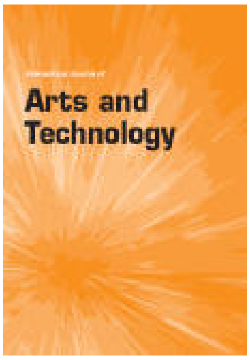


International Journal of Arts and Technology

ISSN online: 1754-8861 - ISSN print: 1754-8853
https://www.inderscience.com/ijart

An analysis of text-to-image generative models as creativity support tools

Emel Cantürk Akyildiz

DOI: 10.1504/IJART.2025.10072474

Article History:

Received:	06 March 2025
Last revised:	05 April 2025
Accepted:	06 April 2025
Published online:	04 August 2025

An analysis of text-to-image generative models as creativity support tools

Emel Cantürk Akyildiz

Department of Architecture,
Mimar Sinan Fine Arts University,
Istanbul, Turkey
Email: emel.canturk@msgsu.edu.tr

Abstract: This study examines the role of text-to-image (T2I) generative models as creativity support tools in architectural design education. It investigates how 28 first-year architecture students engage with prompt engineering to generate architectural representations using a T2I model. The methodology involved a structured task where students crafted and iterated prompts to recreate architectural forms, followed by a Creativity Support Index (CSI) survey and qualitative analysis of generated prompts and images. Findings indicate that structured prompts with precise modifiers significantly enhance output quality, while the trial-and-error nature of prompt crafting presents challenges. Students struggled to articulate complex visual concepts due to limited domain-specific vocabulary and lacked control over composition. CSI results showed strong support for exploration and expressiveness, though immersion and collaboration were limited. The study underscores the potential of T2I models in architectural pedagogy while highlighting the need for structured training in prompt engineering and improved AI-human interaction strategies.

Keywords: generative AI; GAI; text-to-image models; prompt engineering; creativity support tools; human-AI interaction; architectural design education.

Reference to this paper should be made as follows: Cantürk Akyildiz, E. (2025) 'An analysis of text-to-image generative models as creativity support tools', *Int. J. Arts and Technology*, Vol. 15, No. 3, pp.257–282.

Biographical notes: Emel Cantürk Akyildiz is an Associate Professor of Architecture at Mimar Sinan Fine Arts University. She received her BArch degree in Architecture (2006), MSc degree in Architectural Design (2009) and PhD degree in Architectural Design (2017) at Istanbul Technical University (ITU). She teaches undergraduate and graduate courses, such as design studio, architectural representation, 20th-century housing architecture, and generative AI in design. Her research areas are mainly architectural education, architectural design, architectural theory and criticism. Currently, her research focuses on emerging generative AI technologies as a new design method and their transformative impact on design thinking and architectural design education.

1 Introduction

Artificial intelligence has been a prominent topic in technology over the past decade. However, the emergence of generative AI (GAI) has propelled it into global attention, driving an unprecedented wave of AI advancements and adoption. Recent breakthroughs in GAI technologies at the intersection of artificial neural networks and linguistics led to the introduction of text-to-image (T2I) generative models. Text-guided generation of images with machine learning (ML) technology has made significant advances due to the deep learning architecture innovations (Van den Oord et al., 2017; Vaswani et al., 2017) and has seen an increasing interest since 2021 when OpenAI introduced Contrastive Language-Image Pre-training (CLIP), which efficiently learns visual concepts from natural language supervision and can generate images using Generative Pre-trained Transformer 3 (GPT-3) (Brown et al., 2020; Radford et al., 2021; Ramesh et al., 2021). Today, there is a growing ecosystem of T2I tools operating through easily accessible web-based services that can synthesise photorealistic images from textual descriptions (known as *prompts*), such as Midjourney (URL-1), DALL-E (Ramesh et al., 2021), Stable Diffusion (Rombach et al., 2022), Imagen (Saharia et al., 2022), Muse (Chang et al., 2023), LookX AI (URL-2), Leonardo AI (URL-3), etc. which significantly surpass previous architectures in image synthesis capabilities (Dhariwal and Nicol, 2021; Chang et al., 2023; Oppenlaender et al., 2024).

GAI, particularly T2I models, transforms creative disciplines such as art, design, and architecture. These tools have revolutionised workflows, enabling rapid visualisations of conceptual ideas and accelerating iterative design processes (Ploennigs and Berger, 2023; Jaruga-Rozdolska, 2022). Despite these advancements, the integration of T2I models into creative industries also raises philosophical and ethical considerations, particularly concerning authorship, originality, and the evolving role of human creativity in AI-assisted workflows. Consequently, in recent years considerable research and debates have focused on the creativity of T2I generative models (see Oppenlaender, 2022, 2023; Oppenlaender et al., 2024; Paananen et al., 2023; Rezwana and Maher, 2022; Wingström et al., 2022; Karimi et al., 2020; Reynolds and McDonell, 2021; Deckers et al., 2023; Marrone et al., 2024).

Within the field of architectural design, T2I models have emerged as powerful tools for exploring new realms of creativity. Recent studies suggest that they hold significant promise in supporting architects during the early stages of ideation (Bolojan and Vermisso, 2020; Desouki et al., 2023; Albaghajati et al., 2023; Cheung and Dall'Asta, 2024; Paananen et al., 2023; Tan and Luhrs, 2024). Despite their potential, integrating T2I technologies into architectural education remains a nascent area of inquiry, with limited research exploring their pedagogical applications (see Iranmanesh and Lotfabad, 2024; Chandrasekera et al., 2024; Cheung and Dall'Asta, 2024; Stigsen et al., 2023; Paananen et al., 2023). This study aims to address this gap by examining the engagement of first-year architecture students with T2I generative models and evaluating their potential as creativity support tools.

This study explores the prompt engineering experience of 28 first-year architecture students using a T2I generative model for a creative task. In T2I systems, prompting – natural language text inputs – is the primary interaction method between the user and the model, alongside other techniques such as image-based inputs, text-accompanied prompts, and in painting – filling in missing parts of an image. Prompt engineering, the creative process of writing and refining prompts, is central to human-AI co-creation process (Oppenlaender, 2023; Deckers et al., 2023). However, it poses significant challenges due to its steep learning curve and reliance on trial and error (Oppenlaender, 2022; Sanchez, 2023). This research examines these challenges from a human-computer interaction (HCI) perspective, focusing on how first-year architecture students navigate prompt engineering in T2I models. The study addresses four key research questions:

RQ1 What strategies do students develop for writing prompts?

RQ2 What challenges do students face when writing prompts in T2I generative models?

RQ3 To what extent can T2I models act as creativity support tools for first-year architecture students?

RQ4 What are the implications of integrating T2I generative models into architectural design education?

Understanding the strategies and challenges associated with prompt engineering is crucial for effectively integrating T2I models into architectural design education. By exploring these research questions, this study aims to provide insights into how T2I models can serve as creativity support tools and the broader implications of their integration into architectural design education.

2 A framework for AI-driven creativity and co-creativity in design

Creativity is a complex and multifaceted phenomenon with no commonly agreed definition. Various researchers have approached creativity from diverse perspectives, defining it as a process (Stein, 1953; Welsch, 1980; Torrance, 1993), an act (Koestler, 1964; Boone and Hollingsworth, 1990), or an ability/capacity (Guilford, 1950; Mumford and Gustafson, 1988; Vernon, 1989) that results in a product or idea that is both novel (unexpected, unusual) and valuable (appropriate, adaptive, useful) (Stein, 1953; Rogers, 1954; Welsch, 1980; Boone and Hollingsworth, 1990; Sternberg and Lubart, 1999; Boden, 2001; Mumford, 2003).

Kampylis and Valtanen (2010) reviewed numerous definitions and found a common understanding of creativity as generating novel and valuable solutions by synthesising existing facts or ideas. Koestler (1964, p.120) remarked, "The creative act is not an act of creation in the sense of the Old Testament. It does not create something out of nothing: it uncovers, selects, reshuffles, combines, and synthesises already existing facts, ideas, faculties, and skills." Several researchers have linked creativity to the formation of novel associations between existing elements, "forming of associative elements into new combinations" [Mednick, (1962), p.221], "products by transformation of existing products" [Welsch, (1980), p.97], "forming connections between ideas based on shared predicates" [Sasso, (1980), p.131].

Building on earlier works, Torrance (1993) identifies 'thinking by analogy' as a foundational element of creative thinking. Drawing from Ribot's (1906) earlier work, Torrance (1993, p.232) characterises analogy as recognising similarities between seemingly disparate objects or concepts. Similarly, Bronowski (1985) frames this as 'unexpected likeliness', where novel connections arise from the interplay of unrelated elements. Torrance (1993) argues that analogical thinking is a powerful tool for generating novel ideas. However, he also draws attention to the 'non-rational' aspect of creative thinking and underlines the importance of critical evaluation, or what he termed 'discrimination' or 'choice' as a crucial act of the creative process. Creativity is not merely about forming connections but also involves critically evaluating those connections, discerning between valuable and nonsensical analogies, and selecting the most promising ones. This dual process – of generating and filtering – ensures that creative outputs are not only novel but also relevant and meaningful.

