Inkspire: Supporting Design Exploration with Generative AI through Analogical Sketching

David Chuan-En Lin
Human-Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
chuanenl@cs.cmu.edu

Hyeonsu B. Kang
Human-Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
hyeonsuk@cs.cmu.edu

Nikolas Martelaro
Human-Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
nikmart@cmu.edu

Aniket Kittur
Human Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
nkittur@cs.cmu.edu

Yan-Ying Chen
Toyota Research Institute
Los Altos, California, USA
yan-ying.chen@tri.global

Matthew K. Hong
Toyota Research Institute
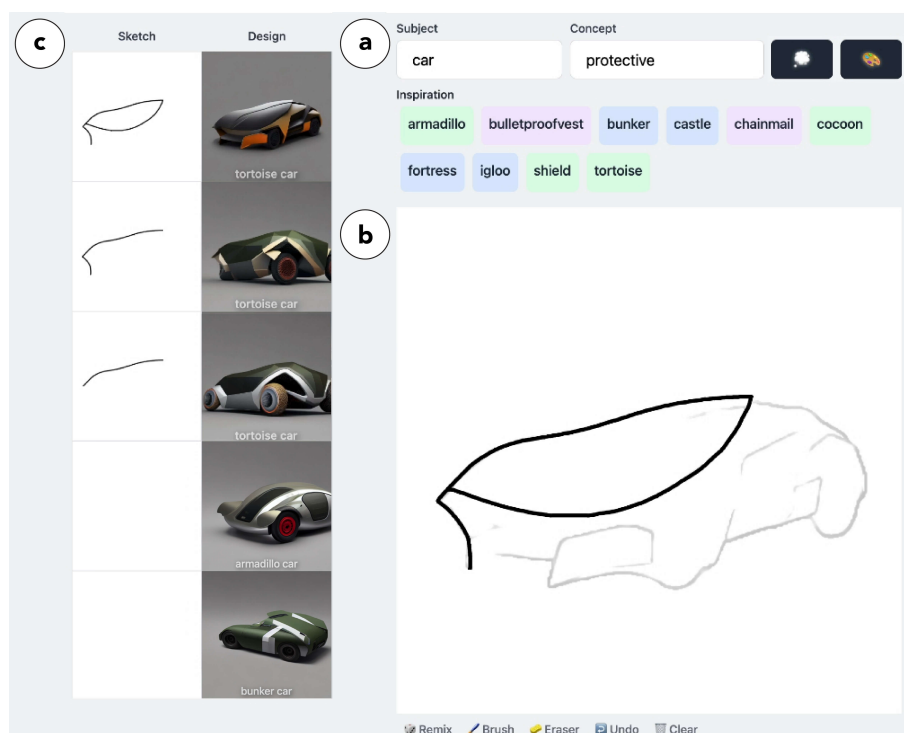Los Altos, California, USA
matt.hong@tri.global

Figure 1: The Inkspire interface. The designer may use the Analogical Panel (a) to ideate analogical inspirations for abstract concepts (e.g., "protective car" → "tortoise car"). The designer may sketch on the Sketching Panel (b) to iteratively guide AI design generations. For each iteration, we display a sketch scaffolding under the canvas. This scaffolding is created through abstracting AI designs into lower fidelity. Finally, the designer may view the history of iterations on the Evolution Panel (c).

Abstract

With recent advancements in the capabilities of Text-to-Image (T2I) AI models, product designers have begun experimenting with them in their work. However, T2I models struggle to interpret abstract language and the current user experience of T2I tools can induce design fixation rather than a more iterative, exploratory process. To

address these challenges, we developed Inkspire, a sketch-driven tool that supports designers in prototyping product design concepts with analogical inspirations and a complete sketch-to-design-to-sketch feedback loop. To inform the design of Inkspire, we conducted an exchange session with designers and distilled design goals for improving T2I interactions. In a within-subjects study comparing Inkspire to ControlNet, we found that Inkspire supported designers with more inspiration and exploration of design ideas, and improved aspects of the co-creative process by allowing designers to effectively grasp the current state of the AI to guide it towards novel design intentions.

CCS Concepts

• **Human-centered computing → Interactive systems and tools**; **Human computer interaction (HCI)**.

Keywords

generative AI, sketching, iterative design, co-creative design

1 Introduction

We have seen significant progress in the capabilities of text-to-image (T2I) models, many of which are now able to generate realistic images using text [9]. These models not only accelerate the process of converting thoughts into visuals but can also potentially create serendipitous inspirations for users [25]. Recent research has also opened up new possibilities for translating one image representation into another, such as transforming a sketched drawing into detailed designs [72]. Consequently, many designers have begun embracing the use of T2I models to enhance their creative work.

However, despite the proposed benefits T2I models, integrating them into the designer's creative workflow can be challenging. In particular, recent works have observed that designers experience *a high level of fixation* [35] when using generative AI [63] leading them to explore fewer novel ideas than may otherwise be beneficial for innovation. For instance, a designer might write a text prompt and hit generate. Upon viewing the result, they might adjust a few words in their prompt and hit generate again [20]—repeating this process akin to using a slot machine [2]. Their final prompt tends to be conceptually similar to their original one, often with small, incremental modifications aimed at getting the AI to create their *initial design intention*, instead of exploring new diverse design spaces [63]. As found in prior work, unsupported text-only prompting can be a limiting interface for generating good outputs from GenAI [70].

To understand how professional designers might experience these challenges with GenAI in their own processes, we conducted a comprehensive day-long exchange session with a team of professional designers from a large automotive company. The team

outlined their design process, from conceptualizing a product design sketch based on a specific guiding theme to presenting final designs to stakeholders. From our discussions with the designers, we identified three key challenges in their use of GenAI:

C1 Designing through text prompts feels *unnatural* as compared to their traditional sketching and ideation process.

C2 T2I models struggle with generating inspiring designs from *abstract concepts* (e.g., "protective" car), a technique the team uses to create more novel designs.

C3 It is difficult to directly build on generated designs—they appear *"too complete"*, potentially leading to fixation.

To help designers working with GenAI avoid design fixation, we might take inspiration from design research working to break designer fixation. Two common strategies in line with what we learned from the professional design team include analogy-driven design [30, 36, 53], where concepts from outside the domain area are used to foster inspiration in the domain area (addressing C2) and providing multimodal interfaces that work visually and with lower-fidelity assets (addressing C1 & C3).

Several recent works in HCI have focused on developing new interfaces for GenAI to help designers explore a larger design space, including interactive prompting which aims help people who often have limited prompting ability [11], multimodal search [61], and visually navigating a 2D latent space interface [24]. Nonetheless, being able to help designers move beyond one thread of thinking and explore a wider design space while still taking inspiration and *building on* generated designs remains a significant limitation preventing GenAI to be fully integrated into designers' work. While the ML community offers methods to increase editing freedom of generated images, including local inpainting [57] and instruction-based image editing [12], these methods remove designers away from their natural rhythm of interaction with ideas.

In this paper, we introduce a workflow of continuous exploration with T2I models, encouraging designers to adopt a mindset that facilitates more iterative and exploratory design generation. Specifically, we built Inkspire, a proof-of-concept tool that provides a more familiar interaction built around iterative sketching and by leveraging the concept of analogical design to facilitate inspiration around design ideas. Our tool integrates analogical inspiration to promote concept-level ideation from abstract themes, allowing designers to recognize creative possibilities without needing to come up with them and write prompts manually. This reduces cognitive friction, enabling fluid exploration of new ideas [14]. Additionally, inspired by prior works on drawing assistance systems that convert photographs into sketches to teach people how to draw [46], we introduce a new mechanism that converts high-fidelity AI designs into high-quality but low-resolution sketch scaffolds, directly underlaid on the designer's canvas. The scaffolding provides a transparent view into the current state of the AI and helps designers build on AI designs without being overly fixated on photorealistic renders. Finally, we enable a new design generation every time a new pen stroke is drawn (with near real-time performance), encouraging designers to consider small but meaningful changes in form and refinement with each pen stroke, thereby creating many more opportunities for designers to explore new directions. All of

these components are designed to be seamlessly integrated into the sketch process familiar to designers.

To understand how designers use Inkspire, we invited both professional designers and novice users to design everyday products. We also asked them to use a baseline condition of a state-of-the-art ControlNet [72], which allows designers to use sketches and user-written prompts to guide generative design, but does not provide our proposed analogy inspirations or convert designs into lower resolution sketches, and requires the user to choose when they generate explicitly. The results show that users rated Inkspire as providing significantly more inspiration and exploration over baseline ControlNet. The interaction that users had with Inkspire was drastically different than using a baseline ControlNet—designers using Inkspire generated many more concept sketches and showed a process that appears more co-creative, whereas when using ControlNet they focused on manual sketching and refining prompts before handing off to the generative system. Inkspire enabled designers to more effectively co-create designs with T2I models with significantly increased partnership, controllability, communication, and sense of attribution over their creations. This research thus contributes:

- **Inkspire**, a proof-of-concept, iterative sketching tool that helps ground abstract concepts into analogies and converts AI-generated images into abstracted sketch scaffolds to support fluid design iteration and avoid fixation.
- A **within-subjects study** showing that Inkspire embodies a more iterative and exploratory workflow with T2I models, with designers rating Inkspire with significantly higher inspiration, exploration, and attributes of co-creation.

2 Related Work

This work builds on prior research in human-AI co-creativity for helping people generate ideas [52]. Specifically, we build upon prior works aiming to help designers create new visual concepts through text to image models and visual inputs [32, 51, 54]. We review how generative AI systems have been shown to increase idea generation, but counterintuitively can lead to more design fixation. We then review possible solutions for overcoming design fixation including analogy-driven design, computational sketching tools, and reducing the fidelity of generated images.

2.1 Generative AI and Design Fixation

One of the proposed benefits of generative AI is to help designers avoid *design fixation*, where designers remain stuck in a single line of thinking and a limited set of ideas, thus limiting the conceptual novelty of ideas and potential for innovation [35]. While many of the works suggest that generative AI can help people generate more ideas, potentially helping them to move into new conceptual spaces, recent experimental work has found contradictory results on generative AI's impact on design fixation. DesignAID [16] leverages large language models and image generators to help people explore a large, diverse space of ideas and was found to provide more inspiration than search-based tools. However, people rated the generative AI ideas as less valuable, with the paper authors suggesting that the ideas may not have been diverse enough, not well matched to the problem, or just not enough to break people

design fixation. Bordas et al. [10] find that people using ChatGPT 3.5 to help generate ideas to protect and egg falling from 10 meters increased their idea generation but remained overly fixated on specific solutions. Wadinambiarachchi et al. [63] find that participants using a text-to-image generator to create a new chatbot avatar concept had significantly higher design fixation with fewer ideas, less variety, and less originality when using a text-to-image generator as compared to using an image search engine or coming up with ideas unassisted. The authors suggest that people's limited prompting ability, a known issue when lay people try to use LLMs [70], led them to simply copy keywords from the design brief, limiting the language used to generate the images. Davis et al. [24] and Zhang et al. [71] report similar findings, where designers working with text-only image generators showed limited creative exploration, again due to people's limited abilities with text prompting.

To overcome the fixating issues of text-based generative systems, Davis et al. [24] developed a generative system taking example images of dresses and provided graphical buttons and sliders to generate more realistic vs. creative ideas and alter shape, color and texture. This visual generative system was preferred by designers and lead to more creative idea explorations. However, the high-fidelity images presented by many image generation tools, which Wadinambiarachchi et al. [63] suggest could lead to more fixation based on prior design research showing that high-fidelity images lead to more fixation over rougher sketches [18, 21].

