DesignWeaver: Dimensional Scaffolding for Text-to-Image Product Design

ANONYMOUS AUTHOR(S)

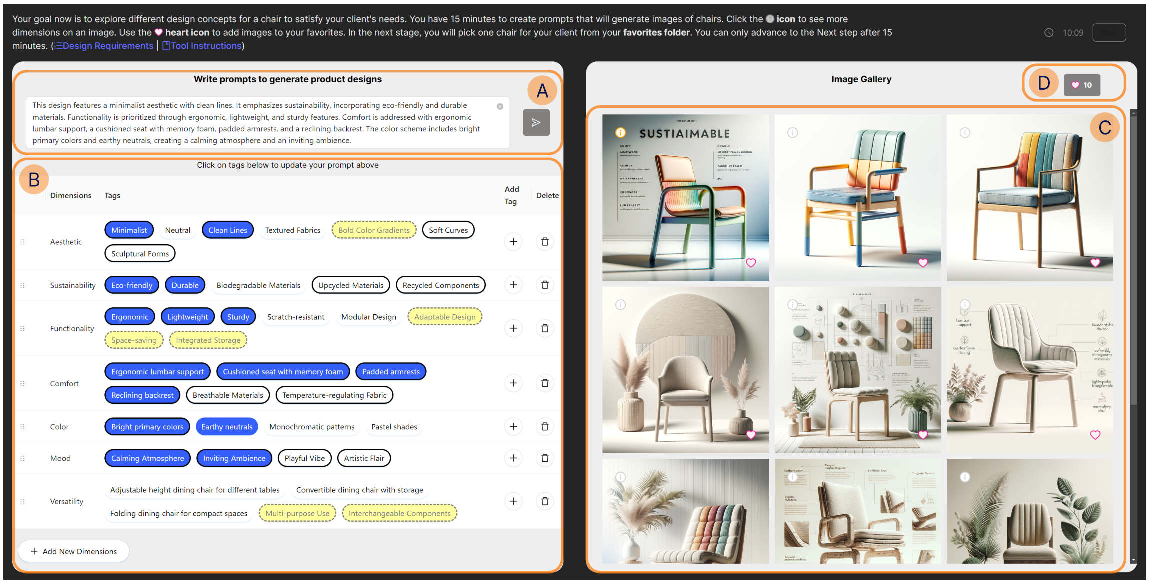

Fig. 1. *DesignWeaver*: An AI-enabled Product Design User Interface Designed for Novice. The components include (A) Prompt Box, (B) Dimension Palette, (C) Image Gallery, and (D) Favorite Folder

Generative AI has enabled novice designers to quickly create professional-looking visual representations for product concepts. However, novices also have limited domain knowledge that could constrain their ability to write prompts that effectively explore a product design space. To understand how experts explore and communicate about design spaces, we conducted a formative study with 12 experienced product designers and found that experts — and their clients less-versed in the domain — often use visual references to guide co-design discussions, in lieu of specific descriptive language. These insights inspired DesignWeaver, an interface that helps novices generate prompts for a text-to-image model by surfacing key design dimensions from generated images. In a study with 52 novices, DesignWeaver enabled participants to craft longer prompts with more domain-specific vocabularies, resulting in more diverse, innovative product designs. However, the more nuanced prompts raised participants' expectations, which were not met by current text-to-image models.

CCS Concepts: • **Human-centered computing** → **Empirical studies in interaction design**; **Text input**; *Graphical user interfaces*; *Collaborative interaction*; Natural language interfaces.

Additional Key Words and Phrases: Creativity support tools, design ideation, idea management, human-AI interaction, text-to-image models

1 INTRODUCTION

Recent advancements in Generative AI (GenAI) [13, 22, 48, 60] have revolutionized creative processes across various disciplines, enabling new possibilities for ideation and content generation [12, 26]. These models show significant

2025. Manuscript submitted to ACM

potential across multiple modalities, including text, images, 3D models, and video [47, 50, 54]. In the Human-Computer Interaction (HCI) community, efforts to improve interactivity with GenAI have focused on refining prompt engineering interfaces, such as structuring prompts through familiar templates [11, 65, 66], studying how designers craft prompts [10, 43], and leveraging visual manipulation of GenAI-generated images [6, 69], as well as iterative refinement using large language models chaining [14, 35, 64].

Recently, researchers have explore using text-to-image (T2I) models to generate renderings for product design concepts [25, 36]. Leveraging T2I models to generate visuals for product designs places the burden on users to write prompts that embed specific details about the product space and different design dimensions, while also grappling with the challenge of articulating tacit preferences and design knowledge in natural language [10]. Also, studies like DesignPrompt [46] show that supporting prompt engineering through bespoke UIs —such as through multi-modal interactions—can introduce unpredictability and hinder practical use. Even expert designers cultivate the ability to navigate complex design spaces by understanding design principles, technical details, and available options within a specific domain, such as home products. This expertise allows them to effectively balance high-level creative goals with the technical constraints of their tools, ensuring that their prompts or instructions are clear and aligned with the system's capabilities [63]. As for novice designers, even when experienced at writing text prompts for AI, often lack an understanding of complex concepts and specialized vocabulary within a design domain [42, 68].

To understand how experts explore a design space and communicate with their clients who generally lack domain knowledge, we conducted a formative study and interviewed twelve experienced designers (2-20 years) from diverse fields on how they communicate about key dimensions and options with clients. We found that designers often presented multiple alternatives to provide choices for their clients while maintaining their own creative control. Both the experts and clients rely on visual representations to articulate and negotiate a vision. This suggests tools should allow novices to grapple with design concepts and choices by emphasizing visual representations, rather than requiring precise language.

Based on these insights, we developed *DesignWeaver*, an interface powered by GPT-4 and DALL-e 3, designed to support novice designers through a technique we call *dimensional scaffolding* (surfacing key design dimensions based on curated images to the guide the design process). *DesignWeaver* enables users to generate detailed prompts by extracting key design dimensions from curated images as visual references. Users can use these extracted design dimensions as a palette – toggling which ones to include vs not, adding new dimensions and changing by inspecting output images. Key dimensions, such as geometry, style, color, and material, emerge organically through bidirectional interaction between text prompts and generated images. We hypothesize this technique helps users discover and adjust key dimensions, compensating for their lack of domain-specific knowledge, leading to more effective and creative prompts, while avoiding cognitive overload from advanced AI features.

To evaluate the effectiveness of *dimensional scaffolding*, we focused on three key research questions:

(1) **How does dimensional scaffolding affect the quality, length, and use of domain-specific language in text prompts compared to, compared to a standard text-based interface?**
(2) **How do any differences in the text prompts enabled by dimensional scaffolding affect the diversity and quality of generated product designs compared to standard text input?**
(3) **How do participants engage with dimensional scaffolding, and what are their experiences using this feature?**

To investigate these research questions, we conducted a between-subjects study (n=52), where participants were asked to design a chair based on a design brief using either *DesignWeaver* or a text-based prompting interface (like ChatGPT). o gather quantitative and qualitative insights, we collected user-generated design artifacts, system logs, survey ratings, and semi-structured interviews.

Our findings show that *DesignWeaver*'s *dimensional scaffolding* enabled users to confidently explore various design combinations, uncovering dimensions they might not have considered with a text-based prompting interface. Participants using *DesignWeaver* produced more detailed and nuanced prompts (measured by its length & substance - unique design vocabulary) from the outset. Dimensional scaffolding allows for the structuring of key design dimensions and smoother navigation between text and image, which leads to more nuanced prompts and, in turn, results in better design outcomes, as evaluated by CLIP-based similarity [48] and blind-to-condition expert designers. Participants using *DesignWeaver* added on average 3.1 new dimensions and 19.1 new tags to the palette, and included them in the design prompts them 25.7% of the time in every round. Participants reported that *DesignWeaver* gave them confidence to try different combinations and helped them tap into language they would not have used otherwise. However, we find the more nuanced prompts raised participants' expectations (measured by their satisfaction level with both the process and the result), which sometimes creates frustration when current T2I models can't fulfilled their more nuanced request.

Our core contributions are:

- A novel system, *DesignWeaver*, introducing dimensional scaffolding for iteratively authoring prompts and discovering key design dimensions through output inspection.
- Empirical quantitative and qualitative insights that suggest that *DesignWeaver* enhances prompt nuance, creative exploration, and design innovation, with AI-based image diversity analysis aligning closely with expert ratings.

2 RELATED WORK

2.1 Designers need to explore Design Space Dimensions. But it's hard, especially for novices

Exploring the design space is a critical part of the creative process, enabling designers to navigate through various possibilities before arriving at an optimal solution. Multiple studies have emphasized the importance of systematic exploration in creativity and design, highlighting its role in both problem-solving and innovation [15, 37] Design space exploration (DSE) involves identifying and testing various alternatives across dimensions such as form, function, aesthetics, and usability [67]. However, the complexity of design problems often makes the exploration process overwhelming, particularly for novices who may lack experience in identifying relevant dimensions and exploring trade-offs between solutions [38]. While expert designers can rely on tacit knowledge and heuristics, novices often struggle to understand the structure of the design space and how to traverse it effectively [1]. Research has shown that novices tend to focus on surface-level features and have difficulty framing problems and generating alternatives [2, 21]. This issue is exacerbated in complex domains like product design, where multiple conflicting constraints must be balanced [28].

Interactive tools have been proposed to support design exploration, with some success in guiding users through multi-dimensional trade-offs [18]. Visualization tools, for instance, allow designers to navigate and explore design dimensions [7, 56]. In this paper, *DesignWeaver* helps users visually explore the design space by generating and iterating potential solutions.

2.2 Visuals help with exploring Design Space dimensions

Visuals play a pivotal role in design space exploration by bridging the gap between abstract ideas and tangible solutions. Prototypes such as sketches, diagrams, and digital models, help designers articulate and explore complex design problems [7]. The use of visuals allows designers to externalize their thought processes, making tacit knowledge more explicit and facilitating better communication and collaboration [8, 20]. For instance, early design sketches can capture a wide range of possibilities and support iterative exploration by providing a visual reference for evaluating different design alternatives [20]. Conceptual diagrams and storyboards enable designers to organize and manipulate information, which can lead to deeper insights and more informed decisions [55]. The iterative nature of prototyping allows designers to explore various design dimensions, including functionality, usability, and aesthetics, and discover unforeseen issues and opportunities, which can be difficult to identify through conceptual thinking alone [16].

Recent advancements in Generative AI (GenAI) have significantly expanded the possibilities for generating visual content in the design process. Tools like DALL·E [50], Stable Diffusion [51], and MidJourney have demonstrated tremendous potential in producing high-quality images from textual descriptions, which can help accelerate design exploration and visualization [13, 22, 49]. These tools can generate content across various modalities, including text [3, 41, 58, 59], images [5, 9, 33, 34, 50–52], 3D models [19, 29–31, 40, 47, 53, 70], and video [4, 23, 27, 54, 61].

While these tools have enhanced creative autonomy, they still require users to construct precise prompts, which poses a challenge for novice designers. Additionally, in the field of Human-Computer Interaction (HCI) and graphics, image-image interaction techniques such as in-painting, point-based manipulation, and sketch-based control have been explored to enhance user control over AI-generated outputs without relying solely on textual inputs [24, 45, 69]. These approaches allow designers to refine outputs visually, providing more flexible and interactive methods of engaging with AI-generated images. Maybe you generate lots of bad alternatives that land in "valleys" not on "hills" in the notion of stimulated annealing [39]. Therefore, despite these advances, novice designers still encounter difficulties in refining prompts or controlling the generated outputs through purely visual methods, underscoring the need for better tools that integrate both visual and text-based interactions to support more intuitive design exploration.

2.3 Prompt Engineering for Dimension-Based Design

Prompt engineering is a critical component of GenAI interactions, as it directly influences the relevance and quality of the generated outputs. While general-purpose AI systems have made significant strides in supporting prompt crafting, they often lack the specificity required for dimension-based design exploration, where users must balance multiple design aspects like geometry, style, and functionality. Novice designers are particularly disadvantaged in these tasks, as they struggle to manage these competing design considerations. Novices struggle to construct meaningful prompts and often lack the domain-specific knowledge required to generate effective outputs [42, 68]. Novices may also find themselves overwhelmed by the complexity of generative AI systems, leading to suboptimal or irrelevant results. Even expert designers face the problem of abstraction matching: when the user has a well-formed intent, how do they select an utterance from the near infinite space of naturalistic utterances that they believe the system will reliably map to a satisfactory solution? This involves “matching” the utterance to the right level of “abstraction”, by specifying the utterance at a level of granularity and detail that matches the set of actions the system can take, and selecting suitable words and grammar [32]. A study of how professional designers craft prompts shows that they focus more on the strategies and practices involved in prompt crafting, while novice designers often struggle with surfacing key design dimensions and lack effective prompt writing practices [10]. This gap in prompt crafting knowledge among novices presents a major barrier to utilizing AI tools effectively.