2.1 Concept of human-AI co-creativity

Several researchers (Rhodes, 1961; Torrance, 1993; Sasso, 1980; Boden, 2004) characterise creativity as a fundamental aspect of human intelligence, intricately connected to various elements such as cognition, perception, critical thinking, motivation, emotion, and environmental and social factors. Rhodes' (1961) 4Ps framework (categorising creativity into person, process, product, and press) introduces a broad scale of factors that influence creative expression, from personal traits such as intelligence, motivation and openness to environmental and social factors. Similarly, after examining numerous definitions of creativity, Kampylis and Valtanen (2010) observed that creativity tends to follow an individual and intentional process. Based on these definitions, creativity is considered a trait of the human mind, yet AI poses a challenge to the conventional understanding of creativity. Wingström et al. (2022) argue that the integration of AI into professional workflows introduces the concept of 'co-creativity', which describes a collaborative human-AI creative process. 'Co-creativity' is defined as a kind of 'blended' creativity where the creativity of humans and AI interact on a shared task [Karimi et al., (2020), p.22] This emerging paradigm suggests that the boundary between human and artificial creativity is becoming increasingly blurred, redefining the creative process in a technologically augmented context.

The rise of GAI models has sparked debate over whether they should be considered autonomous artists or designers or regarded as tools or mediums utilised by human creators (Chang et al., 2023). This discussion is closely tied to broader philosophical and cognitive questions about the nature of creativity, agency, and authorship. Hertzmann (2018) notes that media narratives often anthropomorphise algorithms, attributing human-like qualities such as intention and consciousness; some even describe algorithms as independent artists. Yet, the nature of consciousness remains unresolved, and it is unclear whether AI can simulate human-like subjective experience (Chalmers, 1996). Hertzmann (2018) argues that 'art is an interaction between social agents', emphasising the inherently social and contextual nature of artistic creation. The absence of environmental awareness, cultural context, and personal intent in GAI highlights its distinction from human creativity. While AI can generate aesthetically compelling works, it lacks the lived experiences, emotions, and cultural understanding that shape human creativity (Floridi and Sanders, 2004).

Within this framework, GAI models can best be understood as interactive agents rather than autonomous creators. Human creators play a critical role in interpreting, curating, and contextualising AI-generated outputs, ensuring that meaning, intent and cultural relevance are embedded within the final work. This aligns with the theories of creativity discussed above and the perspective that AI serves as a collaborator rather than a replacement for human creativity (Manovich, 2018; McCormack et al., 2019).

2.2 GAI as a creative partner in architectural design

Architectural design, relying on experience and creativity to create new designs (Pena et al., 2021), encompasses creative activities and tasks such as idea generation, conceptualisation, and design development [Rafizadeh et al., (2004), p.80]. During the conceptual stage of an architectural project, the design requirements are not yet clearly established, making this process more of an exploration of both the requirements themselves and the potential solutions to address them (Logan and Smithers, 1992; Gero, 1994; Maher, 2000). Maher et al. (1996) argue that before beginning the design synthesis, designers often lack a comprehensive understanding of the problem. During the conceptual design phase, they explore ideas to gain clarity about the problem instead of immediately seeking a solution; therefore, design is an iterative process involving the exploration of both the problem space and the solution space (Pena et al., 2021). In this context, GAI technologies, particularly T2I models, offer significant potential to assist architects during the design ideation and conceptual design phase by expanding and diversifying both the problem space and the solution space; and accelerating and automating the exploration process. Therefore, the emergence of GAI underlines a paradigm shift in design thinking, where the computer is not just an aid but more a 'creative design partner' in the context of human-AI co-creation, human and AI collaborating on a shared creative product.

Recent research has demonstrated that AI can generate novel ideas and provide inspiration by producing sketches with varying degrees of similarity to existing concepts (Kantosalo and Toivonen, 2016; Karimi et al., 2020; Maher, 2012; Wingström et al., 2022). Multiple studies have highlighted the efficacy of GAI in assisting architects during early design ideation (Bolojan and Vermisso, 2020; Desouki et al., 2023; Albaghajati et al., 2023; Cheung and Dall'Asta, 2024; Paananen et al., 2023; Tan and Luhrs, 2024). As del Campo and Carlson (2022, p.179) observe, "Synthetic images generated by AI can stimulate the human mind to push the boundaries of architectural creativity. Because they are based on existing information, they are familiar enough to be construed as architecture but strange enough to provoke us and challenge us as designers." This capability enables architects to explore unconventional ideas that might not emerge through traditional design methods.

As Wingström et al. (2022, p.188) highlight, "AI cannot work intentionally or consciously; humans are required to interpret, develop, and create meaning for the outcomes that AI produces." This emphasises the necessity of human intuition, critical thinking, and design sensibility in the AI-assisted design process. Human-AI co-creation fosters a dynamic relationship where designers leverage AI-generated outputs, interpreting and refining them to align with their artistic vision and functional requirements. In this context, the most significant shift is observed in the designer's role. No longer the sole creator, the designer transitions into a curator and evaluator of AI-generated possibilities. Instead of manually crafting each design iteration, architects now assess, refine, and contextualise AI-generated outputs.

2.3 Creative work of prompt engineering

Prompt engineering is the systematic process of crafting and refining textual inputs (prompts) to control and optimise the outputs of GAI models, particularly T2I systems. Robertson et al. (2024) emphasise its role in enhancing human-AI collaboration, positioning prompts as a key interface in knowledge creation. Researchers highlight that prompt engineering is not merely an intuitive practice but a learned skill developed through systematic experimentation and iteration (Oppenlaender, 2023; Chang et al., 2023).

A defining characteristic of prompt engineering is its iterative and experimental nature. Dang et al. (2022) highlight that writing effective prompts is primarily a trial-and-error process where users refine their inputs based on the model's responses. This iterative refinement is crucial in steering the model toward the desired output, requiring an understanding of how GAI interprets language and latent representations (Robertson et al., 2024). Oppenlaender (2022) further underscores that proficiency in prompt engineering is cultivated through practice, community engagement, and the application of structured methodologies.

The design of effective prompts has become a key research focus within HCI. Several studies approach prompt engineering as a creative process itself, identifying and addressing several factors, including lexical choice, syntactic structure, and the use of modifiers as key elements in the effectiveness of prompts. Chang et al. (2023) argue that precision in vocabulary selection significantly impacts AI-generated results, as nuanced linguistic variations can lead to distinct outputs. However, Liu and Chilton (2022) found that rephrasing a prompt while preserving key terms does not necessarily enhance image generation quality, suggesting that selecting robust subject and style keywords is more influential than minor syntactic adjustments.

Research suggests that successful prompt crafting relies on structured composition and the inclusion of targeted keywords. Robertson et al. (2024) stress that effective prompts require an awareness of AI model biases and training data constraints. Other researchers (Barros and Ai, 2024; Oppenlaender, 2022, 2023) highlight the role of modifiers in refining GAI outputs in predictable ways, making them integral to prompt optimisation. Various prompt taxonomies have been proposed to aid structured prompting. Xie et al. (2023) categorise prompt modifiers into subject, form, and content, while Sanchez (2023) expands this framework to include medium, artistic influence, detail, lighting, colour, and composition. Oppenlaender (2023) classifies modifiers into six types: subject term, style modifier, image prompt, quality booster, repeating term, and magic term. These taxonomies illustrate the importance of structured linguistic input in achieving high-quality AI-generated results (Oppenlaender, 2023).

Despite variations in taxonomic structures, there is consensus that prompts function as structured linguistic frameworks rather than simple text inputs. Researchers also emphasise that prompt effectiveness extends beyond text alone, with multimodal approaches – such as integrating image prompts – enhancing subject representation and compositional accuracy in AI-generated content (Qiao et al., 2022).

3 Materials and methods

3.1 Data collection

The data collection process was structured into four key stages:

1. Pre-task survey: At the outset, participants completed a pre-task survey to assess their general knowledge of AI, T2I generative models, and any prior experience with these models. This stage established a baseline understanding of participants' familiarity with the technology.
2. Task execution sessions: Participants were then introduced to the procedures and given detailed instructions for completing the assigned task within three hours. The

task involved crafting written prompts and generating images using a T2I tool. Throughout the session, participants' inputs (*text prompts*) and outputs (*generated images*) were systematically recorded for analysis.