In our work, we explore how to overcome the fixating issues that many generative AI systems have today. First, we look to help designers explore more novel ideas by providing interfaces that overcome their limited prompting abilities and help them generate more novel ideas. For this, we look to analogy-driven design as a potential solution. Second, we move away from text-only generation and provide more visual interfaces for design exploration (C1), helping better match how designers come up with ideas. We explore how iterative, computational sketching can be used as an visual input to image generation. Third, we break from showing only *high-fidelity* images which could lead to design fixation and explore how AI generated images can be altered and presented to designers in *lower resolution* forms to see if such representations may scaffold new ideas.

2.2 Analogy-Driven Design

Analogy-driven design, or design-by-analogy, is an approach to drawing inspiration from a known domain, including concepts and products, to find novel solutions to a target domain. An abundance of text-based research systems and prototypes such as DANE [31], Idea-Inspire 4.0 [60], and BioTRIZ [62] have been developed to support design ideation by offering different design-by-analogy capabilities that include retrieval and mapping of analogies to a target problem based on inferred similarities.

Recent proliferation of text and image data repositories combined with advances in vision and language models gave rise to new multimodal approaches that expand analogy-driven design to the visual domain [36]. For instance, Kwon et al. developed an approach that leverages visual similarity to discover visual analogies for generating new ideas [43]. In addition, Zhang and Jin proposed an unsupervised deep learning model, Sketch-pix2seq, to extract

shape features from Quickdraw sketches, creating a latent space that enables defining visual similarities and searching for analogical sketches [74]. Jiang et al. developed a CNN-based model to create feature vectors representing patent images, which combine visual and technological information to enhance visual stimuli retrieval [37]. While these models present promising methods to support image-based analogy-driven design, the specialized nature of these models reduces their practical appeal to designers seeking exposure to out-of-distribution inspirations.

Large pre-trained models provide exciting opportunities to support domain-agnostic analogy based image retrieval. While T2I models alone struggle to generate images from an abstract concept such as "mystery" [40], machine learning research has demonstrated the use of LLMs to convert abstract concepts into semantically meaningful physical representations, thereby streamlining the process for downstream T2I generation tasks [28, 49, 66]. For instance, Fan et al. introduced an approach that extends an abstract concept such as peace with concrete objects (e.g., white doves, olive branches), then rewriting the original prompt to incorporate the objects in a scene [28]. Liao et al. proposed the Text-to-Image generation for Abstract Concepts (TIAC) framework that builds on a three-layer artwork theory to clarify the *intent* of the abstract concept with a detailed definition, then transforming it into semantically-related physical *objects* and concept-dependent *form* [49]. However, much of this work focused on scene illustration with multiple objects. Moreover, techniques that rely on automated prompt enrichment reduce the steerability of T2I models. In Inkspire, we apply a similar process through the use of LLMs and and analogical reasoning to convert an abstract concept into individual physical objects from across multiple domains. We investigate how providing users with a menu of analogical concepts can help the user rapidly generate diverse analogical-grounded designs with T2I models, to overcome the current challenges of creating inspiring designs from abstract design prompts (C2) in ways that also offer them more control over design space exploration.

2.3 Guided Sketching

Research has explored various computational techniques that provide sketch guidance to aid in several use cases including skill building, serving as reference points, and encouraging creative exploration.

Many existing systems have explored guided sketching tools that help novices learn how to sketch. ShadowDraw [46] offers real-time shadow-based feedback, while systems such as The Drawing Assistant [34] and Painting with Bob [8] focus on translating photographs to sketches. Several works have explored crowdsourcing-based approaches. Systems like Limpaecher et al. [50] and Sketchy [58] leverage collective human knowledge to provide guidance. These systems excel at teaching specific drawing techniques but are less suited for open-ended creative exploration. Similarly, portrait-specific systems such as DualFace [33], which employs a two-stage drawing guidance for freehand portraits, and PortraitSketch [67], which provides face sketching assistance, demonstrate the value of domain-specific guidance but are constrained to a narrow use case.

Most relevant to Inkspire is Creative Sketching Partner [23], which retrieves sketches that are visually and conceptually similar to the user's sketches as a means to stimulate exploration and inspire new designs. This system demonstrates the potential of computational guided sketching systems to inspire new designs. However, it has the limitation of relying on existing sketch databases, whereas in this work, we leverage the generative capabilities of AI.

In our work, we draw inspiration from empirical research by Williford et al. [65], whose analysis of 240 concept sketches revealed that ambiguous sketch underlays could reduce fixation on conventional forms and promote divergent thinking, guided sketching techniques such as ShadowDraw [46], and existing design practices of sketch scaffolding [19]. With Inkspire, we propose a novel computational pipeline for converting GenAI designs into abstracted, yet high-quality sketching scaffolds that underlay the user's canvas. To the best of our knowledge, this work is the first to support the full closed-loop-cycle of sketching GenAI designs and abstracting GenAI designs into sketches. Through this technique, we address the challenge of building on GenAI images that are "too complete" (C3) by drawing designers' attention away from the high-fidelity generated image and inspire designers to iterate on top of the silhouette of GenAI designs with low friction.

3 Formative Session with Design Professionals

To understand how professional designers use Text-to-Image (T2I) models in their work, we conducted a day-long exchange session with a team of seven professional product designers from a large automotive company. The designers in our exchange session work at a top 5 automotive manufacturing company and cover multiple disciplines of training spanning a wide range of roles in the company, including creative director, modeling lead, interior designer, exterior designer, artistic creator, conceptual lead creator, UX/UI and strategy. There is significant cross-collaboration across departments globally and with other companies in the industry. The designers also have experience in using T2I tools such as Midjourney [4] and Vizcom [5]. The designers showed our team their design processes in presentations with specific examples from their past work. This included sketches, concept boards, and various documentations of collaborative meetings for a wide range of mobility concepts. The designers also showed their process of using current T2I tools, and we discussed new ideas on how to design tools to support them. From our interaction with the designers, we identified key challenges they face when using T2I models and summarized them into three design goals to inform the development of Inkspire.

3.1 Design Goals

3.1.1 Design Goal 1. Sketching as a Natural Method of Interaction. Designers emphasized that prompting is an unnatural approach to designing. They expressed difficulty in effectively conveying design ideas through language [70]. They often felt constrained by the need to craft comprehensive prompts, which limited their ability to explore a wider range of ideas. In contrast, designers expressed a preference for approaching design tasks through *sketching* [15], often beginning with just a simple line or silhouette. Therefore, our first design goal is to allow designers to interact with AI via sketching as a natural method of interaction. We aim to support

designers in starting with simple abstract lines and assist them in progressing towards complete sketches.

3.1.2 Design Goal 2. Visually-Concrete Inspirations. Designers mentioned that design briefs are typically inherently abstract, for example, "design a vehicle that conveys a sense of protectiveness". However, they found T2I models to generally produce poor and generic results when prompted with such abstract terms [69]. Even when resorting to prompt engineering tricks, they find it challenging to visualize abstract concepts in concrete forms. Therefore, our second design goal is to assist designers in visualizing abstract design themes through visually-concrete inspirations. Inspired by the way designers may draw inspiration from nature [26], we aim to recommend analogical inspirations [39] to designers and make it easy for them to quickly visualize a variety of diverse inspirations.

3.1.3 Design Goal 3. Complete the Feedback Loop. Designers expressed difficulty in iterating on AI-generated designs and they described the process of using T2I models as a one-way process. They feel that the generated designs look "too complete", making it difficult for them to envision new ways to build on them. Often, designers find themselves in a position where they can either choose to use a design or discard it entirely. Therefore, we aim to bridge the gap in the feedback loop – while T2I models transform ideas into images, our goal is to *transform images back into abstractions (i.e., sketches), to allow designers to continue the iteration process.*

4 Inkspire

We first illustrate how a user would use Inkspire through a concept car design example. We then describe the technical implementation of Inkspire, which consists of two primary components: Sketch2Design (generating AI designs from sketches and analogies) and Design2Sketch (converting AI designs into lower fidelity sketch scaffolding).

4.1 System Walkthrough

DeLorean is an automotive designer tasked with creating a concept car design that embodies a sense of "protectiveness" (Figure 1).

4.1.1 Ideating a Design Concept. To begin, DeLorean uses the *Analogy Panel* (Figure 1a) to generate visually-concrete inspirations for the abstract concept (Design Goal 2). He uses "car" as the subject and "protective" as the abstract concept, then clicks on the inspiration button. He is presented with a selection of inspirations, color-coded by categories: nature , fashion , and architecture . DeLorean experiments with several inspirations (such as bunker, armadillo, tortoise) by clicking on them. For each inspiration he selects, the AI generates a design in the *Evolution Panel* (Figure 1c). He can manually edit inspirations in the concept box and manually click on the generate button. DeLorean decides to use tortoise as his inspiration.

4.1.2 Iterating on Designs through Sketching. DeLorean now iteratively guides the AI through sketching in the *Sketching Panel* (Figure 1b) (Design Goal 1). He starts off with a simple silhouette line, and the AI generates an initial design in the Evolution Panel. In the Sketching Panel, DeLorean sees a *scaffolding* abstracted from the initial AI design (Design Goal 3). This allows him to take inspiration

from the AI designs without being overly fixated on photorealistic renders. He is drawn to the bold curve of the windshield area shown in the scaffolding and loosely traces this part to add to his sketch. The AI then generates a new design and scaffolding.

DeLorean repeats this back-and-forth process with the AI (Design Goal 1), continuing to iterate on his designs through sketching until he achieves a design that satisfies him. To navigate between iterations, restart a sketch, or refine parts of the sketch, he uses the *Undo*, *Clear*, and *Eraser* tools. To explore design variations using the same input (the same inspiration and sketches), he uses the *Remix* tool.