Efforts in HCI have focused on developing UI interfaces for prompt engineering, which assist users in crafting more effective prompts by providing structured input fields or offering suggestions [6, 56]. For example, systems like Jamplate structure AI prompts using familiar templates, helping guide users through the process of generating effective outputs without fully surfacing key design dimensions [66]. These tools simplify prompt creation, but still do not provide comprehensive guidance for dimension-based exploration. In addition to structured UIs, the concept of LLM chaining has been explored, where users iteratively refine prompts and outputs by chaining multiple steps together [14, 35, 64]. This multi-step approach allows for more complex interactions and refinement of generated content, supporting more sophisticated design workflows. However, even these advancements do not fully address the needs of novice designers, who require deeper support for balancing and exploring multiple design dimensions simultaneously. In this paper, *DesignWeaver* integrates both visual and text-based interactions to support more intuitive design exploration – surfacing key design dimensions based on curated images to the guide the design process.

3 FORMATIVE STUDY

This study aims to understand how designers explore options to meet client goals, communicate with clients to clarify preferences, and negotiate constraints to converge on a solution for further development. We conducted semi-structured, one-on-one interviews with twelve experienced designers, who specialize in diverse design fields, had professional experience ranging from over two years in furniture design to more than twenty years in speculative and architectural design (Table 1). Their roles spanned design studios, corporate settings, and independent design shops, with responsibilities including managing teams and projects in consumer electronics, medical devices, startups, bespoke furniture, and outdoor environments. The recruitment process employed targeted outreach on LinkedIn, Reddit, and Discord, specifically within design-focused subreddits and community group chats. Interested individuals were invited to sign up through a Google Form. The interviews, conducted via Zoom, lasted between 30 to 60 minutes. The interviews covered questions related to their design processes, client interactions, key considerations, and challenges faced, emphasizing their strategies for managing trade-offs, constraints, and client feedback to navigate project obstacles. Qualitative analysis was conducted on the Zoom session transcripts using an open coding scheme, with quotes grouped into themes based on recurring patterns.

Table 1. Details and characteristics about participants in the formative study are shown.

Experts	**Years of Experience**	**Design Domain(s)**
E3, E8	more than 2 years	Product, Furniture Design
E9, E10	more than 3 years	Product, Furniture Design
E11, E12	more than 5 years	Landscape Architectural, Industrial Design
E1, E4, E7	more than 10 years	Architectural, Product, Furniture, Industrial Design
E5, E6	more than 15 years	Product, Industrial Design
E2	more than 20 years	Speculative, Architectural Design

3.1 *Insight 1: Clients often struggle to articulate specific preferences and rely on visuals to express ideas.*

[16, 57] Few clients can clearly articulate their vision with specific preferences, often turning to visual aids or physical materials to communicate their ideas more effectively. As *[E3]* points out, *"Sometimes, on top of rough sketches and photos of inspiration, clients even provide fabric swatches, color samples, or other physical materials to guide the design process."* These visuals help bridge the gap between vague ideas and concrete preferences. Designers, in turn, must carefully analyze these materials to *"identify key themes, categorize preferences, and highlight main priorities."* When given a visual representation of a concept, clients provide more specific feedback, as *[E4]* explains, *"When I send a sketch to the client, they might circle parts they don't like and say, 'Okay, I don't like this.'"* However, even with the help of visual aids, *"it can still be difficult for clients to fully communicate their needs,"* which is why designers often offer additional guidance on overlooked aspects like *"environmental impact and cost implications" [E9]*. Some clients go further by bringing external inspirations to the table, as *[E8]* highlights: *"Some clients bring pictures from places like France or from friends, and others share videos of the furniture they want. When designing together, we sketch and discuss details over the phone, finalizing them in person."* Visual tools are also crucial in preventing unwanted outcomes, as *[E7]* emphasizes the importance of avoiding negative surprises, *"with a product that surprised you, but not in a good way,"* where clients may react with *"Oh, I wasn't expecting this, and this doesn't feel right."*

3.2 *Insight 2: Designers present multiple visual alternatives to support tacit communication about a design space.*

This approach typically starts with *[E3] "exploring various forms and ideas before narrowing down"* to the most feasible ones. This initial exploration leads to the presentation of *[E5] "three strong options, detailing why each one is specific and discussing one standout option."* By presenting multiple alternatives, designers can involve clients early in the design process. As *[E6]* notes, *"we impress clients by involving them throughout the brainstorming session"* as *"we display ideas on the wall and present numerous concepts that we feel strongly about."* This method allows for the sharing of creative ideas and timely reference materials, while quickly ruling out what works and what doesn't for better outcomes. *[E7]* further validates the claim by engaging in this activity *"before moving to a full-scale model."* This method also helps designers and clients identify key constraints, such as budget and materials, and assess where compromises can be made. *[E3] "By offering different design options that vary in cost, materials, and aesthetics, this helps to identify what aspects are non-negotiable for the client and where they are willing to make compromises. For instance, if a particular material is too expensive, we present a more affordable alternative that still aligns with their design goals."* With a lack of control, designers often face challenges when moving into the manufacturing phase, where decisions about form and shape can cause regret. *[E6]* reflects, *"In hindsight, we should have designed a different product to simplify the manufacturing process… ensuring the texture and color were correct was notably difficult."* When collaborating with engineering and manufacturing teams, where the high cost of prototyping—often in the thousands—demands efficiency, effective communication and refinement of choices before implementation is even more crucial: *[E6] "Through meticulous refinement, we successfully achieved a workable model by reducing prototyping rounds from three to one. This efficiency not only saved costs but also expedited the development process."* Throughout their design process, several designers consistently preferred displaying typically three versions with detailed breakdowns to highlight the impact of individual changes on the overall project, aiming to exceed client expectations and minimize major revisions.

3.3 *Insight 3: Designers convey trade-offs to educate and manage unrealistic expectations.*

As noted by *[E11] "I will have to work within their budget, but they will also be more idealistic. Because sometimes I find that an idea, and the practicality of it at times don't go hand in hand."* For physical products, especially furniture, space constraints often require adjustments to form and materials to ensure suitability: *[E10] "I may find the need to communicate that the size constraints they've mentioned are not aesthetically suitable or practical for their space."* Social media can further complicate these discussions, as clients often want materials they've seen online, without considering practical factors. As highlighted by *[E11] "They want the exact same materials as seen on Tiktok or Instagram, and do not consider other things like their local environment or the most suitable materials, especially for plants."* Budget limitations are another frequent challenge, as clients often underestimate the costs of their desired designs, learning to misalignment between their expectations and the final deliverable. *[E4]* shared *"Someone will come to me demanding 'I want it in this color,' but their budget doesn't even meet half the actual cost."* Designers must guide clients through these decisions and explain why certain choices are necessary, often convincing them to trust their expertise and carefully consider the information before forming opinions. As *[E7]* noted, *"There are always situations where clients have specific preferences about the look or colors, and I need to help them understand that my choices are based on professional expertise and are the best fit for the project."* In many cases, clients prioritize appearance over durability, making it challenging for designers to convey the importance of longevity over aesthetics for a lasting, valuable purchase. Similarly, *[E10]* shared *"Sometimes, clients may ask for cheaper alternatives 'cause their initial budget is insufficient for the desired look."* Timelines and material availability create additional obstacles. As *[E3]* explained *"Certain materials may be unavailable or be too expensive, especially if the client has a specific aesthetic in mind. They often have deadlines for when they need the furniture, which can be tight due to events, renovations, or move-in dates,"* also supported by *[E10] "When clients see a chair they like in a restaurant, they ask for a similar one under 2 meters long. However, I find their request impractical for their space, and when I adjust the design, it exceeds their budget and deadlines, forcing me to redo the work."* Despite these challenges, most designers find that clients are typically satisfied with the final outcome as long as the design captures their original vision.

4 A SYSTEM FOR DIMENSIONAL SCAFFOLDING

4.1 Design Goals

The Formative Study revealed a significant challenge in client-designer interactions: clients often struggle to articulate what they like or dislike about their preferred product concepts, relying on images from the web or social media rather than using technical language. This leaves designers to interpret these visuals, decipher client preferences, and explain design decisions, all while helping clients navigate trade-offs in the design process. To address this, we aim to equip users with domain-specific language that enhances their ability to construct precise prompts for generating reference images. By capturing key dimensions and enabling a structured exploration of options, we seek to support designers—especially novices—not just in generating accurate product renderings through large language models (LLMs), but in navigating critical design aspects more systematically, fostering a more informed and thoughtful exploration of the design space.

Based on these findings and insights from theory and practice, we aim to create a tool with the following design goals:

- ***Goal 1: Absorb known preferences, requirements, and constraints.*** Establishing a context-aware foundation requires a thorough understanding of the client's identity, vision, and specific needs, which shape and inform the design process [44]. Gaining deep insight into who the client is, what they value, and their project goals, requirements, and constraints allows the designer to make decisions—whether about form, materials, or style—that align with the client's core intentions. By integrating these elements from the beginning, the designer ensures that every step of the process resonates with the client's objectives. Collecting tangible examples and personal inspirations of what is already known is crucial to developing a clear sense of the client's aesthetic preferences and functional requirements, as well as guiding the creative direction for the rest of the design process effectively.

- ***Goal 2: Surface specific dimensions from the product design space.*** Current prompt-generation tools often lack the precision and control designers need to tailor outputs effectively, especially when clients struggle to clearly articulate their preferences. This gap highlights the need for designers to surface and refine design dimensions such as form, texture, color, and style—key elements that often emerge only through visual feedback. Sometimes, realizing too late that a different approach would have simplified the manufacturing process or prevented client dissatisfaction can be avoided by anticipating the effects of design decisions early on. By enabling designers to adjust these dimensions in response to visual cues or client input, greater flexibility is achieved in exploring trade-offs and aligning the design with client expectations. This capability not only allows designers to fine-tune prompts based on visual representations but also helps them test different design scenarios in real-time, reducing unforeseen issues and minimizing the need for major revisions.
- ***Goal 3: Enable comparison though multiple visual representations.*** Presenting multiple alternative options simultaneously allows designers to explore a wider range of possibilities and engage clients early in the design process [16, 57]. By showcasing distinct visual variations that emphasize different elements, designers can clearly demonstrate how each choice impacts the final product, facilitating comparing and contrasting options quickly and easily. Additionally, this method provides designers with a way to justify their design choices through visual evidence, ensuring client alignment before moving into advanced stages like prototyping or full-scale modeling.
- ***Goal 4: Facilitate dimensional reasoning through trial and error.*** Novice designers often struggle to navigate complex design spaces due to a lack of domain knowledge and the language needed to clearly communicate ideas and trade-offs. Despite access to large language models (LLMs), this gap can hinder their ability to create effective prompts or engage meaningfully with clients. An iterative feedback loop, focused on trial-and-error experimentation, provides a solution by allowing novices to refine their understanding of design-specific language and principles. Through repeated practice, guided by visual references, they can explore alternatives in materials, form, and aesthetics while learning to articulate the trade-offs involved. This process not only builds the confidence necessary for clear client communication but also fosters a deeper understanding of how to balance creative vision with practical constraints.

4.2 *DesignWeaver* User Experience

In order to address the aforementioned design goals, we developed *DesignWeaver* with the following workflow shown in Figure 2. We will guide you through all the features in *DesignWeaver* using Figure 3.

In *DesignWeaver*, users interact mainly through two primary panels: the Design Panel on the left and the Image Panel on the right. Prior to starting, an initial Design Document is uploaded, detailing the client's persona, vision, specifications, budget, and timeline. This document serves as a crucial guide throughout the design process, helping users make informed decisions while using the tool.

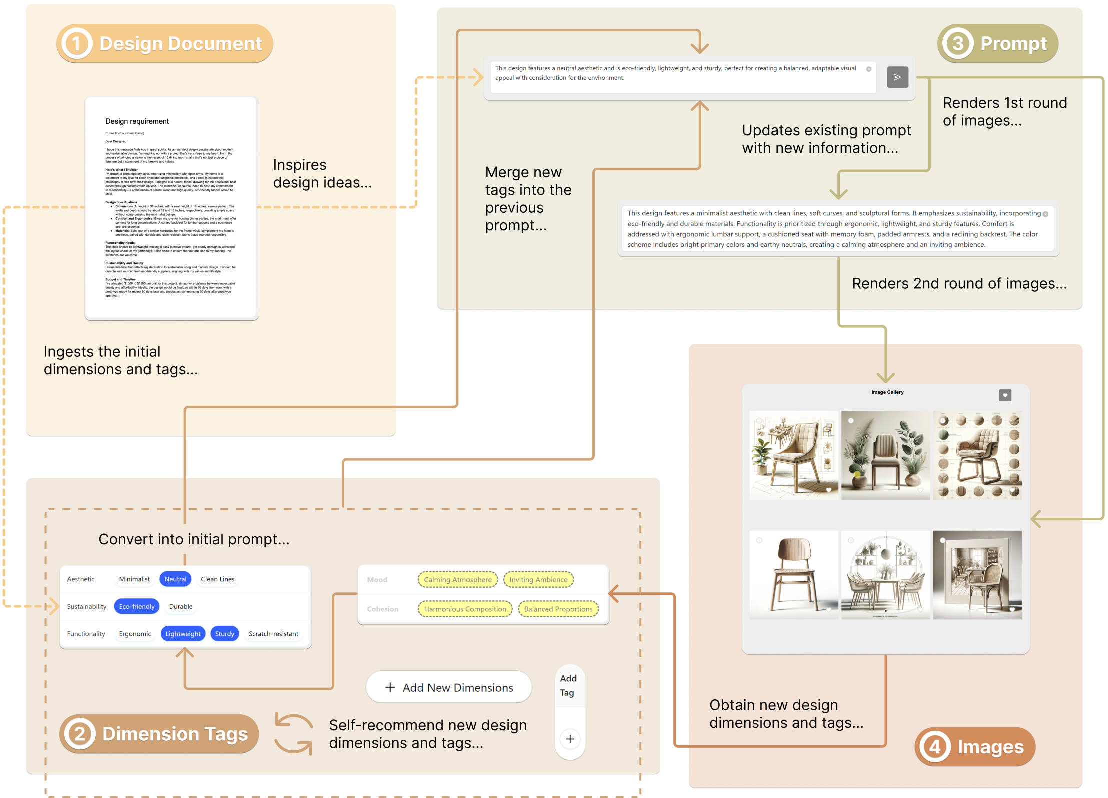

Fig. 2. Overview of the iterative design process using *DesignWeaver*.