3 Creativity support index (CSI) survey: After completing the task, participants responded to the CSI survey, a validated psychometric tool developed by Carroll and Latulipe (2009) to evaluate the level of creativity support provided by various tools, systems, or interfaces. This technique addresses the challenges of assessing creativity, as traditional qualitative methods (e.g., observation and interviews) often lack standardisation and make it challenging to compare tools or derive statistically significant results [Carroll and Latulipe, (2009), p.4009]. CSI provides a standardised framework for evaluating creativity support, offering a quantitative complement to the qualitative insights gained during the study process.

4 Post-task survey: Finally, participants responded to a post-task survey comprising open-ended questions. This survey explored their reflections on the experience, including challenges encountered while using the T2I generation tool, satisfaction with the generated outputs, and their perceived self-improvement during the task execution.

The data collected encompassed both quantitative measures (e.g., CSI scores, survey responses) and qualitative insights (e.g., participants' reflections and recorded study processes), allowing for a comprehensive analysis of the participants' experiences using T2I generative models in an architectural context.

3.2 Participants

The experiment was conducted with 28 first-year architecture students. According to the pre-task survey, 89.3% of participants reported no prior experience with T2I generative models, and 60.7% indicated they had never heard of any T2I tools. Only three students claimed to have some experience with these tools; however, two rated their competency as very low (1 on a Likert scale), while the third rated their proficiency as medium (3 on a Likert scale). In adherence to the Sex and Gender Equity in Research (SAGER) Guidelines (Van Epps et al., 2022), gender information was not collected, as it was deemed irrelevant to the study.

3.3 The task and the procedure

During the experiment, participants were assigned a representation task to be completed in three hours. The task required them to analyse and verbally interpret provided architectural images, then recreate these representations by writing and refining text prompts using Stable Diffusion, Model SDXL-1.0. For this, Stable Diffusion's web interface, DreamStudio, was chosen, as it is easily accessible, offering a user-friendly web user interface and free-of-charge usage for new users. The web-based interface, DreamStudio, was selected for its ease of access and intuitive design. DreamStudio's user-friendly interface eliminates the need for prior technical knowledge or software installation while also offering free-of-charge access for new users, making it an ideal choice for this experiment. As highlighted by researchers, offering user-friendly interfaces for immediate application during in-class exercises is critical in enhancing user acceptance (Cheung and Dall'Asta, 2024). The process and workflow of the 28 participants while executing the task were recorded in the form of written descriptions and visual outputs, providing detailed insights into their interactions with the tool.

Participants were first introduced to DreamStudio and provided with 'tips' for crafting and refining prompts. They were then asked to write prompts for each provided image in three sub-tasks and revise their prompts to test if they could get better results. Each participant was required to generate a minimum of five and a maximum of seven prompts, producing three alternative images per prompt while maintaining the same aspect ratio as the original image. They then evaluated their results to determine which prompt-image pair was most effective. All the data from the study session was collected by recording the input prompt and output image sequences in digital forms.

Curation of the images: A set of three images was curated based on the following criteria: The first image featured simple, easily recognisable architectural elements and spatial compositions, such as arches, temples, repetition, symmetry, and height, which were assumed to be easily defined by the participants. The second image presented a more complex form created by interlocking simple and recognisable shapes at different scales. The third comprised complex, organic, and fluid forms that were challenging to define and recognisable primarily through specific architectural reference and/or style keywords (Figure 1).

Figure 1 Initial image set used for the task sub-tasks (see online version for colours)

3.4 Data analysis

- *Qualitative and quantitative analysis of prompting behaviour:* To understand participants' prompt crafting behaviour, prompts were analysed quantitatively and qualitatively. The study aimed to evaluate how participants structured their prompts when tasked with regenerating initial architectural representations. The analysis encompassed the lexical choices, the taxonomy of the prompts (the structure of the prompts according to specific modifiers), and other qualitative insights gained from the prompts. The taxonomy of the prompts was analysed by assigning each specific word (*type*) in the prompt to a broader corresponding category (*modifier*) (e.g., column → architectural element, corridor → program, tall → scale, sunny → lighting, concrete → material, shiny → surface, etc.) Using NVivo software, all prompts were systematically coded to categorise words associated with specific attributes. This approach ensured a structured classification of terms, allowing for a more precise analysis of the linguistic patterns used in architectural descriptions.

- *Qualitative insights and CSI scores from surveys:* The Creativity Index (CSI) survey was conducted and analysed based on the methodology outlined in Cherry and Latulipe (2014). The survey evaluates six key dimensions of creativity support: Exploration, expressiveness, immersion, enjoyment, results worth effort, and collaboration. Participants provided ratings for each dimension, followed by paired comparisons to assess the relative importance of these factors. To ensure response reliability and minimise bias, the paired questions within each factor were randomised. The CSI score was calculated following the standard procedure described by the researchers.
  1. Individual factor scores were determined by summing the responses to the corresponding agreement statements.
  2. These factor scores were then weighted based on the number of times each factor was selected in the paired comparisons.
  3. The weighted scores were summed and normalised to generate a CSI score on a scale of 0 to 100.

  A higher score indicates better support for creative work. Beyond the SCI survey, additional survey questions were included to explore students' opinions and approaches toward generative T2I models.

4 Findings and results

4.1 Prompt lengths

Participants produced 456 prompts and outputs. The prompt lengths varied between 4 and 122 words, with an average of 38.76 (SD = 19.83) words. The lengths of prompts for sub-tasks are shown in Table 1. Across all three sub-tasks, 18 out of 28 students iterated their prompts the maximum allowed 7 times, while others had less iteration, with an average of 5.8 prompts per student. The iteration patterns suggest that most participants required multiple attempts to achieve their desired results. However, a lower number of iterations did not necessarily indicate success. Instead, it often reflected difficulties in controlling the tool and refining prompts effectively, causing some participants to abandon the task after the second or third attempt.

Table 1 Prompt lengths for subtasks

Sub-task	*Number of prompts*	*Shortest prompt*	*Longest prompt*	*Average length*	*Std. dev. (length)*
1	168	7 words	117 words	42.43	21.44
2	148	4 words	98 words	36.99	18.32
3	140	6 words	122 words	36.86	19.73

4.2 Linguistic patterns and taxonomies of the prompts

According to the systematic analysis of the prompts for each sub-task individually, the top-level categories/modifiers used by students appeared as a medium

b subject

c subject properties

d lighting and atmosphere

e context

f composition.

The use and frequency of top-level and sub-level modifiers varied depending on the initial image.

Sub-task 1

The most frequently used modifier for sub-task 1, after the *subject* modifier, was *light and atmosphere*, whereas this modifier was not used for sub-task 2 and was very rarely used for sub-task 3. Students tried to replicate the atmosphere of the initial image, which featured a strong sense of light and shadow. Regarding sub-categories describing the subject and subject's attributes, the most commonly used modifiers were *architectural elements* (e.g., column, aqueduct, wall) and *scale* (e.g., tall, long, high, narrow), followed by *colour* (e.g., black, dark grey), *geometrical characteristics* (e.g., rectangular, semi-circular), and *program* (e.g., corridor, hallway). In contrast, *material*, *subject composition*, and *surface* were rarely utilised (Table 2). The structure of prompts typically incorporated between four and ten modifiers with an average of seven.

Table 2 Modifiers used by students in prompts for sub-task 1

Modifiers	*No. of students*	*References*
Medium	18	89
Subject		
Architectural elements	28	228
Program	16	81
Subject properties		
Form/geometrical characteristics	17	80
Colour	17	85
Scale	24	169
Material	14	58
Composition	10	58
Surface	4	12
Light and atmosphere	24	215
Context	-	-
Composition	8	29

The word 'column' was the most frequently used subject modifier in the prompts, whereas 'arches' appeared rarely. This likely reflects students' limited knowledge of architectural terminology. As a result, they described the space using geometric terms such as semi-circular, curved, oval, and elliptical. Notably, only one student used 'pointed arch' in their prompts. Scale and lighting emerged as the most defining characteristics of the space. Students described it using terms like long, high, tall, and narrow while also focusing on the interplay of shadow, light, and darkness in the space (Figures 2 and 3).