4.2 Sketch2Design

The *Sketch2Design* component helps users brainstorm design concepts and generate product designs through sketching (Figure 2). First, the user specifies the subject that they are designing for (e.g., car) and an initial abstract concept (e.g., protective). To brainstorm more concrete design ideas for the abstract concept (Design Goal 2), we leverage Large Language Models (LLMs) (GPT-4) [13] to generate analogical inspirations (Figure 2a). We leverage prior techniques in chain-of-thought reasoning [64] to break down the problem of creating analogies based on the given concept word. We take a two-step prompting approach shown in the listings below. We first prompt the LLM to detail the design principles for the given subject (e.g., car design).

```
Describe the key design principles in <subject> design in
one short paragraph.
```

Such design principles can be useful as context [47] to ground the LLM to ideate inspirations that are more suitable for the specific product domain. An example intermediary result from this step for the domain of car design could be

```
Key design principles for car design include aerodynamics
exteriors for fuel efficiency and performance...
```

Given the design principles, we then prompt the LLM to generate analogical inspirations. Our definition of analogy draws on the work of Gentner [29]. This definition involves identifying parallel relations from a source domain to apply to a target domain even when their surface features differ. We have structured our prompt using this definition of analogies by finding visually concrete objects from three source domains (nature, architecture, fashion) that convey concepts of the target domain (abstract concept).

```
You are a <subject> designer. The design principles in <
subject> design are as follows: <design principles from Step
 1>. Brainstorm analogical inspirations for <subject> design
 to convey a sense of <concept> from one of the following
domains: nature, architecture, or fashion. Answer in a
bullet-point list of 10 items (item1\nitem2...\nitem3) using
 visually-concrete objects not adjectives and don't repeat.
```

We empirically found that prompting specifically for the domains of nature, architecture, and fashion leads to particularly interesting inspirations. Furthermore, these domains are outside of the primary
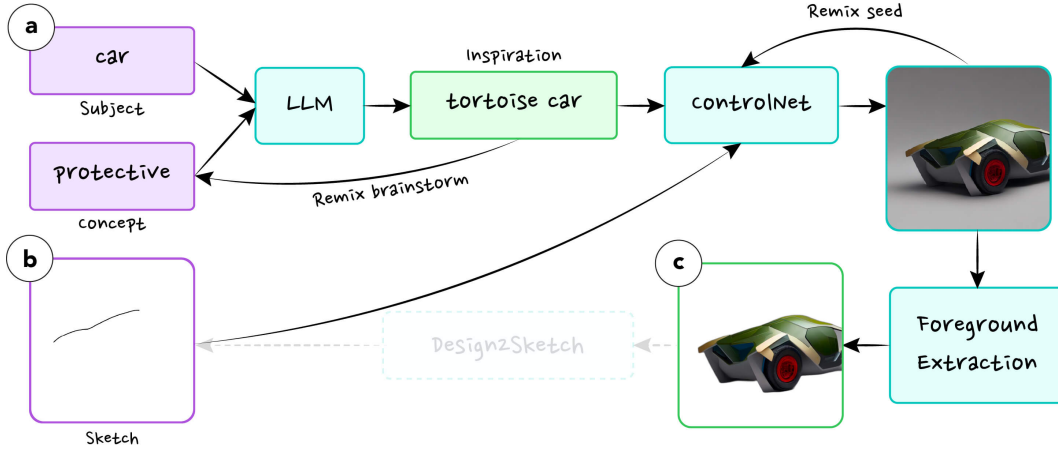
Figure 2: Sketch2Design pipeline, including (a) inspiration generation with LLMs, (b) sketch-guided design generation, and (c) foreground extraction.

product design domain and are common sources of inspiration for designers. The LLM then provides single noun-phrases as results (e.g., protectiveness & → tortoise, armadillo, armor).

The user may select a recommended inspiration and continue branching out to explore further inspirations. For example, selecting tortoise and rerunning the analogy inspiration chain could result in new analogies such as tortoise → tank, backpack, treasure chest). The user may also freely change their concept (e.g., changing *protectiveness* to *freedom*, resulting in a new set of analogical inspirations). In our current implementation, users cannot return to previous inspirations or explore multiple inspiration branches in parallel. When a user selects an inspiration, it serves as a base to generate another set of inspirations. Furthermore, the generated inspirations remain independent of what the user sketches on the canvas. These features suggest potential areas for future work.

After selecting an analogical design inspiration, the user may create product designs by sketching on a canvas. Our conversations with the professional design team revealed that they often start a design with a single silhouette line (Design Goal 1). In Inkspire, the user may start generating images with as little as a single stroke (Figure 2b). Using ControlNet [72] to guide Stable Diffusion [56], we generate a product design guided by the initial stroke. The user may continue adding additional strokes. Each time a stroke is drawn, we generate a new design, making the creation process iterative and implicitly encouraging users to focus on sketching instead of engineering text prompts (i.e., the current paradigm of working with T2I models).

ControlNet does not support per-stroke interaction out-of-the-box as it struggles with incomplete sketches, which is especially problematic during the initial stages of user sketching. Thus, we adapted the ControlNet model with a dynamic guidance scale to enable our desired per-stroke interaction (Equation 1). The guidance scale is a parameter for controlling how closely the model adheres to user input. We initialize with a low guidance scale to handle incomplete sketches. We progressively increase the guidance scale as the designer adds more ink to the sketch, pushing the generations to become more sensitive to the user's sketch over time.

$$G(n) = 7 - 4 \cdot 0.5^{\frac{n}{3}} \qquad (1)$$

where $G(n)$ is the guidance scale at n number of strokes by the user. The guidance scale starts at 3 when n=0 and approaches a maximum of 7 as the sketch becomes more complete ($n \approx 10$). The decay term $0.5^{\frac{n}{3}}$ creates a sharp logarithmic growth in guidance scale that converges gradually at the maximum value.

By gradually increasing the guidance scale, Inkspire allows the user to more precisely guide the generations as their sketches become more complete and well-defined. We maintain the same initial seed (seed specifies the random noise used to initialize image generations) between generations to maintain *consistency* between iterations and for *near-real-time* generations. The user may click on the "remix" button to change to a different seed and generate more diverse designs that break from the current thread that the designer is exploring. Finally, we remove unnecessary backgrounds from the generated design using a foreground extraction method [55] (Figure 2c).

4.3 Design2Sketch

The *Design2Sketch* component helps users build on top of previously generated designs by converting them into *scaffoldings* by abstracting a design into a reduced-fidelity sketch-style with the aim of reducing design fixation on the high-fidelity image (Figure 4). The scaffolding appears as a transparent underlay beneath the user's canvas, functioning like tracing paper [3] that updates in real-time as they sketch. This enables the user to draw inspiration from aspects of the previously generated designs and also helps them overcome the challenge of starting with a blank canvas [38], especially during the early stages of sketching. The user can continue iterating through sketching, completing the sketch-to-design-to-sketch feedback loop (Design Goal 3).

While there are many methods for reducing high fidelity images into lower-resolution, we introduce a novel approach for converting designs to sketch scaffolds. We initially tested existing methods including Canny edge detection [17], HED soft edge extraction [68], a state-of-the-art method for extracting main lines from manga
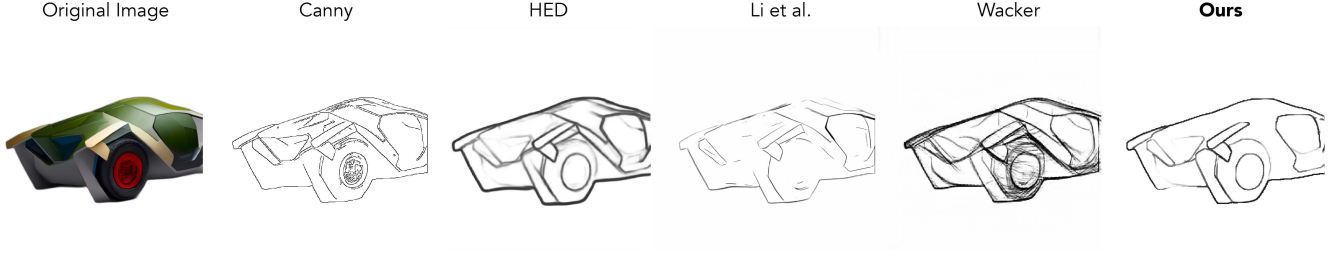
Figure 3: Comparison of our Design2Sketch method with potential alternative methods, such as edge detection, manga line extraction models, and models trained explicitly on pairs of sketches and images.
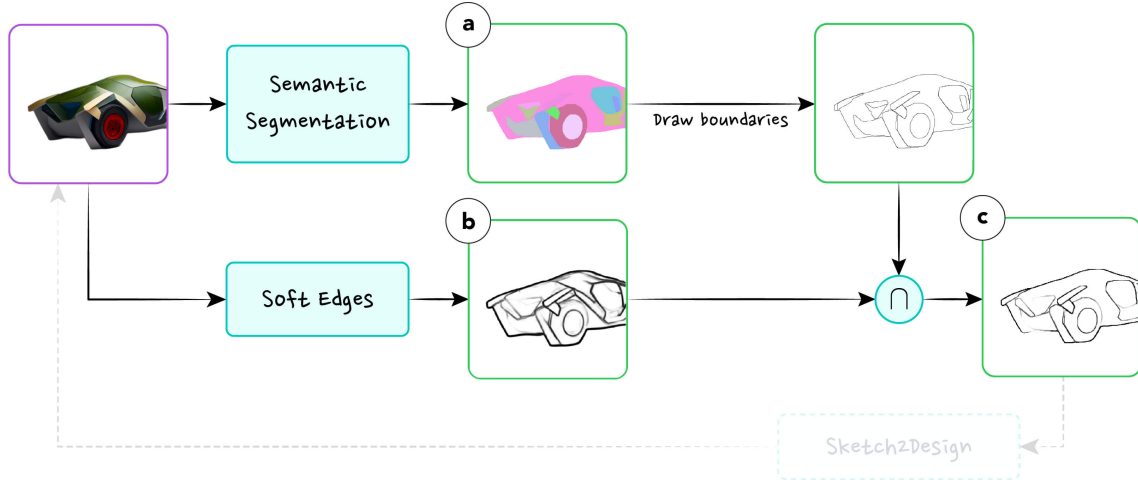

Figure 4: Design2Sketch pipeline, including (a) semantic segmentation, (b) soft edge extraction, and (c) computing an intersection.

illustrations [48], and a neural network method explicitly trained on pairs of sketches and images. (see Figure 3). We observed that edge extraction methods including Canny edge detection and HED soft edge extraction frequently produce unwanted lines caused by the texture of designs. Furthermore, we observed that manga line extraction methods, trained primarily on cartoon illustrations, can lead to a loss of key lines or produce broken lines. Finally, neural networks explicitly trained on pairs of sketches and images on the task of translating images to sketches can create artifacts of excessive sketch stylization, such as shading effects.

In our approach, we combine semantic segmentation and edge extraction. First, we perform semantic segmentation on the design [41] to create a segmentation map that color-codes a design into distinct semantic regions (Figure 4a). Given the segmentation map, we draw the boundaries between the different regions to create an image of semantic boundary lines. Second, we extract soft edges from the design [68] (Figure 4b). These soft edges include varying thickness and line weight, simulating a sketch-like look, though often with many redundant lines caused by texture. Finally, we take a pixel-wise intersection between the segmentation map boundary lines from the first step and the extracted soft edges from the second step as the final scaffolding (Figure 4c):

$$Scaffolding = Boundary(Seg(D)) \cap SoftEdge(D) \qquad (2)$$

where D is the generated design. Through this approach, we are able to acheve the best of both worlds – creating a sketch scaffolding that achieves a natural sketch look while focusing only on the design's key structural lines, filtered through the boundary lines from semantic segmentation step.

5 User Study

We conducted a within-subjects study to understand how Inkspire could address designers' pain points in working with T2I models, its potential to be integrated into design workflows, and identify areas for improvement. We compared Inkspire against a baseline condition using a typical ControlNet [72] setup consisting of a prompt box and a sketching canvas, adopting a similar interface layout as Inkspire (see Figure 14).

5.1 Participants

We invited twelve participants (P1-P12, 10 male and 2 female) to participate in a one-hour user study. Among the participants, six are professional designers who perform product design activities daily or weekly (self-rated confidence in product design μ=6.17, σ=0.75; self-rated confidence in drawing μ=6.00, σ=1.27; 7-point Likert scale) and six are novices who have moderate drawing experience (self-rated confidence in drawing μ=4.17, σ=1.33; 7-point Likert

scale) but do not actively engage in product design. The participants were recruited through known contacts and Upwork [1], a platform for hiring freelancers. They were not exposed to the Inkspire system or concept prior to the user study. Participants accessed Inkspire and the baseline tool through a web browser.

