and modifiers	Early ideation	Rapid divergence and low entry threshold	Linguistic burden and indirect control
Parameters and presets	Sliders, weights, seeds, and preset choices	Generation and variation	Accessible, repeatable tuning	Platform-defined design space
Reference-based input	Images, sketches, masks, and annotations	Concept transfer and variation	Direct communication of visual intent	Misinterpretation, fixation, and provenance
In-canvas manipulation	Local selection, masking, editing, and undo	Refinement	Local, reversible control	Probabilistic outputs and skill demand

4.5.2. Evidence strength and remaining uncertainty

Table 1 should not be interpreted as a ranking supported equally in every row. parameters and presets often appear in large-scale systems, and multimodal and canvas-based projects combine controls that are difficult to separate. Even when direct user evidence is available, the samples limit interpretation: Rafner and others conducted exploratory studies with six and seven participants, Park and others studied sixteen automotive designers [8, 12], and Jang studied 34 third-year undergraduate fashion design students in a within-subject comparison of conventional and generative AI-assisted workflows in Adobe Illustrator [9]. Conditional relationships are therefore more justifiable than universal high, medium, or low ratings. More robust comparisons would require matched tasks and shared measures of efficiency, perceived control, cognitive load, creative agency, and output predictability.

5. Design implications

5.1. Support progressive control

The transition of controls should maintain the previously made decisions. A prompt or parameter is suitable for reference and local editing when the new composition still needs to meet certain constraints and is not just a disposable starting point. This design implication is based on the stage changes shown in practitioner and classroom studies [2, 5, 10]. Controls alone cannot solve this problem; the designer still needs to determine whether the change resulted from an updated prompt, preset, reference, mask, or canvas operation. Visible records of the input would help diagnose the abnormal behaviour [4, 8].

DOI: 10.54254/2755-2721/2026.LD37911

5.2. Preserve designer agency

Undo is consequential because rejecting a generated revision should not require rebuilding the accepted work. Agency and direct-manipulation literature support inspectable changes, version comparison, and recovery of previous states [12, 14]. Protection of user-created assets and manually refined regions also falls under this category; extended regeneration would require an explicit selection, and unchanged areas remain stable. Reports of faster production coupled with a lack of stylistic ownership, satisfaction, or control are particularly relevant [7, 9]. Agency here refers to the capacity to refuse or withdraw a contribution, rather than merely choosing among outputs.

5.3. Match AI roles to workflow stages

The same output can serve as an initial prompt and is not suitable for the final modification. Practitioner and expertise-related findings suggest evaluating AI assistance against the stage of work rather than rewarding output quantity throughout a project [10, 11]. Early exploration benefits from low-cost comparisons without treating the first interpretation as the default. Local constraints and handoff to conventional editing are given more weight during refinement [5, 9]. Design educators can use this division to evaluate which generated materials students choose and modify, as well as how many choices they offer.

5.4. Reduce prompt dependency

Reducing prompt dependency relates to the expression of intent, not the omission of language. Sketches, reference images, annotations, masks, spatial constraints, and parameter controls can all be used to express relationships that are difficult to describe verbally [4, 8]. The contributions of these will be easier to evaluate if a visual history can be constructed linking each output to its inputs and local edits. Such a record would allow designers to compare branches and retrace decisions without rebuilding a hidden sequence [5, 14]. Text could be one of the channels in that record, and it would be supported by visuals.

6. Conclusion

6.1. Summary of findings

This review addressed three research questions. First, it identified text prompting, parameters and presets, reference-based input, and in-canvas manipulation as four interaction patterns (RQ1). Second, it examined their effects on efficiency, perceived control, cognitive load, creative agency, and output predictability (RQ2). Third, it distinguished the stronger support for early ideation from the more conditional support for iterative refinement (RQ3). The core tension in generative AI workflows consists of two conflicting demands: obtaining diverse alternatives during exploration and carrying out approved decisions through precise refinement. The reviewed studies support rapid exploration more clearly than precise, reliable refinement. This explains why a low-workload item may have weak control or stylistic direction: it is relatively simple to execute, but the desired outcome is difficult to achieve.

DOI: 10.54254/2755-2721/2026.LD37911

6.2. Research significance

Tool selection will therefore have to consider the cost of losing the work already won. Generation of another image can help with idea exploration, but it is also relatively expensive if the composition does not need to be changed. These priorities for interface design include continuity across generations and editing, visible constraints, inspectable changes, and recovery to facilitate the evaluation of model contributions.