Figure 2 The word cloud showing the most recurring words in sub-task 1 prompts

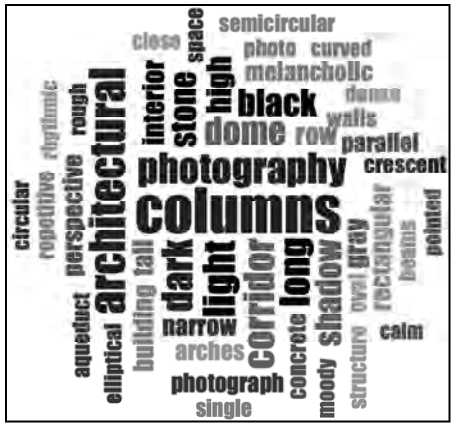
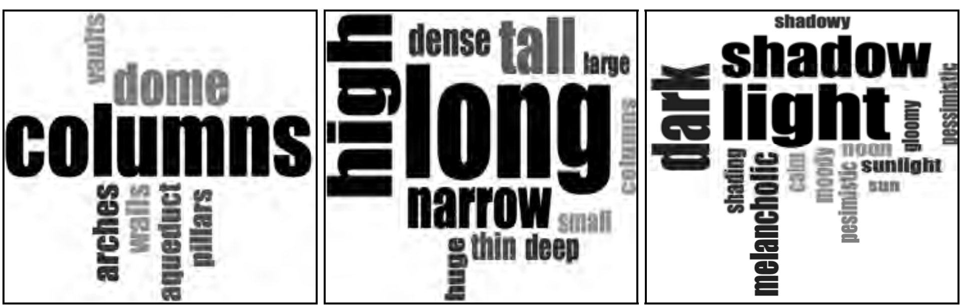
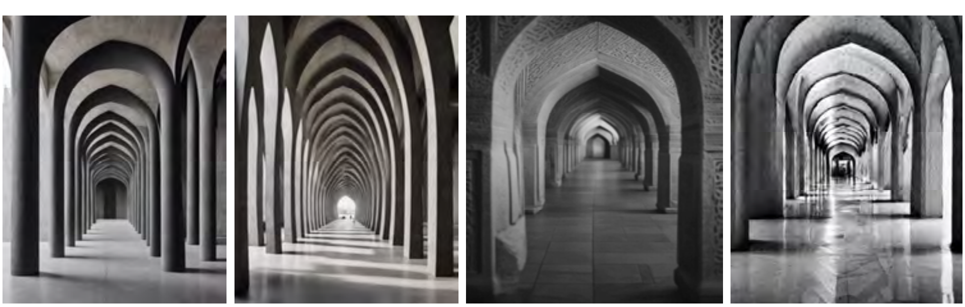


Figure 3 Word clouds showing the recurring words in the first three categories (sub-task 1), (a) architectural elements (b) scale (c) light and atmosphere



(a) (b) (c)

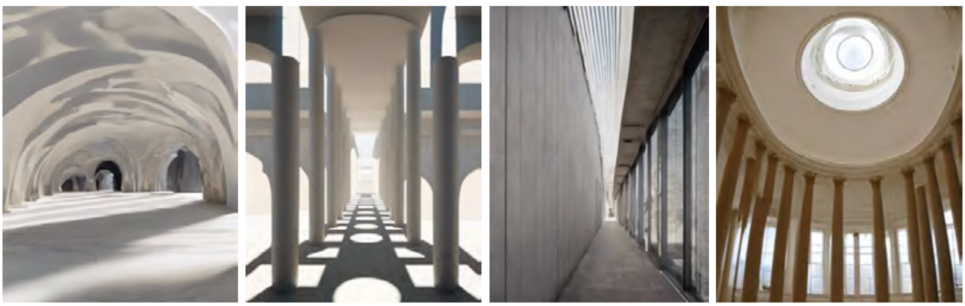
Figure 4 Outputs of prompts with the highest number of modifiers (a) (b) ten modifiers (c) (d) nine modifiers (see online version for colours)



(a) (b) (c) (d)

Results indicated that incorporating more modifiers, particularly those that precisely define the subject and subject characteristics, significantly improved the outputs. Generally, shorter prompts with a high number of well-chosen modifiers yielded better results, provided that the use of precise vocabulary effectively described the visual (Figure 4). In contrast, very short prompts with only a few modifiers failed to achieve the desired outcomes due to a lack of detailed visual descriptions (Figure 5).

Figure 5 Outputs of prompts with the lowest number of modifiers (a) (b) four modifiers (c) (d) five modifiers (see online version for colours)



(a) (b) (c) (d)

Sub-task 2

For sub-task 2, the most frequently used modifiers were *subject* and *context*, while the *composition* modifier was rarely employed, and *lighting* or *atmosphere* were not referenced at all. The most commonly used sub-modifiers included *subject composition* (e.g., irregular, vertical, complex, repetitive, overlapping, stacked), *form/geometrical characteristics* (e.g., cubes, prisms), *program* (e.g., structure, installation, construction, sculpture), and *scale* (e.g., small, large, varying sizes). While *colour properties* were used infrequently, no references to *material* or *surface properties* appeared in the prompts (Table 2). On average, prompts contained six modifiers, with a typical range of three to ten.

Table 2 Modifiers used by students in prompts for sub-task 2

Modifiers	*No. of students*	*References*
Medium	15	75
Subject		
Program	27	115
Subject properties		
Composition	22	184
Form/geometrical characteristics	27	151
Scale	18	114
Colour	16	64
Light and atmosphere	-	-
Context	19	115
Composition	8	35

The most used word across the prompts was ‘cube’. The term ‘cube’ primarily functioned as the defining geometric element of the whole composition, as seen in phrases like ‘a structure made of cubes’, ‘a structure formed by white cubes’, ‘a construction consisting of cube-shaped pieces’, and ‘an installation created by assembling cubes’. Findings revealed that the most challenging aspect was describing the subject composition. Students used 184 reference words and employed a diverse vocabulary to articulate it, including terms such as ‘vertical, complex, irregular, regular, stacked, randomly, intertwined, scattered’, among others. An interesting observation was that some students referred to the cubes as the subject itself, while others described them as part of a larger entity, using a broad range of words like ‘building’, ‘structure’, ‘installation’, ‘sculpture’, ‘construction’, or ‘artwork’ (Figures 6 and 7). Notably, the choice of the *program* had a significant impact on achieving the desired results (Figure 8).

Figure 6 The word cloud showing the most recurring words in the sub-task 2 prompts

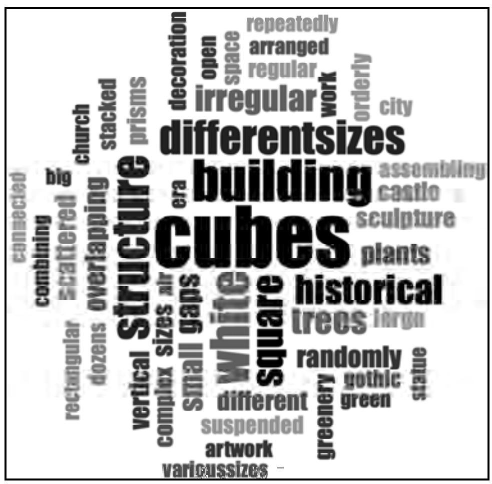
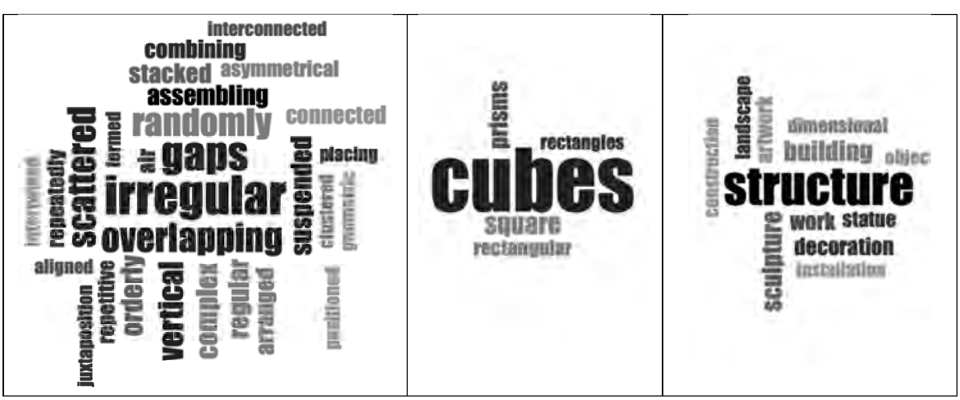


Figure 7 Word clouds showing the recurring words in the first three categories (sub-task 2), (a) subject composition (b) form/geometrical characteristics (c) program



(a) (b) (c)

Figure 8 Output images of different descriptions of the program (a) (b) structure (c) (d) cubes (e) (f) building (g) sculpture (see online version for colours)

Sub-task 3

For sub-task 3, the most frequently used modifiers were *subject*, *subject properties* and *medium*, whereas *light and atmosphere*, *context*, and *composition* were rarely employed. This can be attributed to the dominance of form in the provided image. The most utilised sub-modifiers, respectively, were *form and geometrical characteristics* (e.g., wavy, wave-shaped, curvy, twisted, folded, fluid, organic, inclined), *architectural elements* (e.g., roof, windows), *program* (e.g., building, museum), *material* (e.g., glass), and *scale* (e.g., large, small, higher) (Table 3). Notably, few students used the *style* or *architectural reference* modifier despite the image featuring a well-known building designed by the renowned architect Zaha Hadid, which belongs to a distinct architectural 'style' – *parametricism*. This can likely be attributed to a lack of architectural knowledge among students. Furthermore, prompts including 'Zaha Hadid' or 'Heydar Aliyev' did not necessarily produce the intended results (Figure 9).