5.2 Measures

For both conditions, we asked participants to complete questionnaires to capture their perspectives on using both Inkspire and ControlNet. We assess creativity using the Creativity Support Index [22], measuring exploration, inspiration, engagement, expressiveness, tool transparency, and effort/reward tradeoff. We assess designers sense of the human-AI collaboration using questions from [44] measuring controllability, communication, harmony, partnership, attribution, and ownership. We also asked participants to rate their experiences of sketching by referencing the sketching principles from Bill Buxton's Sketching User Experiences [15], measuring how quick and timely, inexpensive and disposable, and loose and abstract sketching with each tool felt. We asked designers to rate the quality of the final design created with Inkspire and ControlNet and their overall experience satisfaction using each tool. All questionnaire questions were rated on a 7-point Likert scale (7=highly agree, 1-highly disagree). We compared the questionnaire measures using parametric paired t-tests. In addition, we log user interaction data, such as when participants draw a new sketch stroke, edit a prompt, and generate a new design. (Figures 13).

5.3 Procedure

5.3.1 Introduction (5 minutes). Participants provided informed consent and were given an overview of the study procedures. In addition, we briefly explain how components of the system, such as the analogy generation and sketching and scaffolding interaction, works.

5.3.2 Design Tasks (45 minutes). Participants completed a design task with Inkspire (Figure 1) and another design task with the baseline ControlNet tool (Figure 14). The two design tasks are "design a lamp with the theme of serenity" and "design a chair with the theme of fluidity." We counterbalance both the order of the tools and the order of the design tasks. After each condition, the participants completed the questionnaires.

5.3.3 Post-Study (10 minutes). Participants gave feedback during a short interview as well as through a free response questionnaire on the individual subcomponents of Inkspire, their overall experience of using Inkspire, whether they could see Inkspire being integrated into their design workflow, and areas for improving the tool. We reviewed these qualitative data to support the quantitative results.

6 Results

6.1 Creativity

Participants felt that Inkspire improved support for creativity across some attributes of the Creativity Support Index (CSI), shown in Figure 5. Notably, participants reported significantly higher exploration with Inkspire (μ=5.83, σ=1.27) as compared to the baseline (μ=3.83, σ=1.64), (t(11)=3.94, p<0.01, r=0.77, d_s=1.13). Participants also reported significantly higher inspiration (μ=5.92, σ=1.24) as

compared to the baseline μ=4.00, σ=1.41), (t(11)=3.44, p<0.01, r=0.72, d_s=0.99).

From our interviews, participants noted that the [analogy] inspirations feature is "*helpful while doing design ideations*" (P12) and "*a good tool to brainstorm in the early stage of design*"(P5). Participants also found Inkspire to effectively support the exploration of multiple ideas, such as a "*variety of forms, styles, patterns, and proportions*"(P5). P11 explained that they could explore different design directions by manipulating high-level concepts ("*I could just make a basic shape, and change a few keywords and the entire look and feel would change and present me with some great concepts*"). In contrast, we observed that in the baseline condition, many participants focused on sketching extensively on a single idea, a potential sign of the "sunk-cost effect" [7] (the more time spent in a given direction, the harder it is to move to a different one).

We also note that the other attributes of the CSI, Engagement, Expressiveness, Tool Transparency, and Effort/Reward Tradeoff, were not significantly different between Inkspire and the ControlNet baseline. While Inkspire did not improve these aspects of creativity over using ControlNet it did not appear to degrade them either. Overall, we find that Inkspire improved exploration and inspiration, supporting our original design goals and working toward reducing design fixation.

6.2 Human-AI Collaboration

6.2.1 Designer self-ratings of Human-AI Collaboration. Participants felt that Inkspire improved the experience of collaborating with the AI, across dimensions of Human-Machine Collaboration questions from [44] rated on 7-point Likert scales, shown in (Figure 7). Participants reported significantly higher controllability when using Inkspire (μ=5.58, σ=1.00) as compared to the baseline (μ=4.17, σ=1.19), (t(11)=3.56, p<0.01, r=0.73, d_s=1.03). Participants felt that they had significantly higher communication with Inkspire (μ=5.75, σ=1.14) as compared to the baseline (μ=3.92, σ=1.38), (t(11)=4.52, p<0.01, r=0.81, d_s=1.31). In addition, participants rated having a significantly higher degree of partnership with the AI when using Inkspire (μ=5.83, σ=1.03) as compared to the ControlNet baseline (μ=3.42, σ=1.56), (t(11)=4.57, p<0.01, r=0.81, d_s=1.32). Lastly, participants rated having significantly more of their own attribution in the designs when using Inkspire (μ=5.25, σ=1.29) as compared to the baseline (μ=3.67, σ=1.07),(t(11)=3.51, p<0.01, r=0.73, d_s=1.01). Participants did not report significant differences in their feelings of harmony with the AI or ownership over the designs between Inkspire and baseline ControlNet.

From our interviews and open-response questions, many participants commented that they found the scaffolding helpful in being able to *understand the current state of the AI* and plan subsequent sketches ("[The scaffolding] *was useful because it let me know where the current iteration was so I could tell where I'd like to move it next*" (P11)). With scaffolding, participants felt that they could *steer* the direction of the design by building on previous generations ("[scaffolding] *helped with building upon the previous AI-generated design and gave me a direction for what to adjust*" (P6), "*the AI-generated drawing overlay help[ed] me to draw my next line*"(P3), "*having the [scaffolding] as reference helped me [to] combine and remix [old designs] in my [new] sketches*"(P6)). These observations
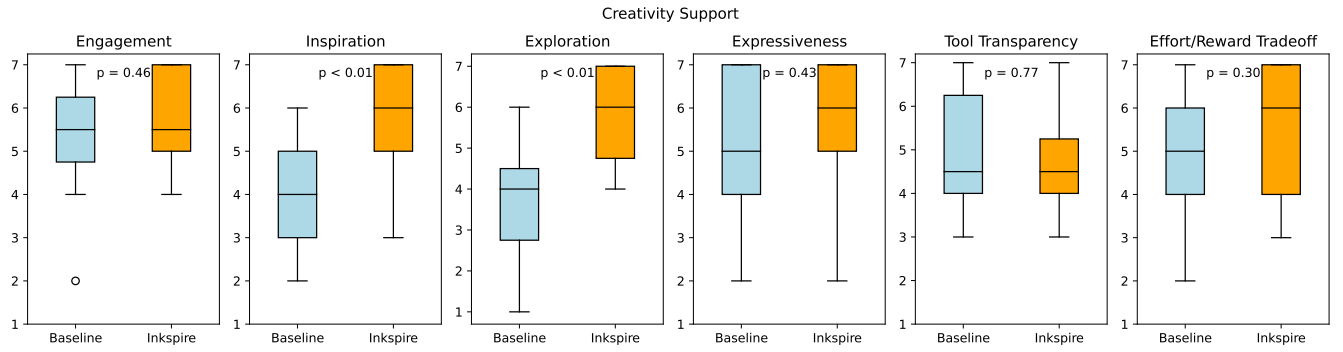
Figure 5: Results on creativity measured with the Creativity Support Index (CSI) [22] (7-point Likert scale, higher is better).
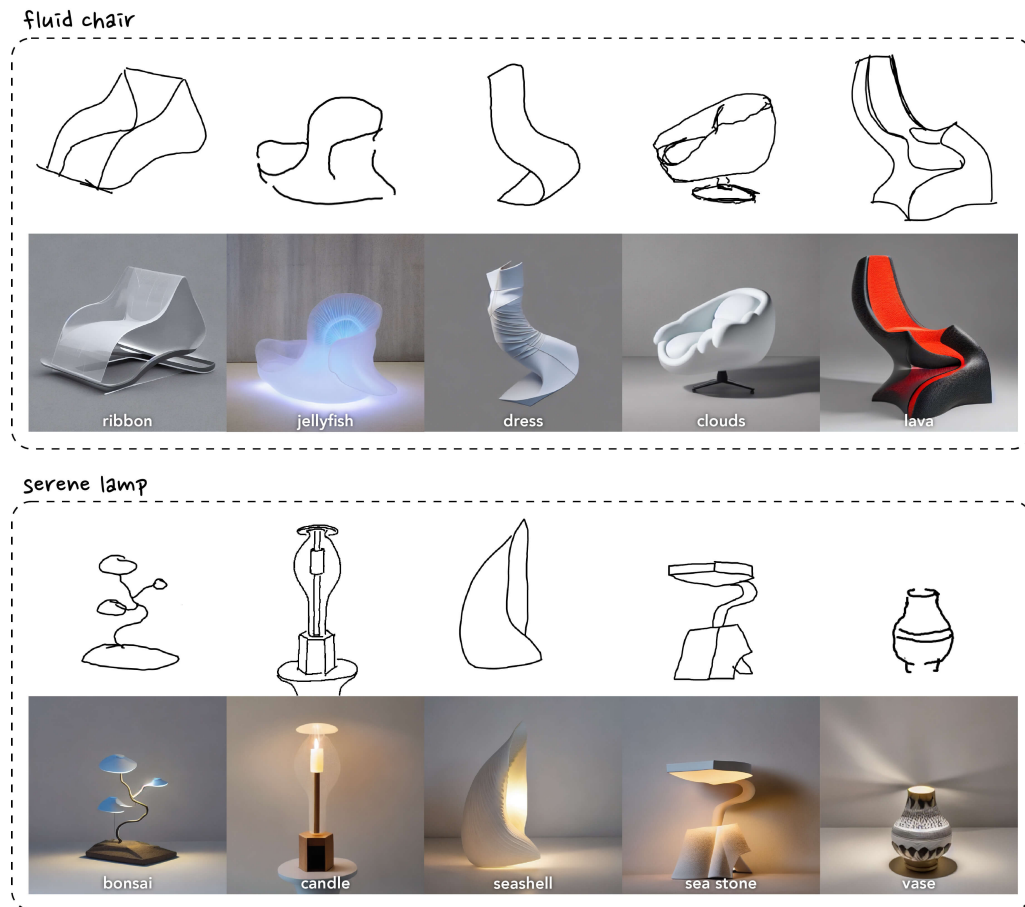

Figure 6: Example designs created by participants using Inkspire for the design tasks of designing a fluid chair and a serene lamp. The final participant-generated sketch is shown on the top, and the generated T2I image is shown on the bottom, along with the selected analogy word chosen by the participant.

and self-reported results suggest that Inkspire improves human-AI collaboration by increasing controllability, communication, and partnership.

6.2.2 *Prompting Behavior.* As expected, we observed that participants generally did less prompt engineering when using Inkspire (μ=8.50, σ=6.31, number of prompts) than when using the baseline (μ=11.9, σ=10.6, number of prompts). When using the baseline ControlNet system, participants often relied on manipulating the
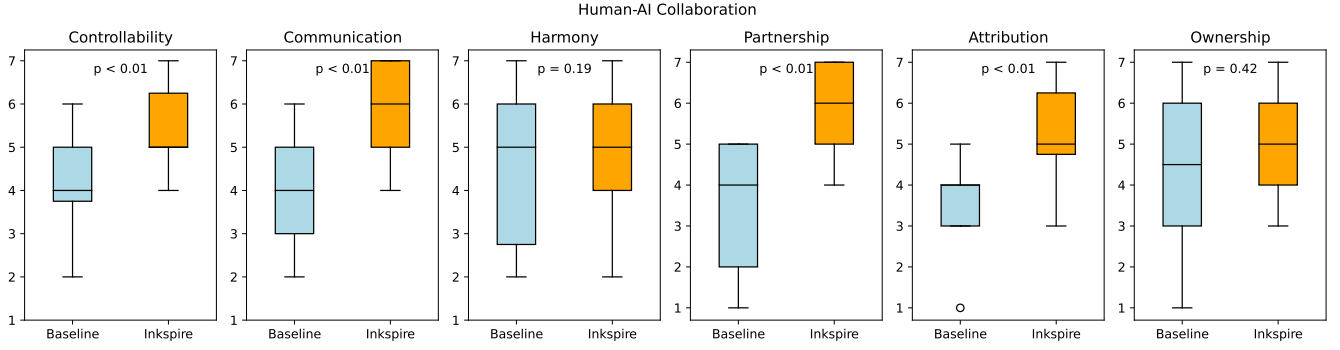
Figure 7: Results on human-AI collaboration measured with Human-Machine Collaboration Questions from [44] (7-point Likert scale, higher is better).

prompt to change their design whereas they used more diverse analogical inspirations as prompts when using Inkspire. Analyzing the semantic similarity between user prompts, we observed a much lower semantic similarity in the Inkspire condition (μ=0.51, σ=0.08) compared to the baseline (μ=0.76, σ=0.13), measured with BERTScore [73].