4.2.1 Design Panel. The Design Panel is further divided into two sections: the Prompt Box at the top and the Dimension Palette at the bottom.

Prompt Box. The Prompt Box is a standard text input area that allows users to type prompts for image generation. Users could either describe their design vision in their own words or use keywords as inputs for the DALL-E 3 API to function. It includes a scroll bar for longer inputs and a Send Button on the right side to submit the prompt. Prompt Box introduces two additional features. The first is an auto-complete feature, which automatically completes any partially written input when activated. The second is a merging feature that combines user-written content with newly generated dimension tags (which will be explained in the next section).

Dimension Palette. The Dimension Palette is the most important feature that distinguishes *DesignWeaver* from other tools. It helps users discover and refine design dimensions by providing categorized style tags based on keywords extracted from the uploaded Design Document. In this context, the term dimension refers to a broad design concept or aspect, such as aesthetics, sustainability, or functionality, which is displayed as the title of each row in the Dimension Palette. For each dimension, such as "Aesthetics," there are associated subcategories like "modern," "classical," or "minimalism," which are located within the same row. These subcategories appear as clickable style tags. Users can add new tags to any row that they feel fits the existing dimension, delete tags, introduce new dimensions (which will

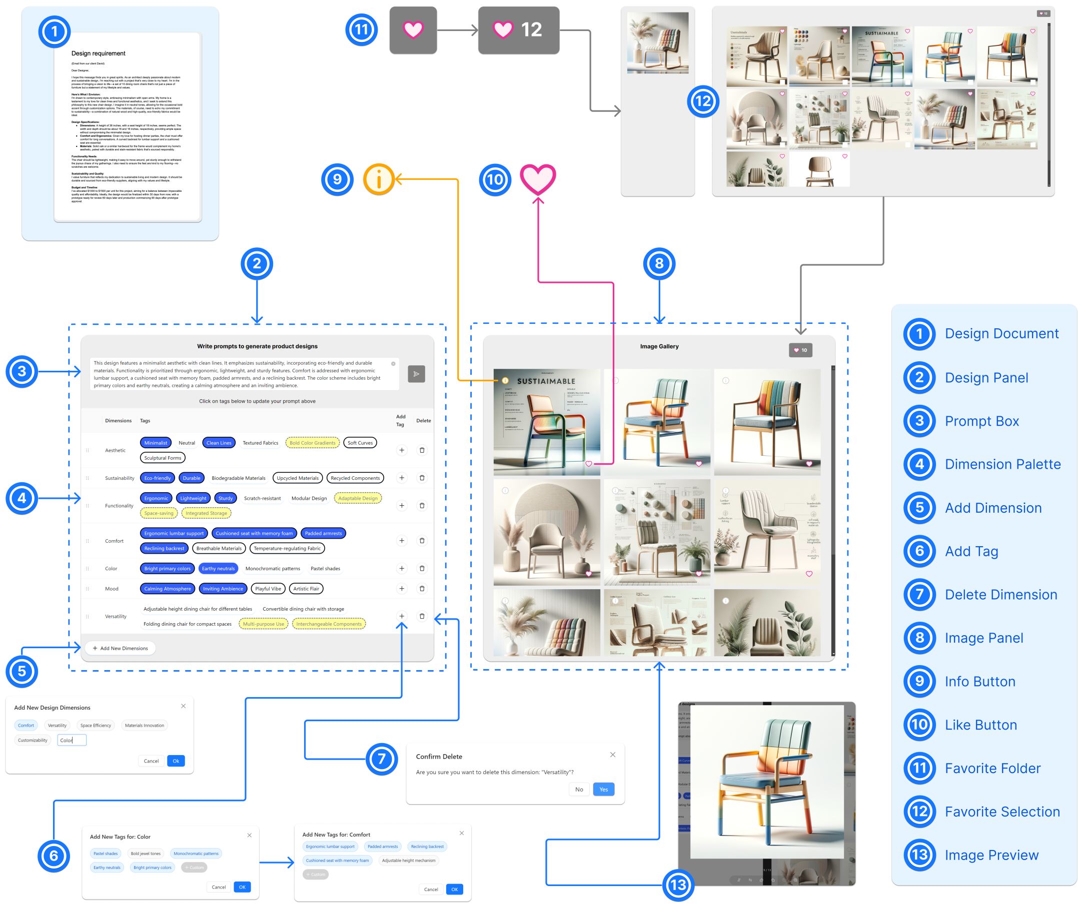

Fig. 3. User Interface of DesignWeaver.

create new rows), remove entire rows, and reorder the rows as needed. When adding new tags, users have the option to select from a list of LLM-generated recommendations or create custom tags. The same process applies when adding new dimensions. Single clicking on a style tag within a dimension changes the tag's color from its default white to bright blue, and simultaneously updates the prompt in the Prompt Box. For instance, selecting the tag "Minimalist" will automatically generate the phrase "The design embraces a minimalist aesthetic" in the Prompt Box. Users can continue to add more tags to further refine the prompt. Once satisfied, they can click the Send Button to generate images. After a brief moment, three images will appear on the right-hand side in the Image Panel, displayed in a row.

4.2.2 Image Panel. Users will start with an empty Image Panel. Each time the Send Button in the Prompt Box is pressed, a row of three images will be generated and displayed in the Image Panel. Each image includes two icons: a heart icon at the bottom right corner and an info icon at the bottom right corner. Users can click on an image to zoom in, zoom out, flip it for alternate perspectives, or expand it in the center of the screen for a detailed preview. After the initial three images are generated, users can modify the text in the Prompt Box and generate new images. This process is repeated for continuous iterations and exploration.

Like Button & Favorite Folder. The heart icon serves as a Like Button to like an image. Liked images are stored in the Favorite Folder, which is located at the top right corner of the panel for easy access. When users click on the Favorite Folder, they can view all the images they have favorited, allowing them to easily filter their preferred images from the entire collection of generated results and compare them with each other. While all images remain visible in the Image Panel by default, the Favorite Folder is a filtering tool that lets users toggle between viewing all pictures and only their favorite ones. To return to the full gallery, users simply click the Favorite Folder again.

Info Button. The info icon is an Info Icon that reveals the style tags associated with a specific image. When users click on it, relevant tags appear on the Dimension Palette, highlighted in bright yellow with dashed outlines to indicate their temporary status. If the tags already exist, they are highlighted with a bold yellow outline; if not, new yellow, dashed-outline tags will appear. A single click on a yellow tag adds it permanently to its respective row, while a double click will activate the tag, turning it bright blue and adding it to the Prompt Box as part of the evolving prompt. The prompt is updated only when a tag has been used. If users decide they no longer want to include the suggested tags from the selected image, they can simply unclick the Info Button to hide all the associated tags.

Users who explore the design space using *DesignWeaver* switch between the Design Panel and Image Panel, creating a natural design iteration loop. The Dimension Palette displays the design dimensions discovered and refined through previous prompt iterations, allowing users to continually improve the prompts for generating new images. The Image Panel, in turn, showcases the generated results, allowing users to explore and interact with the visuals. This iterative process of refining prompts and visual exploration deepens users' understanding of the design space, helping them discover valuable design dimensions and language used in prompting that they may not have previously known. As a result, they can craft more effective prompts and generate images that more accurately reflect their design vision.

4.3 Implementation Details

DesignWeaver is a web application built with React and powered by a Python backend. All logged data, including prompts, is stored in Firebase, while generated images are downloaded via the Python backend and uploaded to Google FireStore. Image generation is handled using OpenAI's DALL-e3 API. To extract language from images and recommend new style tags and dimensions, we utilize the "GPT-4o-mini" model for its speed, ensuring better user control. The "GPT-4o" model is specifically employed to process prompt texts, providing higher quality and stability in text formatting. In comparison to *DesignWeaver*, the baseline is a simpler replica of *DesignWeaver* with all the additional features removed, other than Prompt Box and Like Button & Favorite Folder so it mimics what a standard text-based prompting interface.

4.4 User Scenario

Non-learning based Aesthetic Discovery and Visualization: For novice designers, gaining clarity on their aesthetic preferences can be challenging without tangible examples. *DesignWeaver* assists users by visualizing various design options, enabling them to explore and identify their preferred styles. This visual feedback helps users develop a clearer understanding of their own design tastes and preferences, which is crucial for building design confidence and making informed decisions.

Playground to develop junior designer Skill Development and Experimentation: *DesignWeaver* acts as a playground for junior designers, offering a safe space to explore and experiment with different design dimensions, elements, and styles. By allowing novices to interactively engage with design parameters, the tool supports the development of critical design skills and encourages iterative learning. This hands-on approach helps junior designers quickly adapt to domain-specific knowledge and vocabulary, accelerating their growth and confidence in design practices.

5 USER STUDY

The chair serves as an ideal subject for our user study due to its universal familiarity and design complexity. As an everyday object, the chair is easily understood by novices, providing a practical entry point for exploring design processes. Simultaneously, the chair has been a well-established focus for designers and architects, particularly during their formative training, due to its rich history of variation and evolution across different design movements and eras. Chairs offer a diverse range of design dimensions—such as ergonomics, aesthetics, and material choices—making them a versatile subject that balances functional requirements with creative expression. [17] This combination of simplicity and potential for nuanced exploration allows DesignWeaver to effectively evaluate both novice comprehension and expert critique. By selecting chair design as our focus, we aim to assess how structured dimensional scaffolding can assist in bridging the gap between novice understanding and expert-level process, thus promoting more dynamic and meaningful interactions within tangible design tasks.

5.1 Participants

We recruited a group of 52 participants (19 to 31 years old; 36 females) in a hybrid format (5 remotely and 47 in person) from varying backgrounds and varying experience with design, large-language-model (LLM), and text-to-image-model (T2I). Regarding the respective experience level, Table 2 provides a comprehensive overview.

Table 2. Design, large-language-model, and text-to-image-model Experience

Design Experience	LLM Experience	T2I Experience
No prior experience (29)	Never used (6)	Never used (41)
Beginner (0-1 year) (20)	Tried a few times (24)	Tried a few times (7)
Intermediate (2-5 years) (2)	Few times per month (15)	Few times per month (2)
Advanced (>5 years) (1)	Weekly (6)	Somewhat often (Weekly) (2)
	Daily (1)	

5.2 Controlled Condition

To evaluate the effectiveness of DesignWeaver, we compare it in our study with a controlled various of Design Weaver. In the controlled condition users will have most of the features of DesignWeaver except the initial digestion on the design document, the info button on Image Panel to extract dimension, and the Dimension Palette on Design Panel.

5.3 Study Protocol

Participants were introduced to the study's background, completed a screening survey, and reviewed the consent form. Upon consenting, participants were assigned to either the control group or the *DesignWeaver* group, each receiving an introduction to their respective tools. Before beginning the tasks, a scenario overview was provided, followed by guidance on the tool features, as shown in Figure 4.

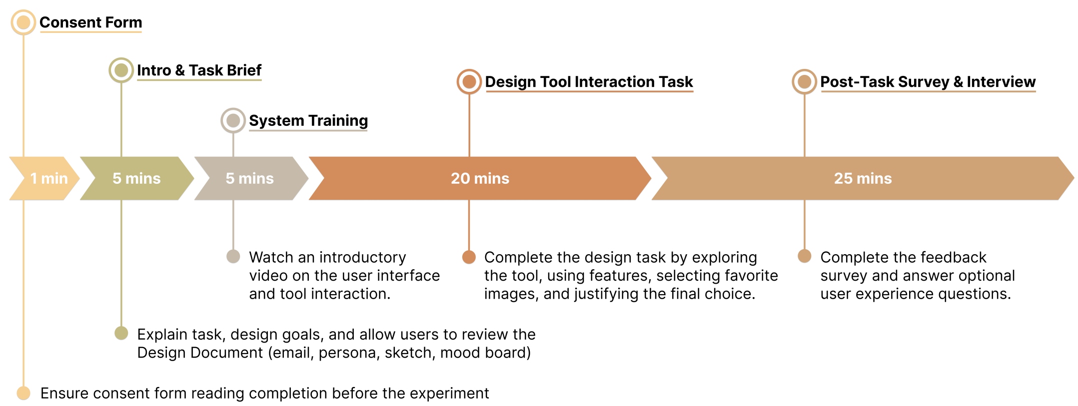

Fig. 4. Workflow of the user study.