The most frequently recurring words in the prompts were *building*, *glass and roof*, followed by *architectural*, *white*, *curves* and *photography* (Figure 10). The subject's program was generally defined as a building, with only a few participants providing more specific descriptions such as *museum* or *opera house*. *Roofs* and *windows* were the most commonly referenced architectural elements. Descriptors related to form and geometrical *characteristics* appeared most frequently, encompassing various terms such as *curves*, *oval*, *organic*, *round*, *fluid*, *wavy*, *twisted*, *sloping*, *inclined*, and *folding* (Figure 11). Beyond using adjectives to describe form, students experimented with detailed articulations to replicate the forms they observed. However, the results indicated that prompts emphasising conceptual clarity, rather than lengthy descriptions, produced better outputs. Highly detailed and overly long descriptions often led to misinterpretations by the algorithm (Figure 12).

Table 3 Modifiers used by students in prompts for sub-task 3

Modifiers	*No. of students*	*References*
Medium	15	76
Subject		
Architectural element	18	114
Program	23	96
Subject properties		
Form/geometrical characteristics	27	196
Material	19	74
Scale	14	71
Colour	17	72
Style	9	34
Surface	8	33
Architectural reference	6	11
Composition	4	16
Light and atmosphere	4	22
Context	4	15
Composition	9	37

Figure 9 Output images of prompts with *architectural reference*, (a) (b) ‘Heydar Aliyev’ (c) (d) ‘Zaha Hadid’ (see online version for colours)

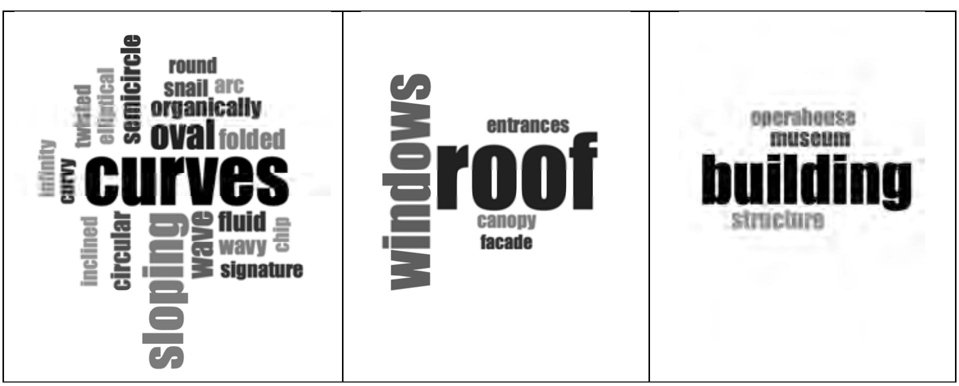
(a) (b) (c) (d)

The variation in prompt structures based on the intended output suggests that no single structure consistently yielded better results. Instead, both the prompt structure and the selected modifiers were tailored to the specific characteristics of the visual concept. However, the analysis revealed that focusing on the subject and incorporating sufficient modifiers related to its characteristics played a crucial role in achieving the desired outputs. In particular, articulating the program and architectural elements with a precise vocabulary made a significant improvement in outputs. Structured prompts that relied on key terms proved to be more effective than overly detailed or unstructured descriptions.

The quality of the generated images depended not only on the prompt taxonomy but also on the precision of the vocabulary and relevant terminology aligning with the intended visual concept.

Figure 10 Modifiers used by students in prompts for sub-task 3

Figure 11 Word clouds showing the recurring words in the first three modifiers (sub-task 3), (a) form/geometrical characteristics (b) architectural element (c) program (see online version for colours)



(a) (b) (c)

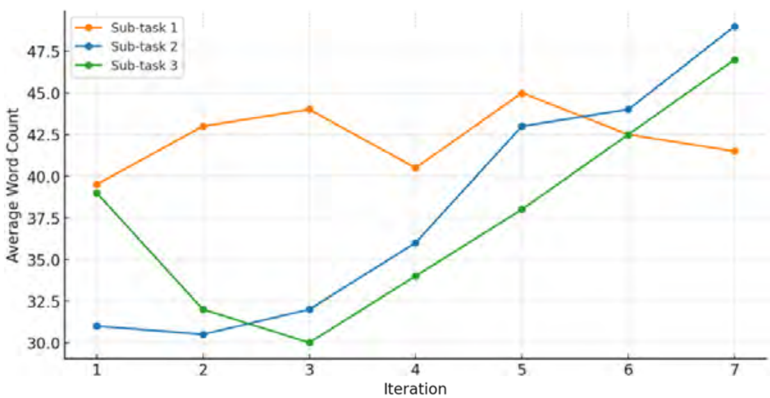
4.3 *Qualitative insights gained from the prompts*

- *Difficulty with optimising prompts*: One strategy participants used during prompt iteration involved removing prepositions and adding more details. Generally, they tend to write longer prompts through iterations. Only a few participants, refined their prompts by eliminating excess details, after expanding with additional details. In the first set of prompts, the prompt length fluctuated across iterations, increasing in some iterations while decreasing in others. In the second and third sets, prompt length steadily increased over iterations (Figure 13).

Figure 12 Image prompt pairs in sub-task 3 (see online version for colours)

		
Architectural photography of a beige building with a curved roof. The roof is C-shaped. The part of the roof in the middle looks like a canopy. The right side of the roof is higher than the left side. There is also a large window on the right side. The building is in a reasonably large area with a shiny floor. The photo of the building was taken from the outside.	Architectural photography, that is a building designed with curves to give the appearance of a signature. The front side is covered with glass.	The right side of the building will consist of a small semicircle with a slope. The left side will be an asymmetrical white building with a larger sloping semicircle. there will be glass under the ceiling. a new generation curved design that looks like a spaceport.

Figure 13 Average word count change over iterations (see online version for colours)



However, changes to prompts did not always produce predictable effects on the generated images. This unpredictability made it challenging for participants to optimise prompts for a desired output. The trial-and-error process did not follow a steady progression or a linear trajectory but required frequent steps backward before achieving improvements.

- *Difficulty with describing prompts:* As Barros and Ai (2004) note, articulating envisioned concepts through words can be challenging – a difficulty clearly reflected in participants' prompts. When attempting to communicate visual ideas, students often relied on vague expressions (e.g., "the light-shadow relationship between the columns draws attention") or analytical descriptions (e.g., "consisting of 20 identical columns"), both of which are difficult for algorithms to interpret.

- *Limited control over composition:* Unlike traditional visual artwork, which allows for manual specification of details and spatial arrangements, generative models operate automatically, limiting users' control over composition (Qiao et al., 2022). This limitation was particularly noticeable in participants' workflows, especially when dealing with organic, fluid forms and intricate compositions.
- *Difficulty with complex prompts:* Generative models often struggle with complex prompts, particularly those requiring multiple levels of detail or specific spatial arrangements (Chang et al., 2023). For instance, in the study, students faced greater difficulty articulating forms in sub-tasks 2 and 3 compared to the simpler forms in sub-task 1.

4.4 Evaluations of CSI scores

The CSI measures how well a tool supports creativity for a given task. It generated an overall CSI score out of 100, with a higher score indicating better creativity support. In this case, novice participants rated DreamStudio with an overall CSI score of 80.96, indicating strong creativity support. Also, the CSI scores were calculated for each individual factor. The average factor scores, factor counts and average weighted factor scores are shown in Table 4.

Table 4 CSI results from the survey (N = 28)

Scale	*Avg. factor score (SD) (out of 20)*	*Avg. factor counts (SD) (out of 5)*	*Avg. weighted factor score (SD) (out of 100)*
Exploration	17 (2.4)	3.5 (1.1)	60.2 (20.1)
Expressiveness	16 (2.9)	4 (1)	56.7 (22.6)
Results worth effort	16 (3)	3 (1)	43 (21.6)
Enjoyment	16 (3)	3 (1)	48 (23.5)
Immersion	14 (4)	1 (1)	20.6 (19)
Collaboration	15 (4)	1 (1)	13.6 (19.4)
Overall			*80.96 (11.43)*

The overall CSI score of 80.96 (SD = 11.43) suggests that DreamStudio provided strong creativity support for novice users. The relatively low standard deviation indicates that user experiences varied, but the scores were generally high and consistent. Among the six creativity factors, DreamStudio particularly supports exploration and creative *expression*, with the highest CSI scores. The highest factor counts suggest that participants also considered exploration and expressiveness key aspects of their creative process. Moderate scores for results worth effort and enjoyment indicate that participants found DreamStudio engaging and enjoyable to a certain extent. The moderate factor count (3/5) suggests that while enjoyment was valued, exploration and expressiveness were slightly more critical. Users generally perceived the output quality as worth their effort, though this factor was not their primary concern in the creative process. Low immersion scores suggest users were not fully engaged or absorbed while using DreamStudio. However, a low factor count (1/5) indicates that immersion was not a significant priority for participants in this task. The lowest weighted score was for collaboration, suggesting that DreamStudio does not strongly support collaborative work. However, since the task did not involve collaboration, this score should not be interpreted as a definitive limitation of the tool. To further refine the CSI evaluation, the results were analysed across two subscales: engagement and effectiveness (Table 5). DreamStudio performs well in engaging users creatively, though immersion could be improved. The lower effectiveness score suggests that collaboration and effort-reward balance could be enhanced.