Looking at the logs of prompts designers used, we observed that participants using ControlNet often stuck to their original prompt, making small edits to make the prompt more and more detailed (see Figure 8). For example, P3 started with the prompt `serenity lamp` (i.e., explicitly prompting the AI with the abstract design task), and then expanded it with additions like `serenity lamp with clear glass and black base`, `serenity lamp with clear glass and black pedestal`, and so on. This result echoes the prompt fixation results of [63] and adds further evidence to the challenges designer have in knowing how to prompt [70]. Using Inkspire, participants made fewer manual prompt edits and frequently utilized the recommended analogical inspirations to *guide* their prompting direction. For example, the same participant above (P3) used various analogical inspirations for creating a *fluid* chair such as `silk`, `river`, and `waterfall`.

6.2.3 Sketching Behavior. We observed that participants drew fewer total sketch strokes using Inkspire (μ=17.3, σ=7.40, number of strokes) as compared to using the baseline (μ=59.8, σ=40.5, number of strokes). When using Inkspire, participants also had a lower sketching frequency (μ=12.5, σ=12.5, strokes/min) as compared to the baseline (μ=20.1, σ=12.6, strokes/min), though this was not significant ($t(11)$=-1.67, p=0.12, r=0.45, d_s=0.48). Participants also spent more time between strokes when using Inkspire (μ=13.8s, σ=12.1s) as compared to the baseline (μ=4.44s, σ=2.94s), ($t(11)$=2.66, p=0.02, r=0.63, d_s=0.77).

Despite participants sketching less and taking longer between adding more ink to their drawing, participants rated that sketch strokes felt significantly more inexpensive when using Inkspire (μ=6.08, σ=0.90, 7-point Likert scale, higher is better, based on Buxton's sketching principles [15]) as compared to using the baseline (μ=3.75, σ=1.60, 7-point Likert scale, higher is better), ($t(11)$=5.01, p<0.01, r=0.83, d_s=1.45). Participants also rated that the sketching was more abstract when using Inkspire (μ=5.75, σ=1.14, 7-point

Likert scale, higher is better, based on Buxton's sketching principles [15]) as compared to the baseline (μ=3.83, σ=1.47, 7-point Likert scale, higher is better), ($t(11)$=4.24, p<0.01, r=0.79, d_s=1.23). Overall, Inkspire appears to have provided designers with an interface for sketching ideas collaboratively with AI that improves over the more stilted sketching and image generation experience of other text-to-image systems.

6.2.4 Analogy Inspirations. We observed that participants generally explored multiple analogical inspirations (μ=4.58, σ=2.43) distributed across the three categories (Nature μ=2.08, σ=1.56; Architecture μ=1.58, σ=1.3; Fashion μ=0.917, σ=0.716) (Table 1). Architecture was the most common final choice (n=6 instances across participants), followed by Nature (n=4) and Fashion (n=2), though participants explored Nature-based inspirations more frequently on average. We also observed that participants frequently switched between analogy categories. For instance, P12 quickly switched between all three categories. Among the switches, we observed that Nature→Architecture (n=6) and Nature→Fashion (n=5) were common transitions. This may suggest that Nature can serve as a common "bridge" category. Many participants' final category deviated from their initial category after exploration. For instance, P2 started with Nature and selected Architecture after exploring 14 different inspirations across all categories. Overall, participants drew inspiration from all categories, with Nature playing a central role, though each category proved to be a valuable source for inspiration.

6.2.5 Overall Usage Patterns. By analyzing our logged interaction data, we observed common usage patterns for Inkspire vs. the baseline ControlNet workflow (please see Figure 10 for an illustrative example and Figure 13 for the full data). First, participants using Inkspire often start the initial ideation phase by experimenting with many analogical inspirations to explore the design space (shown as many thought bubbles during the beginning, Figure 10). This early divergent exploration behavior may suggest that designers were able to consider a wide range of inspirational ideas at the start of their process. Subsequently, participants using Inkspire engaged in a highly iterative manner of sketching and interacted with scaffolding, often drawing one or a small number of strokes, seeing a generation and scaffolding, then pausing a bit to consider where

Figure 8: Example user iterations with Inkspire (top section) vs. the baseline condition (bottom section). For each section, the top row shows the prompts, the middle row shows the sketches and scaffolds, and the bottom row shows the generated designs. Using Inkspire, users create diverse designs via analogies. In addition, users start with a single sketch line and continuously build on their sketch, with the guidance of scaffolding. In contrast, using the baseline, users typically draw a full sketch and update their prompt with incremental modifications. This may lead to a higher degree of fixation and a smaller design space explored.

to move next (shown as many pencils interleaved with sparkles, Figure 10). In contrast, when participants used the baseline system, they often did a complete full sketch (shown as large stretches of a single pencil, Figure 10) before ever hitting the generate button. Designers made more "Undo" actions, removing strokes that they had made (as shown in orange blocks). Overall, participants generally created more new designs with Inkspire than with the baseline condition (shown as significantly more sparkles in Inkspire than the baseline, Figure 10). Together, these patterns show how when participants used ControlNet, they focused more on crafting a whole sketch then handing it off to the generative AI rather than working collaboratively back and forth with the AI.

6.3 Final Design Quality

Participants rated their final creations as having higher design quality when using Inkspire (μ=5.92, σ=1.08) as compared to the baseline (μ=4.58, σ=1.93) on a 7-point Likert scale, shown in Figure 11 (left). Though this difference of ratings for design quality are not statistically significant, ratings showed less spread than when using the baseline ($t(11)$=2.11, p=0.06, r=0.54, d_s=0.61).

In Figure 6, we show several example creations from participants. While the quality of output in general appears relatively high, we do observe that the final results produced with Inkspire do appear to be more diverse than the results produced with Control-Net. Looking at the trend of designs created during a user session,
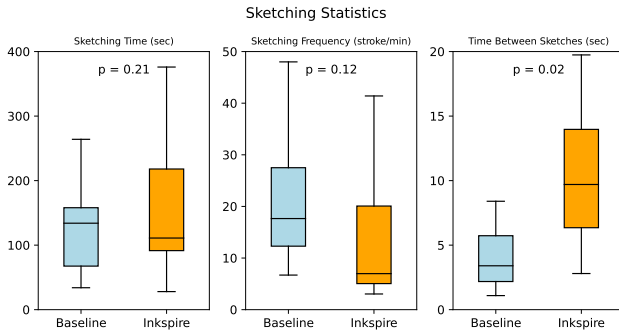
Figure 9: Results on sketching statistics, including total sketching time, sketching frequency, and time between sketches.

we see that the thread of designs created using ControlNet often do not change tracks, whereas those done with Inkspire show more diversity of conceptual exploration. An example showing chair designs from two different users illustrates this in Figure 8. The participant using Inkspire (top) shows a wide exploration of concepts for a `transparent` chair including `glass`, `silicon`, and `jellyfish`. Alternatively, a participant using ControlNet creates a `modern  Italian` chair with a high quality initial sketch and iterations on the prompt from `lancia style` to `lancia style, iconic` to `lancia style, iconic forms`. While both final result are generally of high quality, the final output from the participant using Inkspire is derived from a more exploratory process over design analogies over the more fixated design process from the designer working with ControlNet.

6.4 Designer Ratings of Usage Experience Satisfaction

Participants rated significantly higher satisfaction with their usage experience when using Inkspire (μ=6.08, σ=1.24) as compared to the baseline (μ=4.33, σ=0.98) when rated on a 7-point Likert scale (t(11)=3.54, p<0.01, r=0.73, d_s=1.02), shown in Figure 11 (right).

Based on the interaction logs, Inkspire may improve participants' experiences of sketching with AI by supporting low cost experimentation and helping participants draw abstract sketches that focus on the big picture rather than being bogged down by small details. In the baseline, we observed that participants drew a number of large strokes in succession *before* generating anything, with the majority being "filler" lines (Figure 13). This generally led to designers spending more time between trial-and-error attempts to achieve their intended result, possibly reducing their satisfaction when using ControlNet.

7 Discussion

We introduced Inkspire, a prototype system to explore new ways for designers to leverage generative AI while avoiding design fixation and embedded into their existing practice of sketching new concepts. We focused on a challenge identified in design research literature that exploring a wider range of ideas can lead to improved

design outcomes, though designers often find it difficult to naturally engage in diverse design exploration [27, 42]. Inkspire helped designers explore a more diverse design space by recommending them with diverse analogical anchors based on their initial design concept (e.g., turtle, shield, or bunker for embodying the concept of "protective"). These analogical concepts were used as input along with the sketch from the designer to generate new images. The designer could then continue to iterate on the designs collaboratively with the AI by leveraging stroke-by-stroke generation and building on low-resolution sketch scaffolds of high fidelity renders. Through a user study, we found that participants using Inkspire explored a wider design space and qualitatively changed their ideation process to be more conceptually iterative and collaborative with the AI. In contrast, participants in a control condition, in which they used a state-of-the-art ControlNet, primarily focused their efforts on crafting full sketches and making small changes to prompt inputs before having the AI render an image.

A core challenge we grappled with in this work was helping designers to leverage generative AI while avoiding fixation on the outputs of the AI. We believed this to be especially likely in the context of creative design and sketching because of several factors derived from our formative conversations with professional designers and prior research: text prompting leading to what designers called "unnatural" interaction, text-to-image models struggling with generating inspiring designs from abstract concepts, both often due to designer's limited abilities to prompt well [24, 63, 70, 71] and the AI generating outputs that often look overly "finished", a well-known cause of design fixation [18, 21, 63].

Our results suggest a fundamental issue with the structure of prompt- or sketch-focused generative image interfaces, which is that current controls through prompts or sketches can engender an affordance of needing to feel relatively complete in order to convey the designer's intention. The results from when our participants used ControlNet are similar to results from Wadinambiarachchi et al. [63] where designers focus most of their effort on minor iteration of prompts. Needing to provide a rich enough prompt or a full enough sketch may explain why our participants felt less control over the ideation and less communication and partnership with the AI model.