5.3.1 Consent Form (1 minute). Before starting this IRB exempt experiment, participants were asked to review in advance the consent form the study's purpose, procedures, risks, and confidentiality measures without needing to sign. participants will receive a digital consent form after they signed up and we also confirmed with every participant about having already read through this before every experiment.

5.3.2 Introduction to the Scenario and Familiarization with Tools (10 minutes). After the initial survey, participants in both groups received documents, including a client email with design requirements, a client persona with detailed information, sketches of dining chair concepts, and a mood board. Participants had 5 minutes to review these materials, which remained accessible throughout the design process. Baseline group participants received a tutorial on the baseline tool, while *DesignWeaver* group participants were introduced to the *DesignWeaver* tool, focusing on tag selection and information reveal features. Both groups could access the tutorial slides during the design process to ensure familiarity with the tools. This introduction video stage also takes 5 minutes.

5.3.3 Design Exercise (20 minutes). Participants were then given 15 minutes, plus a 1-minute grace period, to generate images for a dining chair design. The Baseline group used the baseline tool with text prompts, while the *DesignWeaver* group utilized the *DesignWeaver*'s tag selection and information reveal features to craft prompts. Each prompt submission generated three images, and participants could view all generated images, selecting those they 'Liked' for the potential final design. An additional few minutes were provided after the exercise to finalize selections. After the time is up, they will have a few minutes to pick the final image from all the liked images.

5.3.4 Post-Survey and Interview (20-30 minutes). After completing the tasks, participants filled out feedback surveys tailored to their group experience, assessing criteria for final design selection, alignment with client needs, and understanding of design dimensions. The surveys evaluated satisfaction with the tool's outputs, ease of converting ideas into prompts, and overall support in visualizing and refining concepts. They will elaborate on their likert question response by thinking out aloud. The audio is automatically transcribed by used for later quantitative and qualitative analysis. This was followed by an in-depth interview exploring participants' experiences, design processes, tool effectiveness, and any challenges encountered. Feedback from surveys and interviews offered insights into tool interaction and highlighted areas for improvement.

5.4 Data Collection

We gathered data from three main sources throughout the experiment. First, we collected the prompts used and images generated by each participant, along with the dimensions and tags used by each *DesignWeaver* participant. We also tracked the time taken for each generation iteration. In terms of surveys, participants completed a screening survey that collected demographic and background information, including age, gender, language proficiency, and experience with design, LLMs, and T2I models. Following the experiment, participants responded to a post-experiment survey, which included Likert-scale ratings (from 1-7, Strongly Disagree to Strongly Agree) and open-ended questions, covering a broad range of topics we aimed to explore.

Additionally, we recorded audio transcriptions of participants' verbal responses during the Likert-scale ratings and collected elaborations through a post-experiment open-ended interview. To assess the final images selected by participants, we gathered ratings from six design experts (2-5 years of design experience), who rated the novelty and alignment with the client's request on a scale from 1 to 7 (strongly dislike to strongly like).

5.5 Data Analysis

5.5.1 Quantitative. We compared the numerical difference between prompt length, diversity (number of design words it contain), iteration time, and quantitative part of the survey.

For design term extraction, we developed a customized Natural Language Processing (NLP) pipeline with four key components:

- **Stopword Filtering:** Standard preprocessing (lowercasing, stopword removal, lemmatization) plus a custom list to filter non-design terms.
- **POS Tagging:** Extracted *nouns* and *adjectives* representing materials, aesthetics, and key design attributes.
- **TF-IDF:** Ranked terms by importance, filtering out common words and highlighting key design terms.
- **Word Embeddings:** Used Word2Vec to ensure extracted terms aligned with a design-specific vocabulary.

There are 3 sources we analyzed the design terms - 1) design document, 2) prompts produced, and 3) the audio transcriptions.

Additionally, we used the ViT-B/32 CLIP model [48] to calculated both image and prompt semantic similarities for the Baseline and DesignWeaver conditions across iterations. We also added in Levenshtein Distances to measure the words edits in each iteration.

5.5.2 Qualitative. Two authors used Zoom's automatic transcription[1] to transcribe and verify the recorded audio into text scripts for qualitative data, including think-aloud elaborations and open-ended user interviews. They then performed a thematic analysis of these transcripts. The extracted insights and user patterns are reported in the following result presentation.

[1]https://support.zoom.com/hc/en/article?id=zm_kb&sysparm_article=KB0064927

5.6 Data Exclusion Criteria

One user from control condition was excluded due to directly pasting a large portion of the design document directly into the prompt box, which created outliers when analyzing prompt. This removed participant reduced the number of participants in control group from 26 to 25.

6 FINDINGS

6.1 Impact of Dimensional Scaffolding on Prompt Behavior

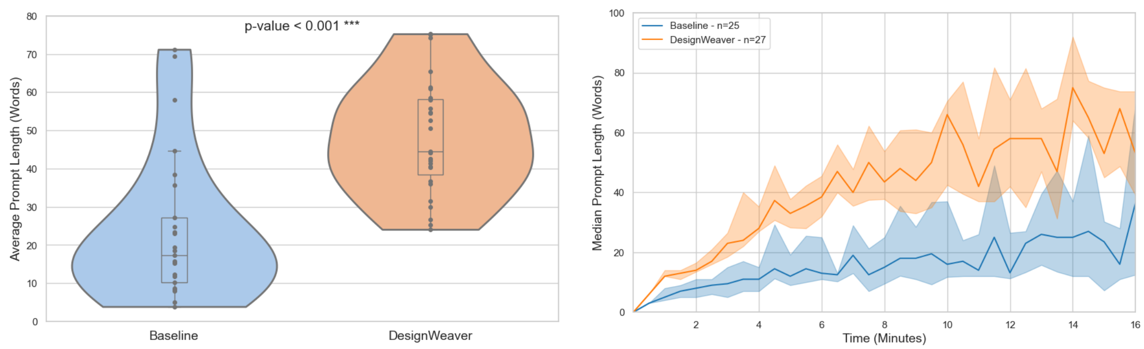

Fig. 5. DesignWeaver has longer average prompt (Left) and its median is consistently higher overtime (Right).

6.1.1 *Longer Prompts with Richer Design Vocabulary.*

Participants using *DesignWeaver* produced significantly longer prompts (M = 48.22, SD = 15.03) compared to the Baseline group (M = 23.73, SD = 19.12), with a statistically significant difference ($U = 97.000, p < 0.001$), as illustrated in Figure 5 (left). Analysis of prompt length over time showed that participants in the DesignWeaver group consistently expanded their prompts, with a steady increase in median prompt length throughout the session(Figure 5 (right)). In contrast, participants in the Baseline group reached a point of stabilization earlier, suggesting a more constrained approach to prompt development and limited iterative exploration.

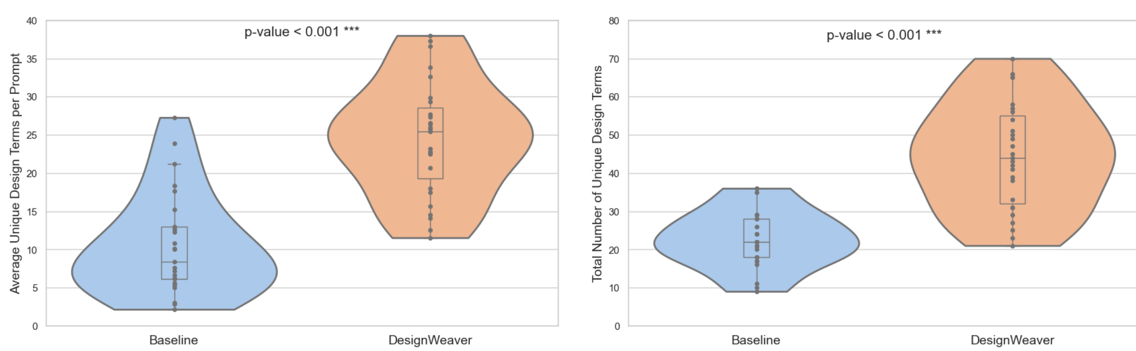

Fig. 6. DesignWeaver has more unique design terms per prompt on average (Left) and in total (Right).

In addition to generating longer prompts, participants in the DesignWeaver group employed a broader range of unique design terms. On average, participants in the DesignWeaver group used 24.48 unique terms per prompt (SD =

7.47), compared to 10.59 unique terms in the Baseline group (SD = 6.72). The total number of unique tags per session was also higher in the DesignWeaver group (M = 44.44, SD = 14.07) than in the Baseline group (M = 22.72, SD = 7.50), with significant differences confirmed by the analysis ($U = 61.0000, p < 0.001$ & $U = 56.5000, p < 0.001$), as illustrated in Figure 6.

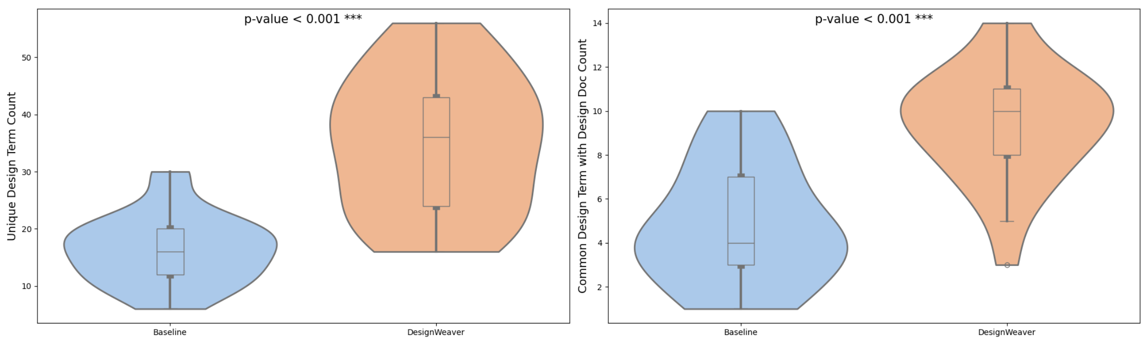

Fig. 7. DesignWeaver generated more unique design terms (left) and adapted more common terms from the design document (right).

To assess the impact of dimensional scaffolding on design vocabulary development, we compared, between the two groups, how many of the design terms are directly adapted from the Design document and how many are developed by the participants themselves while using the tool. Participants using *DesignWeaver* adopted both unique terms ($M = 34.8, SD = 12.0$) and common terms ($M = 9.4, SD = 2.5$) from the Design document more frequently than those in the Baseline group, who used unique terms ($M = 16.4, SD = 6.0$) and common terms ($M = 4.8, SD = 2.7$), as shown in Figure 7. Statistically significant differences were observed in the adoption rates for unique terms ($U = 619.0, p < 0.001$) and common terms ($U = 596.0, p < 0.001$), demonstrating that dimensional scaffolding facilitated not only the use of terms from the Design document but also encouraged participants to learn and adopt new design terms.

6.1.2 Expanded Prompt Strategies with Dimension Scaffolding. Participants using *DesignWeaver* employed a diverse set of strategies during prompt crafting, integrating dimension scaffolding to guide their design process. Initially, all participants(27/27) started with automated prompts generated from selecting tags based on the Design document. Throughout the process, they employed three main strategies: 18 out of 27 used default tags from the discussion panel, 5 out of 257 created custom dimensions and tags, and 12 out of 27 manually edited their prompts. Additionally, 6 out of 27 participants used the information button to map images back to dimension tags, gaining insights into how specific elements influenced the design. Notably, 10 out of 25 participants switched between these strategies. For example, P24 began with default tags such as "Minimalist," "Eco-friendly," " Ergonomic," and "Scratch-resistant." After reviewing the initial images, they added new dimensions and tags from the recommended list. As they continued, P24 customized tags using terms from the design document, like "stain-resistant," and further refined their prompts by manually adding specific details such as "Width is 18 inches and depth is 16 inches." Adopting expanding strategies, participants using *DesignWeaver* found it significantly easier($U = 236.5, p = 0.0296$) to convert ideas into prompts, as reflected in their higher ratings ($M = 5.63, SD = 1.11$) compared to the baseline group ($M = 4.88, SD = 1.51$)(Figure 8).

In contrast, all participants (25/25) in baseline group relied heavily on vocabulary from design documents and adjusted their prompts iteratively based on visual feedback from generated images. For straightforward requirements like "Solid oak", participants easily incorporated terms directly from the design documents. For example, P2 stated, "*I* *mainly just copied and pasted from the document and... included those into the prompt.*" When dealing with more complex requirements, such as ensuring comfort in long conversations, 24 out of 25 participants started with general terms and refined them based on feedback. As P21 explained, they began with "*neutral colored with a cushion seat,*" then added features such as like wood material and color accents based on how images aligned with their vision, iteratively refining to achieve a closer match. Such reliance on image feedback was reported to lead to participants' frustration when images failed to improve, 9 of the 25 participants reported this experience. For instance, P46 noted, "*while I'm... adding more information or... keywords to it, but sometimes... some keywords later, it doesn't change that much. It just gives me some new pictures, but I can't really... tell where... there's a big change from.*"

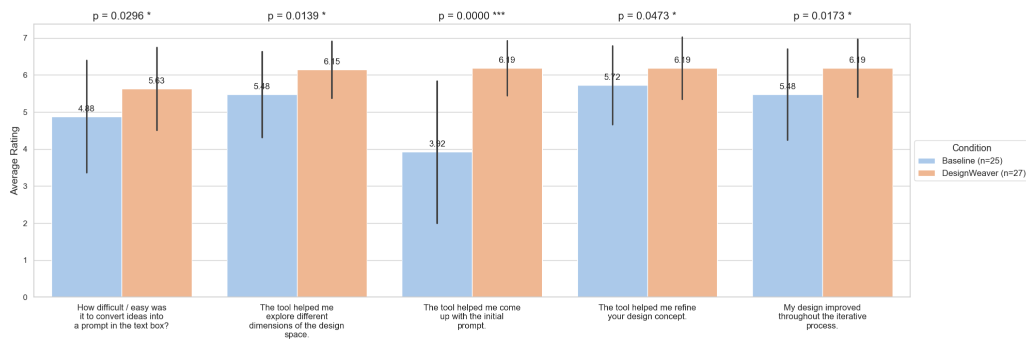

Fig. 8. DesignWeaver consistently got higher ratings in these survey questions.