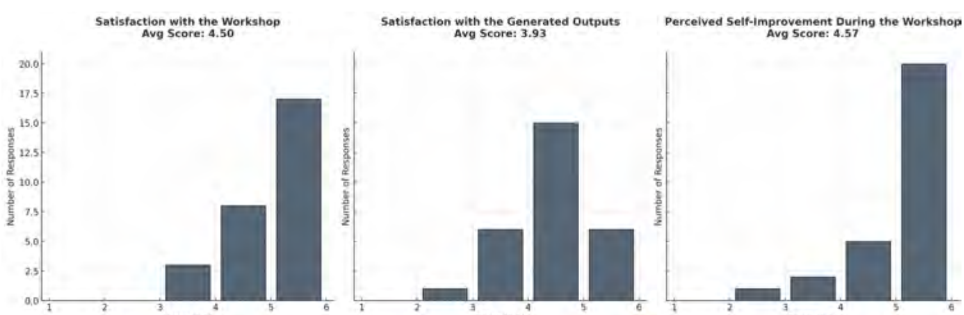
Table 5 The CSI scores of subfactors (N = 28)

Subscale	*Factors Included*	*CSI score*
Engagement	Exploration, expressiveness, immersion, enjoyment	46.4
Effectiveness	Results worth effort, collaboration	28.3

4.5 *Opinions of the participants*

Participants reported high satisfaction with participation in the study. The average satisfaction score was 4.5 out of 5, indicating a high level of approval among participants. Most respondents rated their experience as 4 or 5, with the lowest rating being 3. Participants rated the quality of the visuals they produced with an average score of 3.93 out of 5. While the feedback was generally positive, it was slightly lower than overall satisfaction. The majority of ratings fell between 3 and 5. The workshop was perceived as highly beneficial for learning and self-improvement, with an average score of 4.57 out of 5. Most participants gave ratings of 4 or 5, with only a low number rating lower (Figure 14).

Figure 14 Satisfaction with the workshop and generated outputs and perceived self-improvement (see online version for colours)



Common difficulties reported by participants included a lack of knowledge of architectural terminology and challenges with prompt writing techniques. Many struggled with crafting effective prompts, while a few found the limited DreamStudio credits to be a restriction. To address these challenges, students frequently relied on brainstorming with peers and Google searches to seek guidance for prompt crafting. A common suggestion among participants was the need for structured guidance on prompt engineering, focusing on improving effective prompt writing techniques.

5 Discussion

The findings underscore the dual nature of these technologies as a catalyst for creative exploration and a source of notable challenges, particularly in the realms of prompt articulation and compositional control. One of the primary challenges identified was translating complex visual concepts into effective text prompts. Many participants struggled to articulate architectural forms due to limited knowledge of domain-specific vocabulary. The study reveals that the precision of language and the strategic use of well-chosen modifiers significantly impact the quality of generated images. Moreover, prompt taxonomies must be tailored to the specific task or intended output to ensure optimal results. This underscores the need for structured training in prompt engineering, suggesting that a dedicated educational module on crafting effective prompts could significantly enhance user competence and satisfaction. Such training could include guided exercises, real-world case studies, and interactive prompt refinement workshops to improve users' ability to communicate complex design ideas effectively. Additionally, AI-assisted prompt generation tools could be developed to provide real-time suggestions and refinements based on users' descriptions, helping to bridge the gap between textual input and intended visual output.

Another major limitation was the lack of control over the composition of the generated images. Unlike traditional methods of architectural representation, which allow for direct manipulation of spatial elements and compositional details, the automated nature of current T2I systems restricts the user's ability to fine-tune outcomes. This limitation was particularly evident when participants attempted to recreate complex spatial arrangements. This gap indicates that further refinements in the interface design and underlying algorithms are necessary to offer users more intuitive control over the creative process. A potential solution is the development of hybrid workflows that combine AI generation with refinement and post-processing tools. T2I platforms should incorporate modular editing tools that allow users to adjust spatial elements post-generation. Features such as interactive layering, user-defined constraints, and digital sketching could provide designers with greater control over composition. Furthermore, refining AI models to better interpret explicit spatial instructions – such as those involving scale, symmetry, and spatial hierarchy – would improve the precision of generated images. Collaborations between AI developers and architects could facilitate the creation of domain-specific AI models tailored to architectural workflows.

The complexity of prompts also emerged as a significant factor influencing the quality of the generated images. While simple prompts tended to produce more predictable results, complex prompts – with multiple layers of detail and specific spatial instructions – often introduced inconsistencies and unpredictability in the output. The challenge of managing and optimising these complex prompts was compounded by the nonlinear nature of the iterative trial-and-error process observed during the experiment. Participants frequently oscillated between expanding and contracting their descriptions, underscoring the dynamic learning curve of effective prompt writing. This trial-and-error approach suggests that prompt engineering is an acquired skill requiring both structured learning and experiential practice. Given this, developing AI-driven prompt analysis tools could help users optimise their inputs by highlighting vague descriptions, suggesting alternative phrasings, and visualising probable outputs before final rendering. Additionally, adaptive learning systems – where the AI refines its responses based on user iterations- and real-time feedback mechanisms – that suggest modifications to prompts – could enhance the efficiency of the iterative prompt-engineering process.

Despite these challenges, the overall user experience with DreamStudio was predominantly positive. Participants rated the tool highly for its ability to foster creative exploration and expression, reporting intense satisfaction with its potential as a learning aid in architectural design. The iterative process, though challenging, encouraged users to experiment and refine their creative approaches, ultimately enhancing engagement with the task. However, lower scores in immersion and collaboration indicate that current T2I interfaces are not yet fully optimised for sustained engagement or collaborative design workflows. This raises questions about the broader applicability of GAI in collaborative architectural practices, where iterative dialogue and shared authorship are integral to the design process. To address this gap, future developments should integrate real-time collaborative features, allowing multiple users to co-create and modify AI-generated images simultaneously. Shared workspaces like cloud-based platforms, comment-based design iterations, and version history tracking could foster a more interactive and iterative design process, making AI tools more compatible with professional architectural workflows.

The findings highlight the need for both technological and educational improvements. On the technological front, there is a clear need for T2I models to evolve by incorporating more user-friendly interfaces that allow for greater control over composition and a better understanding of complex visual prompts. Enhancements in the algorithm's ability to process detailed and structured textual inputs could bridge the gap between user intent and visual output. From an educational perspective, integrating structured training on prompt engineering within architectural curricula could equip novice designers with the necessary skills to leverage these emerging tools effectively. This training should focus not only on technical proficiency but also on developing a precise visual vocabulary that aligns with architectural conventions, ensuring that users can articulate their design intent clearly and effectively.

Ultimately, the successful adoption depends not only on the evolution of AI technology but also on the refinement of human-AI interaction strategies. While current T2I generative systems present promising opportunities for creative exploration, the challenges highlighted in this study necessitate a concerted effort to improve both the underlying technology and the pedagogical frameworks that support their use. By addressing the limitations with targeted solutions, T2I systems can transition from experimental design aids to fully intelligent and creative partners in architectural education and practice.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Acknowledgements

This work was supported by the Kocaeli University Scientific Research Projects [Grant No. 3342].

Declarations

The participants were fully informed about the research objectives, procedures, and potential outcomes before enrolment in the study. The participants provided written informed consent, agreeing to participate in the research, to allow the publication of the study findings, and to permit the use of the images they generated during the study.

References

Albaghajati, Z.M., Bettaieb, D.M. and Malek, R.B. (2023) 'Exploring text-to-image application in architectural design: insights and implications', *Architecture, Structures Construction*, Vol. 3, pp.475–497, https://doi.org/10.1007/s44150-023-00103-x.

Barros, M. and Ai, Q. (2024) 'Designing with words: exploring the integration of text-to-image models in industrial design', *Digital Creativity*, Vol. 35, No. 4, pp.378–391, https://doi.org/10.1080/14626268.2024.2411223.