In contrast, with Inkspire the design problem prompt and analogical keyword concepts provide a scaffolding in the design space allowing a single stroke to have expressive meaning while also generating an output that is relevant to the user's goals. Relatedly, our work also explores ways for designers to have a shared mental model of what they are trying to accomplish versus what the AI is trying to do. For example, an interesting point we noticed from our participants was that they found the rough sketches inferred from the AI generation to be useful as control points that showed what the AI was doing and how the user could emphasize or change that, for example by either drawing over or changing a stroke on the roughed sketch, respectively (please see Figure 12). In this way the rough sketch serves not only as a scaffold for the designer's iterative sketching but as a communication tool and shared mental model with a generative AI partner. The continual updating of both sketch and analogical concept ultimately leads to a more iterative and collaborative process where the designer works in turns with the AI. Other generative AI tools could potentially build on this
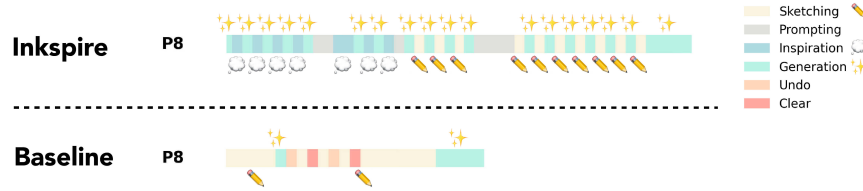
Figure 10: Example user interaction log. We observe that, using Inkspire, the user started by ideating several analogical inspirations (thought bubble). Next, they performed sketching in a highly iterative manner (pencil+sparkles). Overall, the user generated a significant amount of new designs (sparkles). In contrast, using the baseline condition, the user sketched in large stretches with infrequent new design generations. We see that this pattern generally holds true across participants (please see Figure 13 for data on all participants).
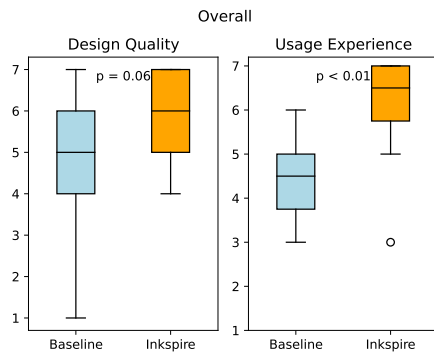

Figure 11: Results on design quality (left) and usage experience (right) (7-point Likert scale, higher is better).

pattern of rapid turning taking to help promote more co-creativity for their creative task.

Overall, the interactions we designed into Inkspire appear to have worked in helping our participants avoid design fixation. As commented on by P6 "*the AI was helpful in providing alternative designs I did not consider previously, especially by adopting a new semantic understanding of the sketch that differed from my initial intention.*" However, even though working with Inkspire pushed the designers into a space they had not originally considered and even resulted in them sketching less in total, they found that we also succeeded in overcoming issues where designers feel less agency when working with generative AI. That the participants felt they had more agency and more ownership of the final design outputs suggests the analogies, stroke-based interaction pattern, and scaffolds led to more fluid interaction through a design space. Again quoting from P6, Inkspire had "*an interesting workflow that I think helps make the tool disappear more in comparison to a prompt-based drawing tool. It allowed me to focus more on sketching than prompt engineering.*"

While our overall results show that Inkspire achieved our design goals, there are some potential issues with Inkspire that could arise and could be considered in future co-creative systems. First, although we see more evidence of designers moving to more distinct ideas throughout a session, they still work along a single thread.

Prior research has shown that designers often prototype ideas in parallel and that such parallel exploration can lead to better design outcomes [27]. One potential way to enable parallel explorations in Inkspire is to generate numerous analogies for a concept and generate an image for each during each sketch step. Second, as Inkspire was primarily designed for early explorations, it may not serve a designer as well during later stages of the design when they have a clear idea in mind that they are trying to render. In this case, a designer may prefer using a tool such as ControlNet where they can focus on the details they want before handing it to an AI. Lee et al. similarly proposed an adaptive multimodal T2I system that initially supports early ideation with a prompt-guided (e.g., protective chair) and sketch-supported (e.g., simple sketch) system, which gradually evolves into sketch-guided (e.g., detailed sketch) and prompt-supported (e.g., "more curved back") system to help refine the concept in later stages of the creative process [45]. This suggests potential opportunities for generative systems to provide controls for how much turn taking the AI should aim to have with the designer. Third, Inkspire and other generative AI approaches may help individual designers break their fixation on specific ideas, however, if common generative models are used, it may promote collective design fixation, where everyone's images begin to look the same as they have a common source [6]. This being said, Inkspire may work to counteract such collective fixation by promoting increased designer engagement by focusing more on suggestions of designs though lower-resolution sketch underlays rather than pixel-level in-painting. Future systems could explore other techniques for avoiding design fixation, such as showing partial photographs [21] instead of sketch underlays. Other ideas for reducing fidelity such as blurring or filtering could also promote designers to fill in details on their own, potentially leading to designs less similar to common output from AI models.

7.1 Future Improvements to Inkspire

There are several avenues for future work on improving Inkspire collected from participants' comments. First, we could improve the sketching canvas with more advanced sketching features such as line weight control and shading tools, which could allow users to sketch with greater detail. Second, participants suggested adding more fine-grained region-based control to the Sketch2Design pipeline.
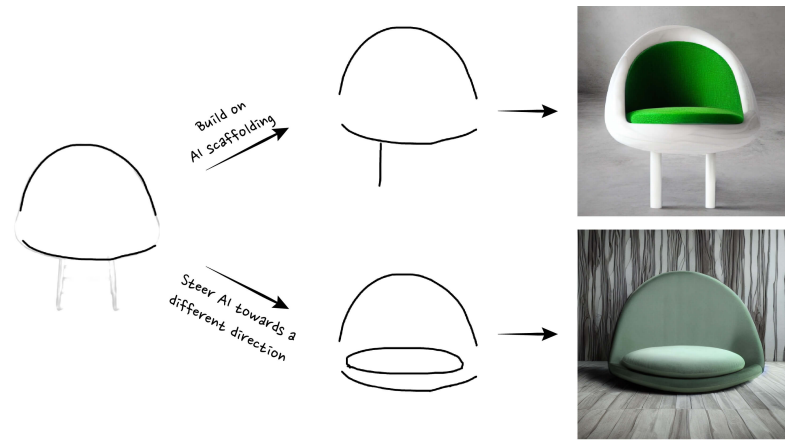
Figure 12: Given a scaffold, the user may choose to build on it (by tracing it) or steer the AI towards a different direction.

For example, a user could specify different material types for different regions of the sketch. Third, we could extend the analogical panel to include the capability to go back to a previous inspiration and perform multiple branches of inspiration explorations in parallel. This could help users compare vastly different design directions. Fourth, we currently adopt a two-step prompting mechanism for generating analogical inspirations: defining the design principles for the target domain, then generating visually-concrete object inspirations for the target domain using nature, architecture, and fashion as source domains. Future work could explore more complex prompting structures, such as traversing hierarchical tree structures [39], that might lead to better analogical inspirations. Fifth, Inkspire contains several features that could all affect a user's behavior, such as automatic per-stroke generation and sketch scaffolding. For future work, we could conduct additional ablation studies. For example, adding dynamic guidance scale to enable automatic per-stroke generation for the baseline condition or hiding sketch scaffolding for Inkspire. Sixth, we adopted a within-subjects design with a relatively small sample size (*n*=12), which could be improved through larger-scale studies. Seventh, participants suggested the possibility of adding additional guidance for the AI, such as mechanical and material constraints, though such a feature may require more research into how to better connect analogically created T2I design concepts with their feasibility to be manufactured. Finally, in this paper, we introduce a sketch-generation-scaffold interaction with generated analogies for abstract concepts through the application of product sketching. We think the sketching strategy could generalize to other forms of drawing, though the analogies might be better suited for product design. For example, very recent work has explored a similar strategy but using an image underlay instead of sketch scaffolding [59]. Future work could explore extending the system to additional design domains and incorporating background elements into designs for mockup or decoration.

8 Conclusion

This research identifies the challenges designers face when using Text-to-Image (T2I) models in design and proposes a set of design guidelines for creating new interfaces for interacting with T2I systems. The research prototype, Inkspire, supports designers in prototyping product design concepts with analogical inspirations and through a complete sketch-to-design-to-sketch feedback loop. Feedback from designers using Inkspire suggests that it could encourage more inspiration and exploration of novel design ideas while allowing designers to focus on iterative sketching collaboratively with the T2I model rather than prompt engineering. These analogy and sketch scaffolding features allowed designers to generate initial ideas for a design concept from a single abstract concept and single stroke of the pen to expand them into multiple alternatives. The iterative interaction and turn-taking between the designer and the AI helped the designer guide the models toward novel design intentions with the potential to overcome design fixation. We are interested in extending Inkspire with more capabilities and hope that this research can help inform future research on sketch-driven tools for co-creating with Generative AI models.

Acknowledgments

References

[1] 2022. *Upwork.* Retrieved August 15, 2022 from https://www.upwork.com/
[2] 2024. *AI is plundering the imagination and replacing it with a slot machine.* Retrieved March 24, 2024 from https://thebulletin.org/2022/10/ai-is-plundering-the-imagination-and-replacing-it-with-a-slot-machine/
[3] 2024. *How to Learn to Draw by Tracing.* Retrieved February 16, 2024 from https://monikazagrobelna.com/2020/08/16/how-to-learn-to-draw-by-tracing/
[4] 2024. *Midjourney.* Retrieved March 24, 2024 from https://midjourney.com/
[5] 2024. *Vizcom.* Retrieved March 24, 2024 from https://app.vizcom.ai/
[6] Barrett R Anderson, Jash Hemant Shah, and Max Kreminski. 2024. Homogenization Effects of Large Language Models on Human Creative Ideation. In *Proceedings of the 16th Conference on Creativity & Cognition* (Chicago, IL, USA) *(C&C '24).* Association for Computing Machinery, New York, NY, USA, 413–425. doi:10.1145/3635636.3656204
[7] Hal R Arkes and Catherine Blumer. 1985. The psychology of sunk cost. *Organizational behavior and human decision processes* 35, 1 (1985), 124–140.

[8] Luca Benedetti, Holger Winnemöller, Massimiliano Corsini, and Roberto Scopigno. 2014. Painting with Bob: assisted creativity for novices. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*. 419–428.

[9] James Betker, Gabriel Goh, Li Jing, Tim Brooks, Jianfeng Wang, Linjie Li, Long Ouyang, Juntang Zhuang, Joyce Lee, Yufei Guo, et al. 2023. Improving image generation with better captions. *Computer Science. https://cdn. openai. com/papers/dall-e-3. pdf* 2, 3 (2023), 8.

[10] Antoine Bordas, Pascal Le Masson, and Benoit Weil. 2024. Switching perspectives on generative artificial intelligence: a design view for humans-generative AI co-creativity. In *R&D Management Conference 2024*.

[11] Stephen Brade, Bryan Wang, Mauricio Sousa, Sageev Oore, and Tovi Grossman. 2023. Promptify: Text-to-image generation through interactive prompt exploration with large language models. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. 1–14.

[12] Tim Brooks, Aleksander Holynski, and Alexei A Efros. 2023. Instructpix2pix: Learning to follow image editing instructions. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 18392–18402.

[13] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems* 33 (2020), 1877–1901.

[14] Raluca Budiu. 2014. Memory recognition and recall in user interfaces. *Nielsen Norman Group* 1 (2014).

[15] Bill Buxton. 2010. *Sketching user experiences: getting the design right and the right design.* Morgan kaufmann.

[16] Alice Cai, Steven R Rick, Jennifer L Heyman, Yanxia Zhang, Alexandre Filipowicz, Matthew Hong, Matt Klenk, and Thomas Malone. 2023. DesignAID: Using Generative AI and Semantic Diversity for Design Inspiration. In *Proceedings of The ACM Collective Intelligence Conference* (Delft, Netherlands) *(CI '23)*. Association for Computing Machinery, New York, NY, USA, 1–11. doi:10.1145/3582269.3615596

[17] John Canny. 1986. A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence* 6 (1986), 679–698.