6.2 Impact of Dimensional Scaffolding on Image Results

Using the ViT-B/32 CLIP model [48], we calculated both image and prompt similarities for the Baseline and DesignWeaver conditions. Results show that:

- Image Similarity (Figure 9): Images in the *DesignWeaver* condition were less semantically similar, suggesting more diverse visual outputs in *DesignWeaver*.
- Prompt Similarity (Figure 10): Prompts in the *DesignWeaver* condition were more semantically similar.
- Distributions (Figure 11): Both image and prompt similarity scores were significantly different from each other group (p<0.001). In image similarity, Baseline has mean similarity of 0.903 (SD=0.015) and *DesignWeaver* has similarity of 0.863 (SD=0.024), which yields a statistically significant difference ($U = 521.0, p < 0.001$). In prompt similarity, Baseline has mean similarity of 0.916 (SD=0.035) and *DesignWeaver* has similarity of 0.964 (SD=0.022), which also yields a statistically significant difference ($U = 78.0, p < 0.001$).

To further support our conclusion that semantically similar prompts in *DesignWeaver* yield more semantically diverse image outcomes and to address potential concerns, we performed a simpler prompt difference comparison using the Levenshtein edit distance. The results in Figure 12 indicate that users in the *DesignWeaver* condition ($M = 174.49, SD = 74.83$) made significantly more prompt modifications compared to the Baseline Condition ($M = 68.21, SD = 44.94$). A Mann-Whitney U Test, shown in Figure 13, produced a U statistic of 76.0 and a p-value less than 0.001, confirming the robustness of our findings.

Lastly, we compared the experts rating data we gathered on novelty and value alignment for the final images picked by all the participants. We found that *DesignWeaver* condition was rated significantly higher on image novelty

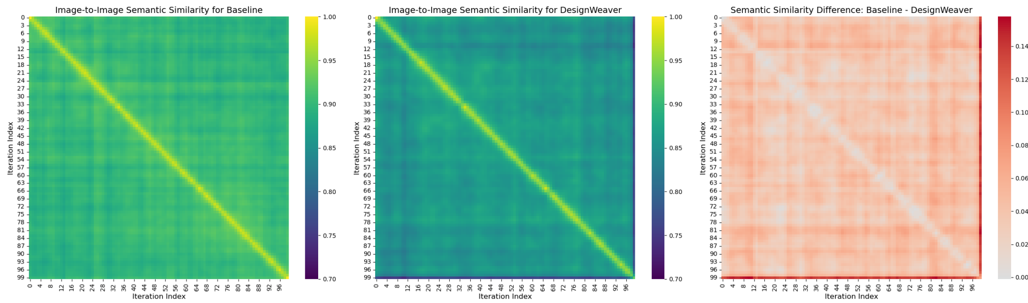

Fig. 9. Image semantic similarity heatmaps for baseline (left), DesignWeaver (middle), and their differences (right). Baseline has higher Image Semantic Similarity.

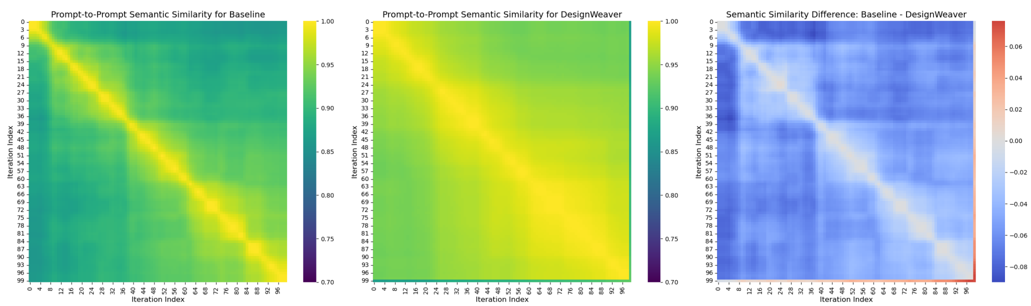

Fig. 10. Prompt similarity heatmaps for baseline (left), DesignWeaver (middle), and their differences (right). DesignWeaver Prompt has higher semantic similarity.

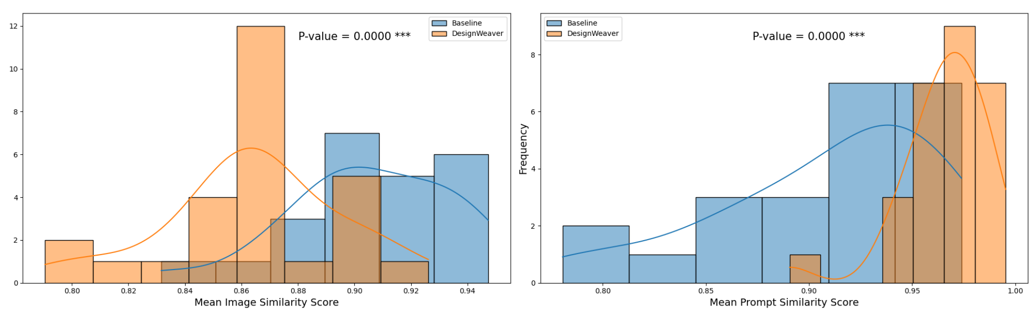

Fig. 11. DesignWeaver has higher semantic image difference (left) & lower semantic prompt difference(right).

($M = 4.09, SD = 1.63$) compared to the Baseline group ($M = 3.54, SD = 1.60$), with a statistically significant difference ($U = 9899.0, p = 0.002$), as illustrated in Figure 14 (left). Analysis of on the alignment with design requirements didn't show a significant difference between *DesignWeaver* condition ($M = 4.5, SD = 1.656$) and Baseline condition ($M = 4.2, SD = 1.67$) with a statistical significance at 0.059 ($U = 10925.5$) (Figure 14 (right)). Figure 15 showcases the top 5 chairs rated by experts for their innovation.

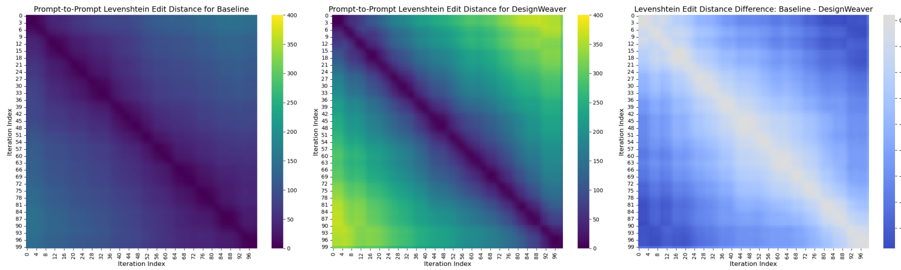

Fig. 12. Prompt Levenshtein edit distance Heatmaps for baseline (left), DesignWeaver (middle), and their differences (right). DesignWeaver has higher prompt Levenshtein edit distance.

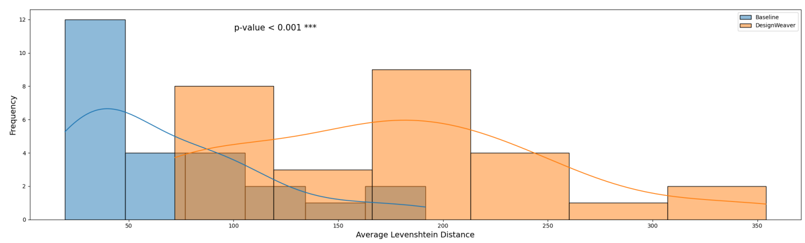

Fig. 13. DesignWeaver group has more text edit in prompts across iterations.

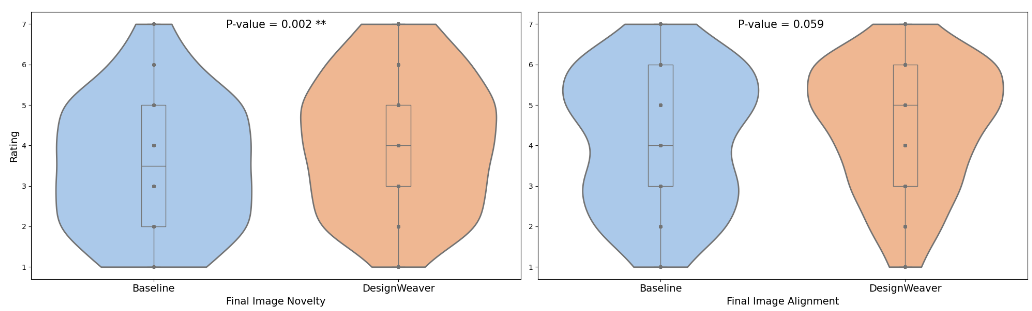

Fig. 14. Expert Rate DesignWeaver only higher on Novelty

6.3 Perceived Support and Challenges from *DesignWeaver*

Participants reported various aspects of their experience with *DesignWeaver*, highlighting both its supportive features and areas where challenges were encountered.

Firstly, *DesignWeaver* was perceived to significantly aid in adapting new design terms, as evidenced by significantly higher($U = 225.0, p = 0.0139$) ratings from participants using *DesignWeaver* ($M = 6.15, SD = 0.77$) compared to the baseline group ($M = 5.48, SD = 1.16$) for exploring different design dimensions((Figure 8)). This adaptability was noticeable from the initial stages of the design process, where *DesignWeaver* generated images that closely aligned with participants' early visions. Participants found *DesignWeaver* highly effective in helping create initial prompts ($M = 6.19, SD = 0.74$), significantly outperforming($U = 115.0, p < 0.001$) the baseline group ($M = 3.92, SD = 1.91$). For instance, P7 remarked, "*it immediately generated something that was very similar to the mood board,*" while P11 noted

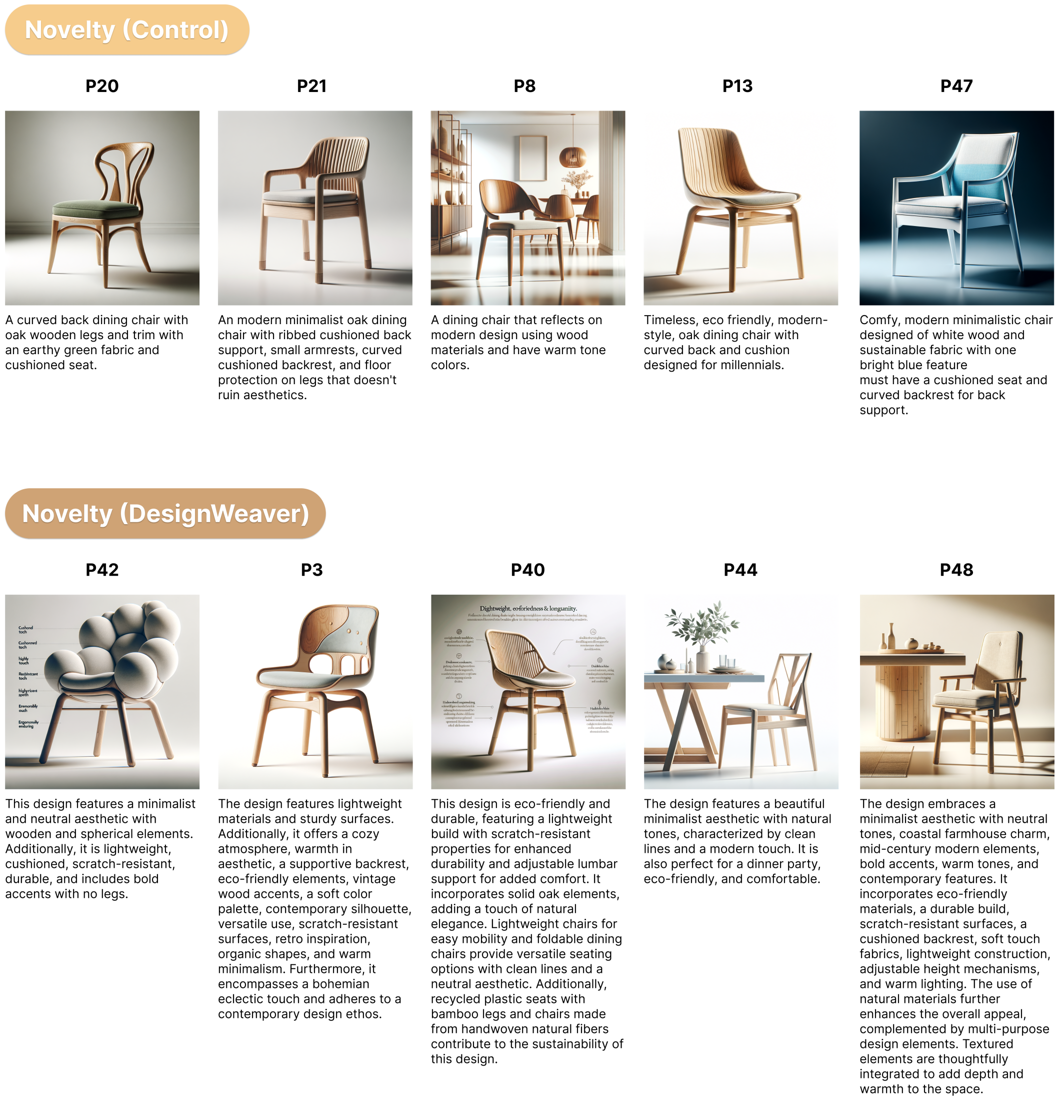

Fig. 15. Top 5 expert rated chair on innovation

that it "*helped generate a sort of baseline on what chairs I could generate.*" Throughout the design process, the tags listed on the panel were instrumental in broadening participants' design exploration, with 8 participants appreciating how these tags introduced new dimensions and terminology they hadn't previously considered. P20 highlighted, "*there were lots of different dimensions that I wouldn't typically think about for chairs. So that was like new information.*" Similarly, P13 valued the exposure to terms like "ergonomic" and "sustainability," which allowed them to refine and expand their design ideas, enhancing their creative versatility.