6.3. Limitations and future research

This interpretation is still conditional because the literature has not provided comprehensive coverage of parameters and presets, and many times it combines multiple interaction modes in one system. Comparative studies can keep the design task constant and measure correction time and recovery effort in relation to generation speed, cognitive load, and perceived control. Longer observation is needed to determine whether short-term assistance changes professional skills or visual diversity. Until then, the evidence supports matching the interaction to the decision at hand rather than using a single best tool for all stages.

References

[1] Li, H., Xue, T., Zhang, A., Luo, X., Kong, L., & Huang, G. (2024). The application and impact of artificial intelligence technology in graphic design: A critical interpretive synthesis. *Heliyon, 10*(21), Article e40037. https://doi.org/10.1016/j.heliyon.2024.e40037

[2] Furtado, L. S., Soares, J. B., & Furtado, V. (2024). A task-oriented framework for generative AI in design. *Journal of Creativity, 34*(2), Article 100086. https://doi.org/10.1016/j.yjoc.2024.100086

[3] Sreenivasan, A., & Suresh, M. (2024). Design thinking and artificial intelligence: A systematic literature review exploring synergies. *International Journal of Innovation Studies, 8*(3), 297–312. https://doi.org/10.1016/j.ijis.2024.05.001

[4] Oppenlaender, J. (2024). A taxonomy of prompt modifiers for text-to-image generation. *Behaviour & Information Technology, 43*(15), 3763–3776. https://doi.org/10.1080/0144929X.2023.2286532

[5] Jiang, J. (2024). When generative artificial intelligence meets multimodal composition: Rethinking the composition process through an AI-assisted design project. *Computers and Composition, 74*, Article 102883. https://doi.org/10.1016/j.compcom.2024.102883

[6] Yu, W. F. (2025). AI as a co-creator and a design material: Transforming the design process. *Design Studies, 97*, Article 101303. https://doi.org/10.1016/j.destud.2025.101303

[7] Li, J., Cao, H., Lin, L., Hou, Y., Zhu, R., & El Ali, A. (2024). User experience design professionals' perceptions of generative artificial intelligence. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (pp. 1–18). Association for Computing Machinery. https://doi.org/10.1145/3613904.3642114

[8] Park, H., Eirich, J., Luckow, A., & Sedlmair, M. (2024). "We are visual thinkers, not verbal thinkers!": A thematic analysis of how professional designers use generative AI image generation tools. In *Nordic Conference on Human-Computer Interaction* (pp. 1–14). Association for Computing Machinery. https://doi.org/10.1145/3679318.3685370

[9] Jang, S. Y. (2026). From designer to curator: Cognitive and creative trade-offs in GenAI-assisted design. *Fashion and Textiles, 13*(1), Article 12. https://doi.org/10.1186/s40691-026-00459-w

[10] Palani, S., & Ramos, G. (2024). Evolving roles and workflows of creative practitioners in the age of generative AI. In *Creativity and Cognition* (pp. 170–184). Association for Computing Machinery. https://doi.org/10.1145/3635636.3656190

[11] Wang, N., Kim, H., Peng, J., & Wang, J. (2025). Exploring creativity in human–AI co-creation: A comparative study across design experience. *Frontiers in Computer Science, 7*, Article 1672735. https://doi.org/10.3389/fcomp.2025.1672735

[12] Rafner, J., Zana, B., Bang Hansen, I., Ceh, S., Sherson, J., Benedek, M., & Lebuda, I. (2025). Agency in human–AI collaboration for image generation and creative writing: Preliminary insights from think-aloud protocols. *Creativity Research Journal*. Advance online publication. https://doi.org/10.1080/10400419.2025.2587803

DOI: 10.54254/2755-2721/2026.LD37911

[13] Hwang, Y., & Lee, J. H. (2025). Exploring students' experiences and perceptions of human–AI collaboration in digital content making. *International Journal of Educational Technology in Higher Education, 22*(1), Article 44. https://doi.org/10.1186/s41239-025-00542-0
[14] Shneiderman, B. (1983). Direct manipulation: A step beyond programming languages. *Computer, 16*(8), 57–69. https://doi.org/10.1109/MC.1983.1654471