[18] Carlos Cardoso, Petra Badke-Schaub, and Ana Luz. 2009. Design fixation on non-verbal stimuli: The influence of simple vs. rich pictorial information on design problem-solving. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 49057. 995–1002.

[19] How To Draw Cars. 2017. Pro Designer Teaches You How To Create Variations on Your Car Design Themes. https://youtu.be/GxFloTllsKE?si=yWYl9w_etZVZPbnK&t=80. Accessed: September 12, 2024.

[20] Minsuk Chang, Stefania Druga, Alexander J Fiannaca, Pedro Vergani, Chinmay Kulkarni, Carrie J Cai, and Michael Terry. 2023. The prompt artists. In *Proceedings of the 15th Conference on Creativity and Cognition*. 75–87.

[21] Peiyao Cheng, Ruth Mugge, and Jan PL Schoormans. 2014. A new strategy to reduce design fixation: Presenting partial photographs to designers. *Design Studies* 35, 4 (2014), 374–391.

[22] Erin Cherry and Celine Latulipe. 2014. Quantifying the creativity support of digital tools through the creativity support index. *ACM Transactions on Computer-Human Interaction (TOCHI)* 21, 4 (2014), 1–25.

[23] Nicholas Davis, Safat Siddiqui, Pegah Karimi, Mary Lou Maher, and Kazjon Grace. 2019. Creative Sketching Partner: A Co-Creative Sketching Tool to Inspire Design Creativity.. In *ICCC*. 358–359.

[24] Richard Lee Davis, Thiemo Wambsganss, Wei Jiang, Kevin Gonyop Kim, Tanja Käser, and Pierre Dillenbourg. 2024. Fashioning Creative Expertise with Generative AI: Graphical Interfaces for Design Space Exploration Better Support Ideation Than Text Prompts. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 167, 26 pages. doi:10.1145/3613904.3642908

[25] A. de Rooij and M. Mose Biskjaer. 2024. Expecting the unexpected: A review of surprise in design processes. In *DRS2024: Boston, 23–28 June*, C. Gray, E. Ciliotta Chehade, P. Hekkert, L. Forlano, P. Ciuccarelli, and P. Lloyd (Eds.). Boston, USA. doi:10.21606/drs.2024.333

[26] Jon-Michael Deldin and Megan Schuknecht. 2013. The AskNature database: enabling solutions in biomimetic design. In *Biologically inspired design: Computational methods and tools*. Springer, 17–27.

[27] Steven P Dow, Alana Glassco, Jonathan Kass, Melissa Schwarz, Daniel L Schwartz, and Scott R Klemmer. 2010. Parallel prototyping leads to better design results, more divergence, and increased self-efficacy. *ACM Transactions on Computer-Human Interaction (TOCHI)* 17, 4 (2010), 1–24.

[28] Zezhong Fan, Xiaohan Li, Kaushiki Nag, Chenhao Fang, Topojoy Biswas, Jianpeng Xu, and Kannan Achan. 2024. Prompt Optimizer of Text-to-Image Diffusion Models for Abstract Concept Understanding. In *Companion Proceedings of the ACM Web Conference 2024* (Singapore, Singapore) *(WWW '24)*. Association for Computing Machinery, New York, NY, USA, 1530–1537. doi:10.1145/3589335.3651927

[29] Dedre Gentner. 1983. Structure-mapping: A theoretical framework for analogy. *Cognitive science* 7, 2 (1983), 155–170.

[30] Ashok K Goel. 1997. Design, analogy, and creativity. *IEEE expert* 12, 3 (1997), 62–70.

[31] Ashok K Goel, Swaroop Vattam, Bryan Wiltgen, and Michael Helms. 2012. Cognitive, collaborative, conceptual and creative—Four characteristics of the next generation of knowledge-based CAD systems: A study in biologically inspired design. *Computer-Aided Design* 44, 10 (2012), 879–900.

[32] Yihan Hou, Manling Yang, Hao Cui, Lei Wang, Jie Xu, and Wei Zeng. 2024. C2Ideas: Supporting Creative Interior Color Design Ideation with Large Language Model. *arXiv preprint arXiv:2401.12586* (2024).

[33] Zhengyu Huang, Yichen Peng, Tomohiro Hibino, Chunqi Zhao, Haoran Xie, Tsukasa Fukusato, and Kazunori Miyata. 2022. dualface: Two-stage drawing guidance for freehand portrait sketching. *Computational Visual Media* 8 (2022), 63–77.

[34] Emmanuel Iarussi, Adrien Bousseau, and Theophanis Tsandilas. 2013. The drawing assistant: Automated drawing guidance and feedback from photographs. In *ACM Symposium on User Interface Software and Technology (UIST)*. ACM.

[35] David G Jansson and Steven M Smith. 1991. Design fixation. *Design studies* 12, 1 (1991), 3–11.

[36] Shuo Jiang, Jie Hu, Kristin L Wood, and Jianxi Luo. 2022. Data-driven design-by-analogy: state-of-the-art and future directions. *Journal of Mechanical Design* 144, 2 (2022), 020801.

[37] Shuo Jiang, Jianxi Luo, Guillermo Ruiz-Pava, Jie Hu, and Christopher L Magee. 2021. Deriving design feature vectors for patent images using convolutional neural networks. *Journal of Mechanical Design* 143, 6 (2021), 061405.

[38] Caneel K Joyce. 2009. *The blank page: Effects of constraint on creativity.* University of California, Berkeley.

[39] Hyeonsu B Kang, David Chuan-En Lin, Nikolas Martelaro, Aniket Kittur, Yan-Ying Chen, and Matthew K Hong. 2023. BioSpark: An End-to-End Generative System for Biological-Analogical Inspirations and Ideation. *arXiv preprint arXiv:2312.11388* (2023).

[40] Mohammed Khaliq, Diego Frassinelli, and Sabine Schulte im Walde. 2024. Comparison of Image Generation Models for Abstract and Concrete Event Descriptions. In *Proceedings of the 4th Workshop on Figurative Language Processing (FigLang 2024)*. 15–21.

[41] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C Berg, Wan-Yen Lo, et al. 2023. Segment anything. *arXiv preprint arXiv:2304.02643* (2023).

[42] Ken Kocienda. 2018. *Creative selection: Inside Apple's design process during the golden age of Steve Jobs.* Pan Macmillan.

[43] E Kwon, A Pehlken, K-D Thoben, A Bazylak, and LH Shu. 2019. Visual similarity to aid alternative-use concept generation for retired wind-turbine blades. *Journal of Mechanical Design* 141, 3 (2019), 031106.

[44] Tomas Lawton, Francisco J Ibarrola, Dan Ventura, and Kazjon Grace. 2023. Drawing with reframer: Emergence and control in co-creative ai. In *Proceedings of the 28th International Conference on Intelligent User Interfaces*. 264–277.

[45] Seung Won Lee, Tae Hee Jo, Semin Jin, Jiin Choi, Kyungwon Yun, Sergio Bromberg, Seonghoon Ban, and Kyung Hoon Hyun. 2024. The Impact of Sketch-guided vs. Prompt-guided 3D Generative AIs on the Design Exploration Process. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 1057, 18 pages. doi:10.1145/3613904.3642218

[46] Yong Jae Lee, C Lawrence Zitnick, and Michael F Cohen. 2011. Shadowdraw: real-time user guidance for freehand drawing. *ACM Transactions on Graphics (ToG)* 30, 4 (2011), 1–10.

[47] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, et al. 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems* 33 (2020), 9459–9474.

[48] Chengze Li, Xueting Liu, and Tien-Tsin Wong. 2017. Deep extraction of manga structural lines. *ACM Transactions on Graphics (TOG)* 36, 4 (2017), 1–12.

[49] Jiayi Liao, Xu Chen, Qiang Fu, Lun Du, Xiangnan He, Xiang Wang, Shi Han, and Dongmei Zhang. 2024. Text-to-image generation for abstract concepts. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 38. 3360–3368.

[50] Alex Limpaecher, Nicolas Feltman, Adrien Treuille, and Michael Cohen. 2013. Real-time drawing assistance through crowdsourcing. *ACM Transactions on Graphics (TOG)* 32, 4 (2013), 1–8.

[51] David Chuan-En Lin and Nikolas Martelaro. 2024. Jigsaw: Supporting Designers to Prototype Multimodal Applications by Assembling AI Foundation Models. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*.

[52] Zhiyu Lin, Upol Ehsan, Rohan Agarwal, Samihan Dani, Vidushi Vashishth, and Mark Riedl. 2023. Beyond Prompts: Exploring the Design Space of Mixed-Initiative Co-Creativity Systems. *arXiv preprint arXiv:2305.07465* (2023).

[53] J. S. Linsey, A. B. Markman, and K. L. Wood. 2012. Design by Analogy: A Study of the WordTree Method for Problem Re-Representation. *Journal of Mechanical Design* 134, 4 (04 2012), 041009. doi:10.1115/1.4006145 arXiv:https://asmedigitalcollection.asme.org/mechanicaldesign/article-pdf/134/4/041009/5606684/041009_1.pdf

[54] Vivian Liu, Tao Long, Nathan Raw, and Lydia Chilton. 2023. Generative disco: Text-to-video generation for music visualization. *arXiv preprint arXiv:2304.08551* (2023).

[55] Xuebin Qin, Zichen Zhang, Chenyang Huang, Masood Dehghan, Osmar R Zaiane, and Martin Jagersand. 2020. U2-Net: Going deeper with nested U-structure for salient object detection. *Pattern recognition* 106 (2020), 107404.

[56] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. 2022. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 10684–10695.

[57] Chitwan Saharia, William Chan, Huiwen Chang, Chris Lee, Jonathan Ho, Tim Salimans, David Fleet, and Mohammad Norouzi. 2022. Palette: Image-to-image diffusion models. In *ACM SIGGRAPH 2022 conference proceedings*. 1–10.

[58] Patsorn Sangkloy, Nathan Burnell, Cusuh Ham, and James Hays. 2016. The sketchy database: learning to retrieve badly drawn bunnies. *ACM Transactions on Graphics (TOG)* 35, 4 (2016), 1–12.

[59] Vishnu Sarukkai, Lu Yuan, Mia Tang, Maneesh Agrawala, and Kayvon Fatahalian. 2024. Block and Detail: Scaffolding Sketch-to-Image Generation. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. 1–13.

[60] L Siddharth and Amaresh Chakrabarti. 2018. Evaluating the impact of Idea-Inspire 4.0 on analogical transfer of concepts. *Ai Edam* 32, 4 (2018), 431–448.

[61] Kihoon Son, DaEun Choi, Tae Soo Kim, Young-Ho Kim, and Juho Kim. 2024. Genquery: Supporting expressive visual search with generative models. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–19.

[62] Julian FV Vincent and Darrell L Mann. 2002. Systematic technology transfer from biology to engineering. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 360, 1791 (2002), 159–173.

[63] 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 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

[64] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, et al. 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems* 35 (2022), 24824–24837.