Participants also felt that *DesignWeaver* provided control over the design process, allowing them to manage their designs through tag selection. 8 of the 27 participants mentioned that tag selection helped them efficiently navigate and refine their designs. P1 described their experience: "*I was able to change the chair and select different [options] that fit the design... and go back if I wanted to add or change... tags.*" Among these 8 participants, 6 participants also highlighted the information button as crucial for staying informed about their current design state and understanding how each modification impacted their design within specific dimensions. P30 explained, "*Whenever I was getting info from the chairs that I liked, it was helpful to see what wording or characteristics I should keep in mind.*" This sense of control is reflected in the significantly higher ratings from participants using *DesignWeaver* ($M = 6.19, SD = 0.74$) compared to the baseline group ($M = 5.72, SD = 1.06$) for refining their design concepts ($U = 251.5, p = 0.047$) (Figure 8). Additionally, participants using *DesignWeaver* also felt their designs improved throughout the iterative process ($M = 6.19, SD = 0.79$), which was also significantly higher than the baseline group's ratings ($M = 5.48, SD = 1.23$; $U = 226.5, p = 0.017$)(Figure 8).

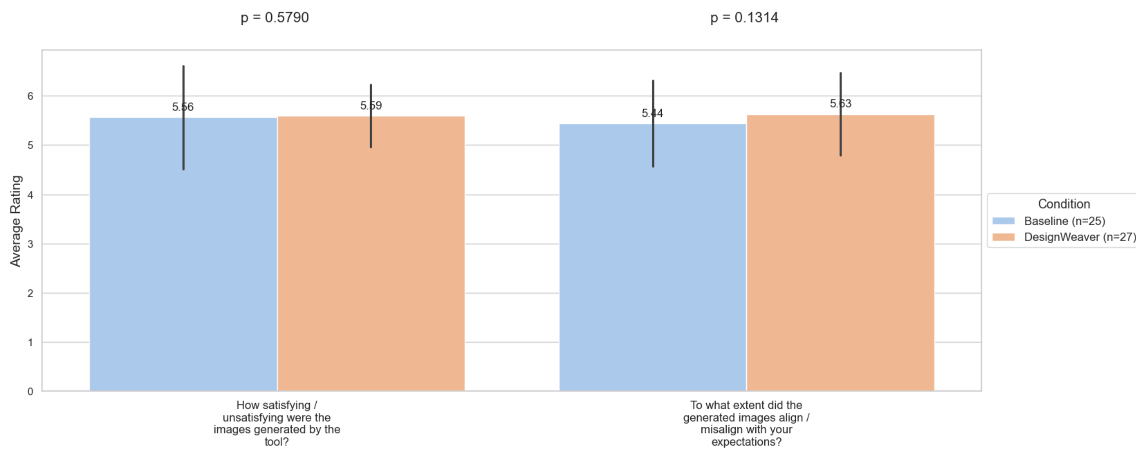

Fig. 16. DesignWeaver and baseline have similar levels of satisfaction reported on result and alignment.

During interviews, participants also reported challenges while using *DesignWeaver*. A notable issue was that choosing too many tags (over-tagging) led to confusion over the design process. Several participants (5 out of 27) indicated that over-tagging led to confusion and a sense of losing control over the design process. P11 mentioned, "*if you choose too many options, it'll get a bit too constricted,*" reflecting a perception that too many tags can make the system overly restrictive. P1 similarly felt that "*at the end, it was kind of already having too many tags, and it was kind of just showing me like a little off to what I needed.*" Participants also expressed concerns about the tool's responsiveness to changes in tags. For example, P17 stated, "*I feel like when I change the culture design tab, not much changed,*" suggesting that adjustments to tags did not always produce the expected variations.

Furthermore, participants reported difficulties when trying to generate designs with specific or detailed characteristics. P47 shared their experience: "*There were a lot of times where I was just clicking around trying to get more variety in an image… but it didn't really do that. Towards the end, I started getting closer, but it wasn't what the client wanted. The client wanted a curved back [on the chairs], but instead of giving me curved backs like it did in the beginning, it started* *giving me just straight backs.*" Similarly, P18 found it challenging to generate specific colors, saying, "*I wanted the colors to not be just tan and white, but they were only giving me tan and white. I didn't know how exactly specific I could be, or if I could just click certain colors.*" P5 also described the increased effort needed as designs became more complex, stating, "*as we got further on into the more nitty, gritty aspects, then it was more on me,*" reflecting the perceived effort needed to refine design as complexity increased.

With perceived challenges and support from *DesignWeaver*, participants using the tool reported similar satisfaction levels with the design images compared to the baseline group, both in terms of quality and alignment with their expectations, as shown in Figure 16.

7 DISCUSSION AND LIMITATION

Our study reveals several important insights into the effects of dimensional scaffolding in *DesignWeaver* on prompt creation, design diversity, and user experience. These findings have practical implications for supporting novice designers and highlight both the potential and limitations of current generative AI tools. Below, we discuss the key contributions of *DesignWeaver* and outline areas for improvement and future research.

7.1 Empowering Novice Designers with Adaptable Domain Knowledge

Dimensional scaffolding in *DesignWeaver* not only increased prompt length but also enriched the design vocabulary of novice users, empowering them to explore a wider range of design ideas. Participants using *DesignWeaver* generated significantly more unique design terms, blending both predefined and self-created tags, which expanded their understanding of design concepts. This flexibility allowed users to move beyond basic prompts and engage with more complex design dimensions.

The tool's structure encouraged dynamic interaction with tags, enabling participants to adapt and refine their designs in real time. For example, P24 began with familiar terms like "Minimalist" and "Eco-friendly" but gradually customized their prompts by adding specific details such as "Ergonomic" and "Stain-resistant." This process of iterative refinement fostered deeper engagement and creativity, as users explored new design dimensions and vocabulary with confidence.

In contrast, participants in the Baseline group tended to stick to initial terms, showing less exploration. Without the guidance provided by dimensional scaffolding, they were hesitant to add new terms, fearing it might confuse the generator. As a result, their design exploration was more limited, leading to narrower and less innovative outcomes.

Overall, *DesignWeaver* empowered novice designers by offering adaptable domain knowledge through dimensional scaffolding, encouraging richer, more versatile design outcomes and helping users confidently navigate complex design challenges.

7.2 User Engagement, Frustration, and Cognitive Overload

Dimensional scaffolding empowered users to experiment more confidently, encouraging them to explore design dimensions and vocabulary they might not have otherwise considered. Participants reported feeling more engaged and in control of the design process when using *DesignWeaver*, as the scaffolding prompted them to craft more detailed and nuanced prompts.

However, this increased confidence also raised user expectations for the quality of AI-generated outputs. As prompts became more complex, participants grew frustrated when the generative models struggled to fulfill the sophisticated design requests. This highlights a key limitation of current text-to-image (T2I) models, which lack the fine-grained control and spatial reasoning needed to handle such nuanced prompts.

Additionally, the phenomenon of over-tagging emerged as a challenge. Participants who added too many tags found that the tool became overly restrictive, limiting their ability to achieve the desired outcomes. This reflects a common issue with scaffolding systems: while they provide helpful structure, they can also introduce complexity and cognitive overload when not managed properly.

These limitations could be mitigated by the advancements in future generative models. Some current work such as ControlNet [69], CLIPasso[62]) offered greater precision and intuitive control. Moreover, introducing dynamic feedback mechanisms that alert users when their prompts become overly complex or recommending ways to reduce dimensionality could prevent users from feeling constrained by the system's output, improving both user experience and creative freedom. As these models evolve, future iterations of *DesignWeaver* could offer even greater creative freedom while aligning more closely with user expectations.

7.3 Generalization and Future Research Directions

While our study demonstrated the benefits of dimensional scaffolding for novice designers working on chair designs, it remains unclear how well this approach would generalize to expert designers or other design domains. The study was limited to relatively simple design tasks, and the results may differ when applied to more complex projects or experienced professionals.

Future research should explore the applicability of dimensional scaffolding across different design fields and with a more diverse set of users. This would help determine whether the tool can support a broader range of creative professionals and design challenges.

Several exciting directions for future research emerge from this study. First, integrating more advanced generative models with real-time feedback and dynamic scaffolding could offer a more seamless and responsive experience for both novice and expert users. Additionally, investigating how *DesignWeaver* can be adapted for more complex design tasks or collaborative design environments could provide deeper insights into its potential for supporting diverse creative workflows.

8 CONCLUSION

In this work, we introduced *DesignWeaver*, a generative AI system that empowers novice designers through dimensional scaffolding, encouraging broader design exploration and more nuanced prompt creation. Our findings demonstrate that *DesignWeaver* led to significantly longer, more diverse prompts and enabled users to discover and adopt new design dimensions. Participants engaged in richer creative processes, integrating both predefined and custom tags to refine their designs iteratively. Compared to a baseline, *DesignWeaver* users produced more unique design terms and exhibited greater confidence in navigating complex design challenges. This suggests that dimensional scaffolding can effectively bridge the gap between novice and expert design strategies, offering a powerful tool for enhancing creative workflows.

References

[1] Saeema Ahmed, Ken M Wallace, and Lucienne TM Blessing. 2003. Understanding the differences between how novice and experienced designers approach design tasks. *Research in Engineering Design* 14, 1 (2003), 1–11.

[2] Cynthia J Atman, Justin R Chimka, Karen M Bursic, and Heather L Nachtmann. 1999. A comparison of freshman and senior engineering design processes. *Design Studies* 20, 2 (1999), 131–152.

[3] Yuntao Bai, Saurav Kadavath, Sandipan Kundu, Amanda Askell, Jackson Kernion, Andy Jones, Anna Chen, Anna Goldie, Azalia Mirhoseini, Cameron McKinnon, et al. 2022. Constitutional ai: Harmlessness from ai feedback. *arXiv preprint arXiv:2212.08073* (2022).

[4] Omer Bar-Tal, Hila Chefer, Omer Tov, Charles Herrmann, Roni Paiss, Shiran Zada, Ariel Ephrat, Junhwa Hur, Yuanzhen Li, Tomer Michaeli, et al. 2024. Lumiere: A space-time diffusion model for video generation. *arXiv preprint arXiv:2401.12945* (2024).

[5] 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.

[6] 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* (San Francisco, CA, USA) *(UIST '23)*. Association for Computing Machinery, New York, NY, USA, Article 96, 14 pages. https://doi.org/10.1145/3586183.3606725

[7] Bill Buxton. 2007. *Sketching User Experiences: Getting the Design Right and the Right Design.* Morgan Kaufmann.

[8] Stuart K Card, Jock Mackinlay, and Ben Shneiderman. 1999. *Readings in information visualization: using vision to think.* Morgan Kaufmann.

[9] Huiwen Chang, Han Zhang, Jarred Barber, Aaron Maschinot, Jose Lezama, Lu Jiang, Ming-Hsuan Yang, Kevin Patrick Murphy, William T. Freeman, Michael Rubinstein, Yuanzhen Li, and Dilip Krishnan. 2023. Muse: Text-To-Image Generation via Masked Generative Transformers. In *Proceedings of the 40th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 202)*, Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett (Eds.). PMLR, 4055–4075. https://proceedings.mlr.press/v202/chang23b.html

[10] Leah Chong, I Lo, Jude Rayan, Steven Dow, Faez Ahmed, Ioanna Lykourentzou, et al. 2024. Prompting for products: Investigating design space exploration strategies for text-to-image generative models. *arXiv preprint arXiv:2408.03946* (2024).

[11] John Joon Young Chung and Eytan Adar. 2023. Promptpaint: Steering text-to-image generation through paint medium-like interactions. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. 1–17.