Boden, M.A. (2001) 'Creativity and knowledge', in Craft, A., Jeffrey, B. and Leibling, M. (Eds.): *Creativity in Education*, pp.95–102, Continuum, London.

Boden, M.A. (2004) *The Creative Mind: Myths and Mechanisms*, 2nd ed., Routledge, London.

Bolojan, D. and Vermisso, E. (2020) 'Deep learning as heuristic approach for architectural concept generation', *Proceedings of the 11th International Conference on Computational Creativity (ICCC'20)*, pp.98–105.

Boone, L.W. and Hollingsworth, A.T. (1990) 'Creative thinking in business organizations', *Review of Business*, Vol. 12, No. 2, pp.3–13.

Bronowski, J. (1985) 'The creative process', *Leonardo*, Vol. 18, No. 4, pp.245–248, https://www.jstor.org/stable/1578075.

Brown, T.B., Mann, B., Ryder, N., Kaplan, J., Dhariwal, P. et al. (2020) 'Language models are few-shot learners', in Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M.F. and Lin, H. (Eds.): *Advances in Neural Information Processing Systems (Proceedings of the NeurIPS 2020)*, Vancouver, Canada, Vol. 33, https://papers.nips.cc/paper_files/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf.

Carroll, E.A. and Latulipe, C. (2009) 'The creativity support index', *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, Association for Computing Machinery, pp.4009–4014, https://doi.org/10.1145/1520340.1520609.

Chalmers, D.J. (1996) *The Conscious Mind: In Search of a Fundamental Theory*, Oxford University Press, New York.

Chandrasekera, T., Hosseini, Z., Perera, U. and Bazhaw-Hyscher, A. (2024) 'Generative artificial intelligence tools for diverse learning styles in design education', *International Journal of Architectural Computing*, https://doi.org/10.1177/14780771241287345.

Chang, H., Zhang, H., Barber, J., Maschinot, A., Lezama, J., Jiang, L., Yang, M., Murphy, K.P., Freeman, W.T., Rubinstein, M., Li, Y. and Krishnan, D. (2023) 'Muse: text-to-image generation via masked generative transformers', *Proceedings of the 40th International Conference on Machine Learning*, Vol. 202, pp.4055–4075, https://proceedings.mlr.press/v202/chang23b.html.

Cherry, E. and Latulipe, C. (2014) 'Quantifying the creativity support of digital tools through the creativity support index', *ACM Transactions on Computer-Human Interaction*, Vol. 21, No. 4, Article 21, pp.1–25.

Cheung, L.H. and Dall'Asta, J.C. (2024) 'Human-computer interaction (HCI) approach to artificial intelligence in education (AIEd) in architectural design', *EÍDOS – Revista Científica de Arquitectura y Urbanism*, Vol. 23, pp.109–131.

Dang, H., Mecke, L., Lehmann, F., Goller, S. and Buschek, D. (2022) *How to Prompt? Opportunities and Challenges of Zero- and Few-Shot Learning for Human – AI Interaction in Creative Applications of Generative Models*, arXiv preprint, arXiv:2209.01390v1, https://doi.org/10.48550/arXiv.2209.01390.

Deckers, N., Fröbe, M., Kiesel, J., Pandolfo, G., Schröder, C., Stein, B. and Potthast, M. (2023) 'The infinite index: information retrieval on generative text-to-image models', in *Proceedings of the 2023 Conference on Human Information Interaction and Retrieval (CHIIR'23)*, pp.172–186, Association for Computing Machinery, https://doi.org/10.1145/3576840.3578327.

del Campo, M. and Carlson, A.I. (2022) 'Strange but familiar enough: reinterpreting style in the context of AI', in Chaillou, S. (Ed.): *Artificial Intelligence and Architecture: From Research to Practice*, pp.172–179, Birkhauser, Basel.

Desouki, M., El-Haddad, T.A. and El-Boshey, B. (2023) 'Revolutionary artificial intelligence architectural design solutions; is it an opportunity or a threat?', *Mansoura Engineering Journal*, Vol. 48, No. 6, Article 11, https://doi.org/10.58491/2735-4202.3091.

Dhariwal, P. and Nichol, A. (2021) 'Diffusion models beat GANs on image synthesis', in Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P.S. and Wortman Vaughan, J. (Eds.): *Advances in Neural Information Processing Systems 34 (NeurIPS 2021)*, https://proceedings.nips.cc/paper_files/paper/2021.

Floridi, L. and Sanders, J.W. (2004) 'On the morality of artificial agents', *Minds and Machine*, Vol. 14, pp.349–379, https://doi.org/10.1023/B:MIND.0000035461.63578.9d.

Gero, J. (1994) 'Towards a model of exploration in computer-aided design', in *Formal Design Methods for CAD, Proceedings of the IFIP TC5/WG5.2 Workshop on Formal Design Methods for CAD*, pp.315–336.

Guilford, J.P. (1950) 'Creativity', *American Psychologist*, Vol. 5, No. 9, pp.444–454.

Hertzmann, A. (2018) 'Can computers create art?', *Arts*, Vol. 7, No. 2, p.18, https://doi.org/10.3390/arts7020018.

Iranmanesh, A. and Lotfabad, P. (2024) 'Critical questions on the emergence of text-to-image artificial intelligence in architectural design pedagogy', *AI & Society*, https://doi.org/10.1007/s00146-024-02111-x.

Jaruga-Rozdolska, A. (2022) 'Artificial intelligence as part of future practices in the architect's work: midjourney generative tool as part of creating an architectural form', *Architectus*, Vol. 3, No. 71, pp.95–104, https://doi.org/10.37190/arc220310.

Kampylis, P.G. and Valtanen, J. (2010) 'Redefining creativity: analyzing definitions, collocations, and consequences', *The Journal of Creative Behavior*, Vol. 44, No. 3, pp.191–214.

Kantosalo, A. and Toivonen, H. (2016) 'Modes for creative human-computer collaboration: alternating and task-divided co-creativity', in Pachet, F., Cardoso, A., Corruble, V. and Ghedini, F. (Eds.): *Proceedings of the Seventh International Conference on Computational Creativity (ICCC 2016)*, Sony CSL, pp.77–84, http://www.computationalcreativity.net/iccc2016/proceedings-2016/.

Karimi, P., Rezwana, J., Siddiqui, S., Maher, M.L. and Dehbozorgi, N. (2020) 'Creative sketching partner: an analysis of human-AI co-creativity', in *Proceedings of the 25th International Conference on Intelligent User Interfaces*, Cagliary, Italy, pp.221–230, https://doi.org/10.1145/3377325.3377522.

Koestler, A. (1964) *The Act of Creation*, Hutchinson, London.

Liu, V. and Chilton, L.B. (2022) 'Design guidelines for prompt engineering text-to-image generative models', in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*, Article 384, pp.1–23, https://doi.org/10.1145/3491102.3501825.

Logan, B. and Smitthers, T. (1992) 'Creativity and design as exploration', in Gero, J.S. and Maher, M.L. (Eds.): *Modelling Creativity and Knowledge-Based Creative Design*, Lawrence Earlbaum, Hillsdale, NJ.

Maher, M.L. (2000) 'A model of co-evolutionary design', *Engineering with Computers*, Vol. 16, pp.195–208, https://doi.org/10.1007/PL00013714.

Maher, M.L. (2012) 'Computational and collective creativity: who's being creative?', in *Proceedings of The Third International Conference on Computational Creativity*, University College Dublin, pp.67–71.

Maher, M.L., Poon, J. and Boulanger, S. (1996) 'Formalising design exploration as co-evolution: a combined gene approach', in Gero, J.S. and Sudweeks, F. (Eds.): *Advances in Formal Design Methods for CAD*, Springer, pp.3–30, https://doi.org/10.1007/978-0-387-34925-1_1.

Manovich, L. (2018) *AI Aesthetics*, Strelka Press, Moscow.

Marrone, R., Cropley, D. and Medeiros, K. (2024) 'How does narrow AI impact human creativity?', *Creativity Research Journal*, pp.1–11, https://doi.org/10.1080/10400419.2024.2378264.

McCormack, J., Gifford, T. and Hutchings, P. (2019) 'Autonomy, authenticity, authorship and intention in computer generated art', in Ekárt, A., Liapis, A. and Pena, M.L.C. (Eds.): *Computational Intelligence in Music, Sound, Art and Design, EvoMUSART 2019, Lecture Notes in Computer Science*, Springer, Vol. 11453, https://doi.org/10.1007/978-3-030-16667-0_3.

Mednick, S.A. (1962) 'The associative basis of the creative process', *Psychological Review*, Vol. 69, No. 3, pp.220–232.

Mumford, M.D. (2003) 'Where have we been, where are we going? Taking stock in creativity research', *Creativity Research Journal*, Vol. 15, Nos. 2/3, pp.107–120.