[65] Blake Williford, Samantha Ray, Jung In Koh, Josh Cherian, Paul Taele, and Tracy Hammond. 2023. Exploring Creativity Support for Concept Art Ideation. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 112, 7 pages. doi:10.1145/3544549.3585684

[66] Shengqiong Wu, Hao Fei, Hanwang Zhang, and Tat-Seng Chua. 2024. Imagine that! abstract-to-intricate text-to-image synthesis with scene graph hallucination diffusion. In *Proceedings of the 37th International Conference on Neural Information Processing Systems* (New Orleans, LA, USA) *(NIPS '23)*. Curran Associates Inc., Red Hook, NY, USA, Article 3468, 20 pages.

[67] Jun Xie, Aaron Hertzmann, Wilmot Li, and Holger Winnemöller. 2014. PortraitSketch: Face sketching assistance for novices. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*. 407–417.

[68] Saining Xie and Zhuowen Tu. 2015. Holistically-nested edge detection. In *Proceedings of the IEEE international conference on computer vision*. 1395–1403.

[69] Yutong Xie, Zhaoying Pan, Jinge Ma, Luo Jie, and Qiaozhu Mei. 2023. A prompt log analysis of text-to-image generation systems. In *Proceedings of the ACM Web Conference 2023*. 3892–3902.

[70] JD Zamfirescu-Pereira, Richmond Y Wong, Bjoern Hartmann, and Qian Yang. 2023. Why Johnny can't prompt: how non-AI experts try (and fail) to design LLM prompts. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–21.

[71] Chengzhi Wang, Weijie Wang, Paul Pangaro, Nikolas Martelaro, and Daragh Byrne. 2023. Generative Image AI Using Design Sketches as input: Opportunities and Challenges. In *Proceedings of the 15th Conference on Creativity and Cognition* (Virtual Event, USA) *(C&C '23)*. Association for Computing Machinery, New York, NY, USA, 254–261. doi:10.1145/3591196.3596820

[72] Lvmin Zhang, Anyi Rao, and Maneesh Agrawala. 2023. Adding conditional control to text-to-image diffusion models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 3836–3847.

[73] Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q Weinberger, and Yoav Artzi. 2019. Bertscore: Evaluating text generation with bert. *arXiv preprint arXiv:1904.09675* (2019).

[74] Zijian Zhang and Yan Jin. 2020. An unsupervised deep learning model to discover visual similarity between sketches for visual analogy support. In *International design engineering technical conferences and computers and information in engineering conference*, Vol. 83976. American Society of Mechanical Engineers.
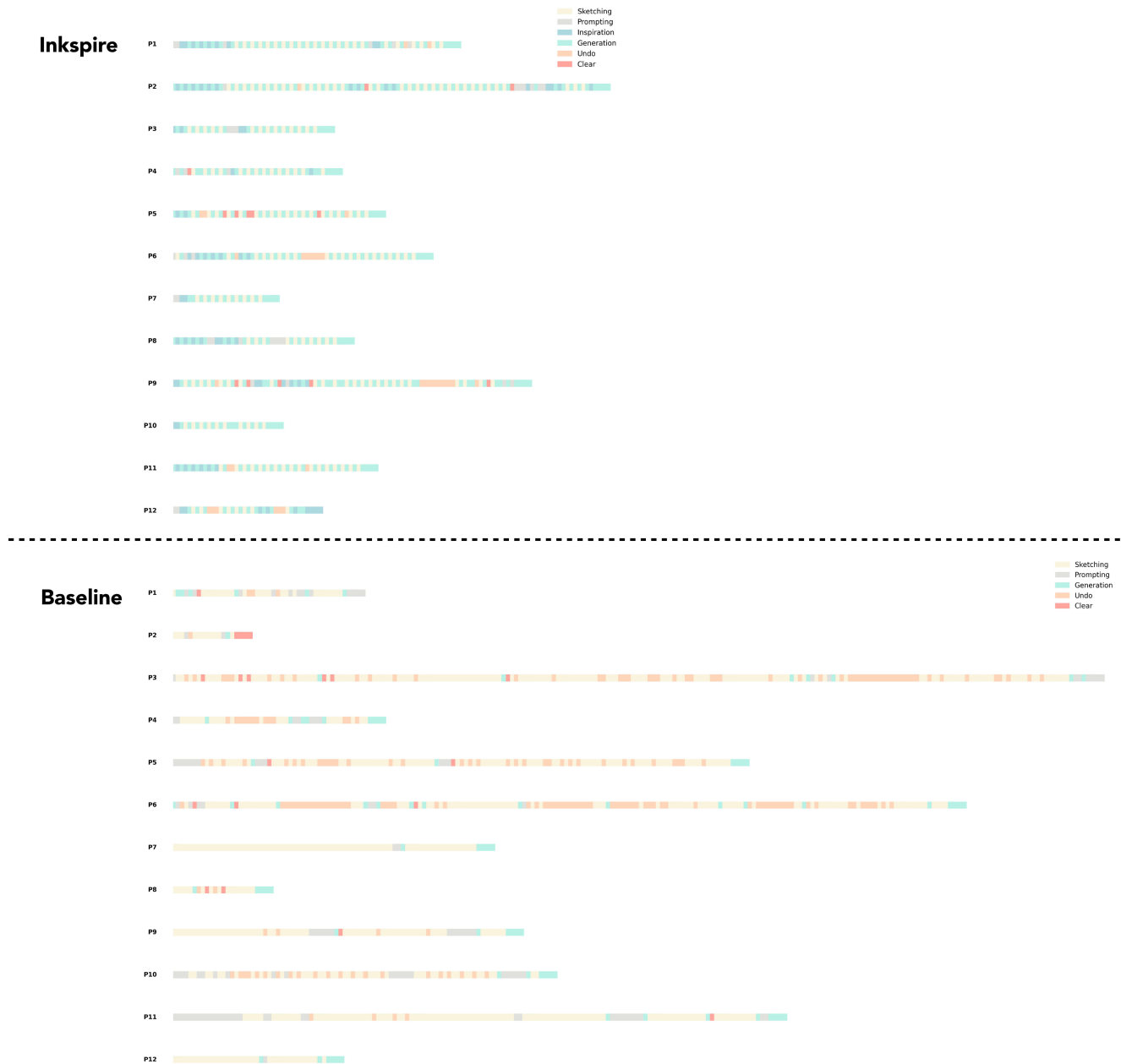
Figure 13: Complete participant usage logs when using Inkspire vs. the baseline. Please see Figure 10 for an annotated illustrative example.
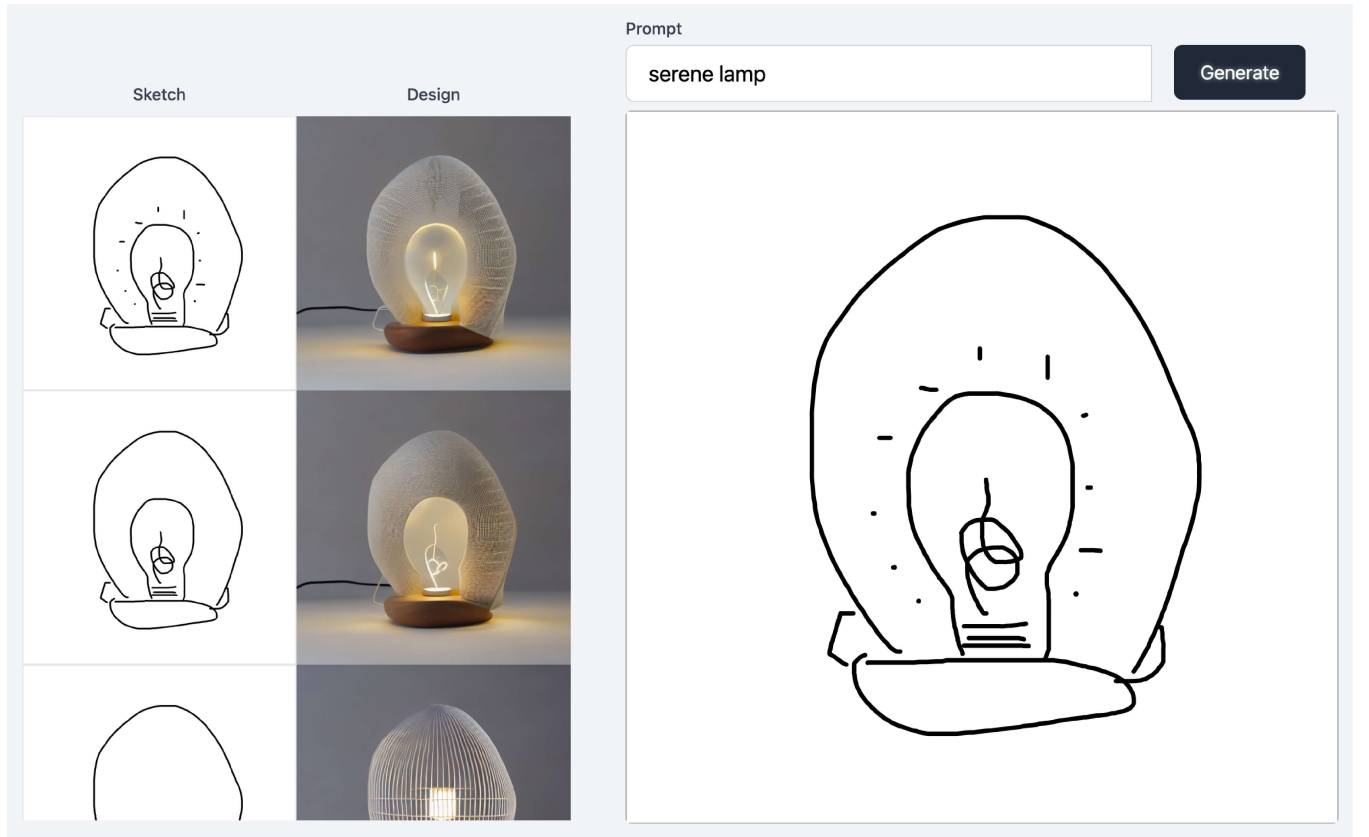
Figure 14: The baseline interface adopts a similar layout as Inkspire, but without analogical inspirations, scaffolding sketch underlays, and stroke-by-stroke generations. The participant would typically manually specify a prompt and draw a complete sketch for the T2I model – a standard workflow of how designers work with T2I models with additional sketch control [72].

| ID | Analogies | Category Count | | | | Final |
		Total	Nature	Architecture	Fashion	Category
P1	running cheetah, pleated dress, catwalk, skyscraper, spiral staircase, pleated dress, jellyfish	6	2	2	2	Nature
P2	waves, waterfall, silk train, flowing dress, flowing dress, waves, running river, running river, waves, willow tree, infinity pool, concrete, silhouettes of brutalist buildings, rubble	10	4	4	2	Architecture
P3	silk, river, clouds, waterfall	4	3	0	1	Nature
P4	zen garden, moonlight, candle, minimalist church	4	1	2	1	Architecture
P5	bamboo, bonsai, japanese tea house, zen garden	4	2	2	0	Architecture
P6	bay window, japanese tea house, zen garden, lotus flower, zen garden, japanese tea house	4	1	3	0	Architecture
P7	lotus	1	1	0	0	Nature
P8	bamboo grove, moonlight, roman column, snow peak, spring creek, bamboo, zen garden, moonlight	7	5	2	0	Nature
P9	waterfall, pawprint, seashell, sea stone, slinky, ribbon	6	4	0	2	Fashion
P10	flowing gown	1	0	0	1	Fashion
P11	floating bridge, lava, silk scarf, winding staircase	4	1	2	1	Architecture
P12	gaudi's architecture, ribbon, waterfall, running track curve	4	1	2	1	Architecture

Table 1: Analogical inspirations explored by participants and statistics on inspiration categories.