[12] N. Davis, C. Hsiao, K. Y. Singh, B. Lin, and B. Magerko. 2017. Quantifying Collaboration with a Co-Creative Drawing Agent. *ACM Trans. Interact. Intell. Syst.* 7, 4, Article 19 (dec 2017), 25 pages. https://doi.org/10.1145/3009981

[13] Prafulla Dhariwal and Alex Nichol. 2024. Diffusion models beat GANs on image synthesis. In *Proceedings of the 35th International Conference on Neural Information Processing Systems (NIPS '21)*. Curran Associates Inc., Red Hook, NY, USA, Article 672, 15 pages.

[14] Giulia Di Fede, Davide Rocchesso, Steven P. Dow, and Salvatore Andolina. 2022. The Idea Machine: LLM-based Expansion, Rewriting, Combination, and Suggestion of Ideas. In *Proceedings of the 14th Conference on Creativity and Cognition* (Venice, Italy) *(C&C '22)*. Association for Computing Machinery, New York, NY, USA, 623–627. https://doi.org/10.1145/3527927.3535197

[15] Kees Dorst and Nigel Cross. 2001. Creativity in the design process: Co-evolution of problem–solution. *Design Studies* 22, 5 (2001), 425–437.

[16] Steven P. Dow, Alana Glassco, Jonathan Kass, Melissa Schwarz, Daniel L. Schwartz, and Scott R. Klemmer. 2011. Parallel prototyping leads to better design results, more divergence, and increased self-efficacy. *ACM Trans. Comput.-Hum. Interact.* 17, 4, Article 18 (dec 2011), 24 pages. https://doi.org/10.1145/1879831.1879836

[17] Charlotte Fiell and Peter Fiell. 2005. *1000 Chairs* (25th anniversary edition ed.). Taschen, Köln, Germany.

[18] Jonas Frich, Midas Nouwens, Kim Halskov, and Peter Dalsgaard. 2021. How digital tools impact convergent and divergent thinking in design ideation. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–11.

[19] Jun Gao, Tianchang Shen, Zian Wang, Wenzheng Chen, Kangxue Yin, Daiqing Li, Or Litany, Zan Gojcic, and Sanja Fidler. 2022. Get3d: A generative model of high quality 3d textured shapes learned from images. *Advances In Neural Information Processing Systems* 35 (2022), 31841–31854.

[20] John S Gero. 1990. Design prototypes: A knowledge representation schema for design. *AI Magazine* 11, 4 (1990), 26–36.

[21] Vinod Goel and Peter Pirolli. 1992. The structure of design problem spaces. *Cognitive Science* 16, 3 (1992), 395–429.

[22] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2020. Generative adversarial networks. *Commun. ACM* 63, 11 (oct 2020), 139–144. https://doi.org/10.1145/3422622

[23] Jonathan Ho, William Chan, Chitwan Saharia, Jay Whang, Ruiqi Gao, Alexey Gritsenko, Diederik P Kingma, Ben Poole, Mohammad Norouzi, David J Fleet, et al. 2022. Imagen video: High definition video generation with diffusion models. *arXiv preprint arXiv:2210.02303* (2022).

[24] Satoshi Iizuka, Edgar Simo-Serra, and Hiroshi Ishikawa. 2017. Globally and locally consistent image completion. *ACM Trans. Graph.* 36, 4, Article 107 (jul 2017), 14 pages. https://doi.org/10.1145/3072959.3073659

[25] Youngseung Jeon, Seungwan Jin, Patrick C. Shih, and Kyungsik Han. 2021. FashionQ: An AI-Driven Creativity Support Tool for Facilitating Ideation in Fashion Design. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 576, 18 pages. https://doi.org/10.1145/3411764.3445093

[26] Pegah Karimi, Jeba Rezwana, Safat Siddiqui, Mary Lou Maher, and Nasrin Dehbozorgi. 2020. Creative sketching partner: an analysis of human-AI co-creativity. In *Proceedings of the 25th International Conference on Intelligent User Interfaces* (Cagliari, Italy) *(IUI '20)*. Association for Computing Machinery, New York, NY, USA, 221–230. https://doi.org/10.1145/3377325.3377522

[27] Dan Kondratyuk, Lijun Yu, Xiuye Gu, José Lezama, Jonathan Huang, Rachel Hornung, Hartwig Adam, Hassan Akbari, Yair Alon, Vighnesh Birodkar, et al. 2023. Videopoet: A large language model for zero-shot video generation. *arXiv preprint arXiv:2312.14125* (2023).

[28] Bryan Lawson. 2006. *How Designers Think: The Design Process Demystified.* Routledge.

[29] Chen-Hsuan Lin, Jun Gao, Luming Tang, Towaki Takikawa, Xiaohui Zeng, Xun Huang, Karsten Kreis, Sanja Fidler, Ming-Yu Liu, and Tsung-Yi Lin. 2023. Magic3d: High-resolution text-to-3d content creation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 300–309.

[30] Minghua Liu, Ruoxi Shi, Linghao Chen, Zhuoyang Zhang, Chao Xu, Xinyue Wei, Hansheng Chen, Chong Zeng, Jiayuan Gu, and Hao Su. 2024. One-2-3-45++: Fast Single Image to 3D Objects with Consistent Multi-View Generation and 3D Diffusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 10072–10083.

[31] Minghua Liu, Chao Xu, Haian Jin, Linghao Chen, Mukund Varma T, Zexiang Xu, and Hao Su. 2024. One-2-3-45: Any single image to 3d mesh in 45 seconds without per-shape optimization. *Advances in Neural Information Processing Systems* 36 (2024).

[32] Michael Xieyang Liu, Advait Sarkar, Carina Negreanu, Benjamin Zorn, Jack Williams, Neil Toronto, and Andrew D. Gordon. 2023. "What It Wants Me To Say": Bridging the Abstraction Gap Between End-User Programmers and Code-Generating Large Language Models. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23, Vol. 4)*. ACM, 1–31. https://doi.org/10.1145/3544548.3580817

[33] Vivian Liu, Rubaiat Habib Kazi, Li-Yi Wei, Matthew Fisher, Timothy Langlois, Seth Walker, and Lydia Chilton. 2024. LogoMotion: Visually Grounded Code Generation for Content-Aware Animation. *arXiv preprint arXiv:2405.07065* (2024).

[34] 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).

[35] Vivian Liu, Han Qiao, and Lydia Chilton. 2022. Opal: Multimodal Image Generation for News Illustration. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (Bend, OR, USA) *(UIST '22)*. Association for Computing Machinery, New York, NY, USA, Article 73, 17 pages. https://doi.org/10.1145/3526113.3545621

[36] Vivian Liu, Jo Vermeulen, George Fitzmaurice, and Justin Matejka. 2023. 3DALL-E: Integrating Text-to-Image AI in 3D Design Workflows. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (Pittsburgh, PA, USA) *(DIS '23)*. Association for Computing Machinery, New York, NY, USA, 1955–1977. https://doi.org/10.1145/3563657.3596098

[37] Mary Lou Maher and Josiah Poon. 1996. Modeling design exploration as co-evolution. *Computer-Aided Civil and Infrastructure Engineering* 11, 3 (1996), 195–209.

[38] James G. March. 1991. Exploration and Exploitation in Organizational Learning. *Organization Science* 2, 1 (1991), 71–87. https://doi.org/10.1287/orsc.2.1.71

[39] Tricia J Ngoon, Caren M Walker, and Scott Klemmer. 2019. The dark side of satisficing: Setting the temperature of creative thinking. In *Proceedings of the 2019 Conference on Creativity and Cognition*. 591–596.

[40] Alex Nichol, Heewoo Jun, Prafulla Dhariwal, Pamela Mishkin, and Mark Chen. 2022. Point-e: A system for generating 3d point clouds from complex prompts. *arXiv preprint arXiv:2212.08751* (2022).

[41] Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul Christiano, Jan Leike, and Ryan Lowe. 2024. Training language models to follow instructions with human feedback. In *Proceedings of the 36th International Conference on Neural Information Processing Systems* (New Orleans, LA, USA) *(NIPS '22)*. Curran Associates Inc., Red Hook, NY, USA, Article 2011, 15 pages.

[42] Srishti Palani, Zijian Ding, Austin Nguyen, Andrew Chuang, Stephen MacNeil, and Steven P. Dow. 2021. CoNotate: Suggesting Queries Based on Notes Promotes Knowledge Discovery. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 726, 14 pages. https://doi.org/10.1145/3411764.3445618

[43] Srishti Palani and Gonzalo Ramos. 2024. Evolving Roles and Workflows of Creative Practitioners in the Age of Generative AI. In *Proceedings of the 16th Conference on Creativity & Cognition*. 170–184.

[44] Srishti Palani, Yingyi Zhou, Sheldon Zhu, and Steven P. Dow. 2022. InterWeave: Presenting Search Suggestions in Context Scaffolds Information Search and Synthesis. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (Bend, OR, USA) *(UIST '22)*. Association for Computing Machinery, New York, NY, USA, Article 93, 16 pages. https://doi.org/10.1145/3526113.3545696

[45] Xingang Pan, Ayush Tewari, Thomas Leimkühler, Lingjie Liu, Abhimitra Meka, and Christian Theobalt. 2023. Drag Your GAN: Interactive Point-based Manipulation on the Generative Image Manifold. In *ACM SIGGRAPH 2023 Conference Proceedings* (Los Angeles, CA, USA) *(SIGGRAPH '23)*. Association for Computing Machinery, New York, NY, USA, Article 78, 11 pages. https://doi.org/10.1145/3588432.3591500

[46] Xiaohan Peng, Janin Koch, and Wendy E. Mackay. 2024. DesignPrompt: Using Multimodal Interaction for Design Exploration with Generative AI. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference* (Copenhagen, Denmark) *(DIS '24)*. Association for Computing Machinery, New York, NY, USA, 804–818. https://doi.org/10.1145/3643834.3661588

[47] Ben Poole, Ajay Jain, Jonathan T Barron, and Ben Mildenhall. 2022. Dreamfusion: Text-to-3d using 2d diffusion. *arXiv preprint arXiv:2209.14988* (2022).

[48] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. 2021. Learning Transferable Visual Models From Natural Language Supervision. In *Proceedings of the 38th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 139)*, Marina Meila and Tong Zhang (Eds.). PMLR, 8748–8763. https://proceedings.mlr.press/v139/radford21a.html

[49] Aditya Ramesh, Prafulla Dhariwal, Alex Nichol, Casey Chu, and Mark Chen. 2022. Hierarchical text-conditional image generation with clip latents. *arXiv preprint arXiv:2204.06125* 1, 2 (2022), 3.

[50] Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever. 2021. Zero-Shot Text-to-Image Generation. In *Proceedings of the 38th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 139)*, Marina Meila and Tong Zhang (Eds.). PMLR, 8821–8831. https://proceedings.mlr.press/v139/ramesh21a.html

[51] 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 (CVPR)*. 10684–10695.

[52] Chitwan Saharia, William Chan, Saurabh Saxena, Lala Lit, Jay Whang, Emily Denton, Seyed Kamyar Seyed Ghasemipour, Burcu Karagol Ayan, S. Sara Mahdavi, Raphael Gontijo-Lopes, Tim Salimans, Jonathan Ho, David J Fleet, and Mohammad Norouzi. 2024. Photorealistic text-to-image

diffusion models with deep language understanding. In *Proceedings of the 36th International Conference on Neural Information Processing Systems* (New Orleans, LA, USA) *(NIPS '22)*. Curran Associates Inc., Red Hook, NY, USA, Article 2643, 16 pages.

[53] Ruoxi Shi, Hansheng Chen, Zhuoyang Zhang, Minghua Liu, Chao Xu, Xinyue Wei, Linghao Chen, Chong Zeng, and Hao Su. 2023. Zero123++: a single image to consistent multi-view diffusion base model. *arXiv preprint arXiv:2310.15110* (2023).

[54] Uriel Singer, Adam Polyak, Thomas Hayes, Xi Yin, Jie An, Songyang Zhang, Qiyuan Hu, Harry Yang, Oron Ashual, Oran Gafni, et al. 2022. Make-a-video: Text-to-video generation without text-video data. *arXiv preprint arXiv:2209.14792* (2022).

[55] Erik Stolterman and Mikael Wiberg. 2010. Concept-driven interaction design research. *Human–Computer Interaction* 25, 2 (2010), 95–118.

[56] Sangho Suh, Meng Chen, Bryan Min, Toby Jia-Jun Li, and Haijun Xia. 2024. Luminate: Structured Generation and Exploration of Design Space with Large Language Models for Human-AI Co-Creation. 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 644, 26 pages. https://doi.org/10.1145/3613904.3642400

[57] Maryam Tohidi, William Buxton, Ronald Baecker, and Abigail Sellen. 2006. Getting the right design and the design right. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada) *(CHI '06)*. Association for Computing Machinery, New York, NY, USA, 1243–1252. https://doi.org/10.1145/1124772.1124960

[58] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. 2023. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971* (2023).

[59] Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, et al. 2023. Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288* (2023).

[60] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Proceedings of the 31st International Conference on Neural Information Processing Systems* (Long Beach, California, USA) *(NIPS'17)*. Curran Associates Inc., Red Hook, NY, USA, 6000–6010.