Mumford, M.D. and Gustafson, S.B. (1988) 'Creativity syndrome: integration, application, and innovation', *Psychological Bulletin*, Vol. 103, No. 1, pp.27–43.

Oppenlaender, J. (2022) 'The creativity of text-to-image generation', in *Proceedings of the 25th International Academic Mindtrek Conference (Academic Mindtrek '22)*, pp.192–202, https://doi.org/10.1145/3569219.3569352.

Oppenlaender, J. (2023) 'A taxonomy of prompt modifiers for text-to-image generation', *Behaviour & Information Technology*, Vol. 43, No. 15, pp.3763–3776, https://doi.org/10.1080/0144929X.2023.2286532.

Oppenlaender, J., Linder, R. and Silvennoinen, J. (2024) 'Prompting AI art: an investigation into the creative skill of prompt engineering', *International Journal of Human-Computer Interaction*, pp.1–23, https://doi.org/10.1080/10447318.2024.2431761.

Paananen, V., Oppenlaender, J. and Visuri, A. (2023) 'Using text-to-image generation for architectural design ideation', *International Journal of Architectural Computing*, Vol. 22, No. 3, pp.458–474, https://doi.org/10.1177/14780771231222783.

Pena, M.L., Adrian, C., Rodríguez-Fernandez, N., Santos, I. and Romero, J. (2021) 'Artificial intelligence applied to conceptual design. A review of its use in architecture', *Automation in Construction*, Vol. 124, p.103505, https://doi.org/10.1016/j.autcon.2021.103550.

Ploennigs, J. and Berger, M. (2023) 'AI art in architecture', *AI in Civil Engineering*, Vol. 2, No. 8, https://doi.org/10.1007/s43503-023-00018-y.

Qiao, H., Liu, V. and Chilton, L. (2022) 'Initial images: using image prompts to improve subject representation in multimodal AI generated art', in *Proceedings of the 14th Conference on Creativity and Cognition (C&C '22)*, Association for Computing Machinery, pp.15–28, https://doi.org/10.1145/3527927.3532792.

Radford, A., Wook Kim, J., Hallacy, C., Ramesh, A. et al. (2021) 'Learning transferable visual models from natural language supervision', *Proceedings of the 38th International Conference on Machine Learning, PMLR*, Vol. 139, pp.8748–8763, https://proceedings.mlr.press/v139/radford21a.

Rafizadeh, H., Teixeira, M.B.F., Donovan, J. and Schork, T. (2024) 'Evolving architectural paradigms: a study of levels of automation in architecture', in *Accelerated Design, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024*, Vol. 3, pp.79–88.

Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., Chen, M. and Sutskever, I. (2021) 'Zero-shot text-to-image generation', *Proceedings of the 38th International Conference on Machine Learning, PMLR*, Vol. 139, pp.8821–8831, https://proceedings.mlr.press/v139/ramesh21a.html.

Reynolds, L. and McDonell, K. (2021) 'Prompt programming for large language models: beyond the few-shot paradigm', in *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21)*, Article 314, pp.1–7, Association for Computing Machinery, https://doi.org/10.1145/3411763.3451760.

Rezwana, J. and Maher, M.L. (2022) 'Designing creative AI partners with COFI: a framework for modelling interaction in human-AI co-creative systems', *ACM Transactions on Computer-Human Interaction*, Vol. 30, No. 5, pp.1–28, https://doi.org/10.1145/3519026.

Rhodes, M. (1961) 'An analysis of creativity', *Phi Delta Kappan*, Vol. 42, No. 7, pp.305–310, https://www.jstor.org/stable/20342603.

Ribot, T. (1906) in Baron, A.H.N. (Ed.): *Essay on the Creative Imagination*, Trans., The Open Court Publishing Company, London.

Robertson, J., Ferreira, C., Botha, E. and Oosthuizen, K. (2024) 'Game changers: a generative AI prompt protocol to enhance human-AI knowledge co-construction', *Business Horizons*, Vol. 67, pp.499–510, https://doi.org/10.1016/j.bushor.2024.04.008.

Rogers, C.R. (1954) 'Toward a theory of creativity', *ETC: A Review of General Semantics*, Vol. 11, No. 4, pp.250–258.

Rombach, R., Blattmann, A., Lorenz, D., Esser, P. and Ommer, B. (2022) 'High-resolution image synthesis with latent diffusion models', in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.10674–10685, https://doi.org/10.1109/CVPR52688.2022.01042.

Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E., Ghasemipour, S.K.S., Ayan, B.K., Mahdavi, S.S., Lopes, R.G. et al. (2022) 'Photorealistic text-to-image diffusion models with deep language understanding', in Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K. and Oh, A. (Eds.): *Advances in Neural Information Processing Systems (NeurIPS 2021)*, Vol. 35, https://proceedings.neurips.cc/paper_files/paper/2022.

Sanchez, T. (2023) 'Examining the text-to-image community of practice: why and how do people prompt generative AIs?', in *Proceedings of the 15th Conference on Creativity and Cognition*, pp.43–61, https://doi.org/10.1145/3591196.3593051.

Sasso, J. (1980) 'The stages of the creative process', *Proceedings of the American Philosophical Society*, Vol. 124, No. 2, pp.119–132, https://www.jstor.org/stable/986207.

Stein, M. (1953) 'Creativity and culture', *Journal of Psychology*, Vol. 36, No. 2, pp.311–322.

Sternberg, R.J. and Lubart, T. (1999) 'The concept of creativity: prospects and paradigms', in Sternberg, R. (Ed.): *Handbook of Creativity*, pp.3–15, Cambridge University Press, New York.

Stigsen, M.B., Moisi, A., Rasoulzadeh, S., Schineggeer, K. and Rutzinger, S. (2023) 'AI diffusion as design vocabulary: investigating the use of AI image generation in early architectural design and education', in Dokonal, W., Hirschberg, U. and Wurzer, G. (Eds.): *Digital Design Reconsidered – Proceedings of the 41st Conference on Education and Research in Computer Aided Architectural Design in Europe (eCAADe 2023)*, Graz, 20–22 September, Vol. 2, pp.587–596, https://doi.org/10.52842/conf.ecaade.2023.2.587.

Tan, L. and Luhrs, M. (2024) 'Using generative AI midjourney to enhance divergent and convergent thinking in an architect's creative design process', *The Design Journal*, Vol. 27, No. 4, pp.677–699, https://doi.org/10.1080/14606925.2024.2353479.

Torrance, E.P. (1993) 'Understanding creativity: where to start?', *Psychological Inquiry*, Vol. 4, No. 3, pp.232–234, https://www.jstor.org/stable/1448974.

Van den Oord, A., Vinyals, O. and Kavukcuoglu, K. (2017) 'Neural discrete representation learning', in von Luxburg, U., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S. and Garnett, R. (Eds.): *Advances in Neural Information Processing Systems (Proceedings of the NeurIPS 2017)*, Guyon, Vol. 30, https://papers.nips.cc/paper_files/paper/2017/file/7a98af17e63a0ac09ce2e96d03992fbc-Paper.pdf.

Van Epps, H., Astudillo, O., del Pozo Martín, Y. and Marsh, J. (2022) 'The sex and gender equity in research (SAGER) guidelines: implementation and checklist development', *European Science Editing*, Vol. 48, p.e86910, https://doi.org/10.3897/ese.2022.e86910.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I. (2017) 'Attention is all you need', in von Luxburg, U., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S. and Garnett, R. (Eds.): *Advances in Neural Information Processing Systems (Proceedings of the NIPS 2017)*, Guyon, Vol. 30, https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.

Vernon, P.E. (1989) 'The nature-nurture problem in creativity', in Glover, J.A., Ronning, R.R. and Reynolds, C.R. (Eds.): *Handbook of Creativity: Perspectives on Individual Differences*, pp.93–110, Plenum Press, New York.

Welsch, P.K. (1980) *The Nurturance of Creative Behavior in Educational Environments: A Comprehensive Curriculum Approach*, Unpublished Doctoral dissertation, University of Michigan.

Wingström, R., Hautala, J. and Lundman, R. (2022) 'Redefining creativity in the era of AI? Perspectives of computer scientists and new media artists', *Creativity Research Journal*, Vol. 36, No. 2, pp.177–193, https://doi.org/10.1080/10400419.2022.2107850.

Xie, Y., Pan, Z., Ma, J., Jie, L. and Mei, Q. (2023) 'A prompt log analysis of text-to-image generation systems', in *Proceedings of the ACM Web Conference 2023*, Association for Computing Machinery, pp.3892–3902, https://doi.org/10.1145/3543507.3587430.

Websites

https://www.midjourney.com/home.

https://www.lookx.ai/.

https://leonardo.ai/.