[61] Ruben Villegas, Mohammad Babaeizadeh, Pieter-Jan Kindermans, Hernan Moraldo, Han Zhang, Mohammad Taghi Saffar, Santiago Castro, Julius Kunze, and Dumitru Erhan. 2022. Phenaki: Variable length video generation from open domain textual descriptions. In *International Conference on Learning Representations*.

[62] Yael Vinker, Ehsan Pajouheshgar, Jessica Y. Bo, Roman Christian Bachmann, Amit Haim Bermano, Daniel Cohen-Or, Amir Zamir, and Ariel Shamir. 2022. CLIPasso: semantically-aware object sketching. *ACM Trans. Graph.* 41, 4, Article 86 (jul 2022), 11 pages. https://doi.org/10.1145/3528223.3530068

[63] Patrick H. Winston. 1970. *Learning Structural Descriptions From Examples.* Technical Report. Massachusetts Institute of Technology, USA.

[64] Tongshuang Wu, Ellen Jiang, Aaron Donsbach, Jeff Gray, Alejandra Molina, Michael Terry, and Carrie J Cai. 2022. PromptChainer: Chaining Large Language Model Prompts through Visual Programming. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 359, 10 pages. https://doi.org/10.1145/3491101.3519729

[65] Tongshuang Wu, Michael Terry, and Carrie Jun Cai. 2022. AI Chains: Transparent and Controllable Human-AI Interaction by Chaining Large Language Model Prompts. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 385, 22 pages. https://doi.org/10.1145/3491102.3517582

[66] Xiaotong (Tone) Xu, Jiayu Yin, Catherine Gu, Jenny Mar, Sydney Zhang, Jane L. E, and Steven P. Dow. 2024. Jamplate: Exploring LLM-Enhanced Templates for Idea Reflection. In *Proceedings of the 29th International Conference on Intelligent User Interfaces* (Greenville, SC, USA) *(IUI '24)*. Association for Computing Machinery, New York, NY, USA, 907–921. https://doi.org/10.1145/3640543.3645196

[67] S. Yilmaz, S.R. Daly, C.M. Seifert, and R. Gonzalez. 2015. Design heuristics in innovative products. *Journal of Mechanical Design* 137, 7 (2015), 071102.

[68] J.D. 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* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 437, 21 pages. https://doi.org/10.1145/3544548.3581388

[69] Lvmin Zhang, Anyi Rao, and Maneesh Agrawala. 2023. Adding Conditional Control to Text-to-Image Diffusion Models. *Proc. of IEEE/CVF International Conf. on Computer Vision (ICCV)* (October 2023), 3836–3847.

[70] Xiaoyu Zhou, Xingjian Ran, Yajiao Xiong, Jinlin He, Zhiwei Lin, Yongtao Wang, Deqing Sun, and Ming-Hsuan Yang. 2024. Gala3d: Towards text-to-3d complex scene generation via layout-guided generative gaussian splatting. *arXiv preprint arXiv:2402.07207* (2024).

A PROMPTS USED IN THE LLM PIPELINES

A.1 Document Digestion

"Get 3 most important design dimensions from the requirement doc. For each dimension, generate 3-5 tags. Be concise."

A.2 Image Generation

```
export async function imageGenerate(
```

```
  prompt_text,
  num = 3,
  size = "1024x1024",
  quality = "standard",
  callback = null
) {
  let promises = [];

  for (let i = 0; i < num; i++) {
    promises.push(
      openai.images
        .generate({
          model: "dall-e-3",
          prompt: prompt_text,
          n: 1,
          size: String(size), // Ensure size is a string
          quality: quality,
        })
        .then(async (response) => {
          if (callback) {
            if (callback.constructor.name === "AsyncFunction") {
              await callback(response.data[0].url, i);
            } else {
              callback(response.data[0].url, i);
            }
          }
          return response.data[0].url;
        })
    );
  }

  const responses = await Promise.all(promises);

  return responses;
}
```

A.3 Prompt Generation

```
export async function convertTagsToPrompt(oldTags, newTags, promptText) {
  const systemMessage = {
    role: "system",
    content:
```

```
  "You are a design generalist that converts design tags and weights into descriptive
       prompts. Your task is to update the prompt according to the given old and new
      tags comparison and their corresponding weights, making sure to remove any
      references to tags that have been removed or neutralized (weight = 0). Preserve
       as much of the original prompt as possible, but reflect all tag changes
      accurately.",
};

function transformTags(tags) {
  return tags.map((tag) => ({
    name: tag.name,
    tags: tag.options.map((option) => ({
      tag: option.optionName,
      weight: option.scale,
    })),
  }));
}

const transformedNewTags = transformTags(newTags);
const transformedOldTags = transformTags(oldTags);

const userMessage = {
  role: "user",
  content: `Update the old prompt by comparing the old and new tags and weights pairs.
      Any tags with a weight of 0 should be removed from the prompt. Any tags with a
      weight of 1 should be included in the prompt.
            New Tags: ${JSON.stringify(transformedNewTags)}
            Old Tags: ${JSON.stringify(transformedOldTags)}
            Old Prompt: "${promptText}"
            Just return the prompt itself. Use complete sentences to describe the
                design.`,
};

const response = await openai.chat.completions.create({
  model: "gpt-4o",
  messages: [systemMessage, userMessage],
  max_tokens: 2000,
});

let updatedPrompt = response.choices[0].message.content.trim();
```

```
  // Remove leading and trailing double quotes if they exist
  if (updatedPrompt.startsWith('"') && updatedPrompt.endsWith('"')) {
    updatedPrompt = updatedPrompt.substring(1, updatedPrompt.length - 1);
  }

  return updatedPrompt;
}
```

A.4 Tag Extraction

```
export async function convertImageToTags(url, cur_categories) {
  const JSON_prompt = cur_categories.map((category) => {
    return {
      name: category.name,
      tags: category.options.map((option) => {
        return option.optionName;
      }),
    };
  });
  console.log(JSON_prompt);
  const system_message = {
    role: "system",
    content:
      "You are a creative and helpful designer who assists in identifying and
          categorizing aesthetic dimensions of product designs. The response should be
          format like: {newtags:[{'name':'Dimention Name', 'tags':['tag1', 'tag2', 'tag3'
           ... }]}",
  };
  // ensure the response JSON follows the format of JSON_prompt

  const text_prompt =
    "What relevant aesthetic dimensions and design element tags are in this image?
        Reference the existing tags and think outside the box and include all relevant
        dimensions.\
     On top of the old tags, generate 1-5 new tags that either append to existing design
         dimensions or create new dimensions while avoiding from creating similar
         dimensions to the old ones. Provide the output in a JSON format." +
    JSON.stringify(JSON_prompt);
  const user_message = {
    role: "user",
    content: [
      {
```

```
          type: "text",
          text: text_prompt,
        },
        {
          type: "image_url",
          image_url: {
            url: url,
            detail: "low",
          },
        },
      ],
    };
    const response = await openai.chat.completions.create({
      model: "gpt-4o-mini",
      response_format: { type: "json_object" },
      messages: [system_message, user_message],
      max_tokens: 3000,
    });
    return JSON.parse(response.choices[0].message.content)["newtags"];
}
```

A.5 Tag Recommendation

```
export async function getRecommendedOptions(currentCategory) {
  const systemMessage = {
    role: "system",
    content:
      "You are a helpful assistant that provides concise 5 distinct design
          recommendations based on existing design tags for dining room chair design.",
  };

  const userMessage = {
    role: "user",
    content: `Based on the current design tags in the category "${currentCategory.name}",
        suggest 5 new distinct design options. Please provide a simple list of options
        separated by commas and nothing else. Don't add numbers or bullet points.`,
  };

  const response = await openai.chat.completions.create({
    model: "gpt-4o-mini",
    messages: [systemMessage, userMessage],
    max_tokens: 50,
```

```
  });

  const recommendedOptions = response.choices[0].message.content
    .trim()
    .split(",")
    .map((option) => option.trim().replace(/"/g, ""));
  return recommendedOptions;
}
```

A.6 Dimension Recommendation

```
export async function getRecommendedDimensions(currentCategories) {
  const systemMessage = {
    role: "system",
    content:
      "You are a helpful assistant that provides concise design recommendations based on
          existing design dimensions for dining room chair design.",
  };

  const categoriesList = currentCategories
    .map((category) => category.name)
    .join(", ");

  const userMessage = {
    role: "user",
    content: `Based on the current design dimensions: [${categoriesList}], suggest 5 new
        distinct dimensions. Please provide a simple list of dimensions separated by
        commas and nothing else. Don't add numbers or bullet points.`,
  };

  const response = await openai.chat.completions.create({
    model: "gpt-4o-mini",
    messages: [systemMessage, userMessage],
    max_tokens: 50,
  });

  const recommendedDimensions = response.choices[0].message.content
    .trim()
    .split(",")
    .map((dimension) => dimension.trim().replace(/"/g, ""));
  return recommendedDimensions;
}
```

B USER STUDY RELATED MATERIAL

B.1 Design Document

B.1.1 Design requirement. (Email from our client David)

Dear Designer,

I hope this message finds you in great spirits. As an architect deeply passionate about modern and sustainable design, I'm reaching out with a project that's very close to my heart. I'm in the process of bringing a vision to life—a set of 10 dining room chairs that's not just a piece of furniture but a statement of my lifestyle and values.

Here's What I Envision: I'm drawn to contemporary style, embracing minimalism with open arms. My home is a testament to my love for clean lines and functional aesthetics, and I seek to extend this philosophy to this new chair design. I imagine it in neutral tones, allowing for the occasional bold accent through customization options. The materials, of course, need to echo my commitment to sustainability—a combination of natural wood and high-quality, eco-friendly fabrics would be ideal.

Design Specifications:

- **Dimensions**: A height of 36 inches, with a seat height of 18 inches, seems perfect. The width and depth should be about 18 and 16 inches, respectively, providing ample space without compromising the minimalist design.
- **Comfort and Ergonomics**: Given my love for hosting dinner parties, the chair must offer comfort for long conversations. A curved backrest for lumbar support and a cushioned seat are essential.
- **Materials**: Solid oak or a similar hardwood for the frame would complement my home's aesthetic, paired with durable and stain-resistant fabric that's sourced responsibly.

Functionality Needs: The chair should be lightweight, making it easy to move around, yet sturdy enough to withstand the joyous chaos of my gatherings. I also need to ensure the feet are kind to my flooring—no scratches are welcome.

Sustainability and Quality: I value furniture that reflects my dedication to sustainable living and modern design. It should be durable and sourced from eco-friendly suppliers, aligning with my values and lifestyle.

Budget and Timeline: I've allocated $1000 to $1500 per unit for this project, aiming for a balance between impeccable quality and affordability. Ideally, the design would be finalized within 30 days from now, with a prototype ready for review 60 days later and production commencing 90 days after prototype approval.

Additional Considerations: Ease of assembly and eco-friendly packaging that ensures safe transport without compromising our planet's health is crucial.

I'm thrilled at the prospect of working together to create a piece that serves its purpose and does so with style and conscience. Your talent in design and understanding of functional aesthetics make you the perfect partner for this venture.

Looking forward to your thoughts and ideas.

Best regards, David Thompson

B.1.2 Client's Persona.

- David Thompson
- Age: 35
- Occupation: Architect
- Location: San Francisco, California

- Marital Status: Married
- Interests:
  - Passionate about modern and sustainable architecture.
  - Enjoys reading about interior design and home renovation.
  - Loves outdoor activities like hiking and cycling.
- Lifestyle:
  - Lives in a well-designed, modern home.
  - Prefers a minimalist and functional aesthetic.
  - Often hosts dinner parties and small gatherings.
  - Has a few dogs of varying ages
- Purchase Motivations:
  - Seeking furniture that reflects his taste for modern design.
  - Values sustainability and eco-friendly materials.
  - Wants durable, high-quality furniture that can withstand regular use.
- Functional Needs:
  - Comfortable for long dinner conversations, easy to move, and matches with a diverse range of dining tables.
  - Able to stand pets' daily activities
- Aesthetic Preferences:
  - Prefers neutral tones with occasional bold accents.
  - Likes clean lines and uncluttered spaces.
  - Appreciates furniture that makes a statement but remains timeless.

B.2 Sketch

Here is the sketch (Figure 17).

B.3 Mood Board

Here are the mood boards (Figure 17).

B.3.1 Post-Experiment Survey Questions.

- What criteria did you use to select the final image for the client?
- Why do you think this image best fits your client's needs?
- List all design dimensions and options you learned about the chair design space.
- How satisfying or unsatisfying were the images generated by the tool?
- Why was the image result satisfying or unsatisfying to you?
- Did the generated images align with your expectations? Why or why not?
- How difficult or easy was it to convert ideas into a prompt?
- Why was it easy or difficult to generate prompts?
- Describe your process for crafting prompts for image generation.
- To what extent did the tool help you visualize your ideas?
- Did the tool help you explore the different dimensions of the design space?

Fig. 17. Design Document Sketches

- Did the tool help you come up with the initial prompt?
- Did the tool help you refine your design concept?
- Did your design improve throughout the iterative process?
- Did your text prompts become more detailed and nuanced over time?
- What aspects of the tool were most intuitive or useful, and which were confusing or difficult?
- How could the Design Assistant tool be improved?

Fig. 18. Design Document Mood Board 1

Fig. 19. Design Document Mood Board 2

Fig. 20. Design Document Mood Board 3