Design Research Society
DRS Digital Library

IASDR Conference Series

Dec 2nd, 9:00 AM - Dec 5th, 5:00 PM

The Unseen Hand: The Influence of Visual and AI Literacy on AI Text-to-Image Generation

Alessandro Canossa
Royal Danish Academy, Copenhagen, Denmark

Lisa Toender
Royal Danish Academy, Copenhagen, Denmark

Lucas Fellner
Royal Danish Academy, Copenhagen, Denmark

Jesper Juul
Royal Danish Academy, Copenhagen, Denmark

Willem Van der Maden
IT University of Copenhagen, Copenhagen, Denmark

See next page for additional authors

Follow this and additional works at: https://dl.designresearchsociety.org/iasdr

Part of the Art and Design Commons

Citation

Canossa, A., Toender, L., Fellner, L., Juul, J., Maden, W.V.,and Zhu, J.(2025) The Unseen Hand: The Influence of Visual and AI Literacy on AI Text-to-Image Generation, in Chang, C.-Y., Chen, C.-H., & Hsu, Y. (eds.), *IASDR 2025: Design Next*, 02-05 December, Taipei, Taiwan. https://doi.org/10.21606/iasdr.2025.1177

This Research Paper is brought to you for free and open access by DRS Digital Library. It has been accepted for inclusion in IASDR Conference Series by an authorized administrator of DRS Digital Library. For more information, please contact library@thedrs.org.

Authors
Alessandro Canossa, Lisa Toender, Lucas Fellner, Jesper Juul, Willem Van der Maden, and Jichen Zhu

This research paper is available at DRS Digital Library: https://dl.designresearchsociety.org/iasdr/iasdr2025/fullpapers/337

International Association of Societies
of Design Research Congress 2025
DESIGN NEXT

Taipei, Taiwan
2-5 December 2025

The Unseen Hand: The Influence of Visual and AI Literacy on AI Text-to-Image Generation

Canossa, Alessandro[*a]; Toender, Lisa[b]; Fellner, Lucas[a]; Juul, Jesper[a]; Van der Maden, Willem[b]; Zhu, Jichen[b]

[a] Royal Danish Academy, Copenhagen, Denmark
[b] IT University, Copenhagen, Denmark C
* acan@kglakademi.dk

Text-To-Image (TTI) generators are becoming widely used and are often promoted as "democratizing" the making of images independently of the skill set of the user. But relatively little is known about whether people with different educations, skill sets and literacies use these tools differently, and how their backgrounds influence the quality of the results. In this article we investigate the impact of Visual Literacy (VL), AI Literacy (AIL) and Prompt Engineering Literacy (PEL) on prompt use. More precisely we are examining how they each impact prompt usage patterns, linguistic and semantic composition, and the structure and morphology of prompts. Additionally, we employed Process Analysis to investigate how participants with different literacy approach the creative design process with TTI. Our results show that individuals scoring high on Visual Literacy (VL) tend to employ a larger vocabulary and more nuanced language, greater prompt variety and make more references to art genres, styles, and movements. In contrast, participants scoring high AI Literacy (AIL) demonstrated a deeper understanding of AI interaction, influencing their prompting strategies, and tended to adhere closely to the exact wording of the brief, treating it as precise specifications. The study's findings have the potential of significantly impacting educations, both by actively teaching students to experiment with prompt variations and deepen expressive visual vocabulary as well as designing curricula that incorporate structured exercises on iterative prompt refinement, explicitly teaching how to adjust prompts based on AI output to achieve desired visual outcomes.

1 Introduction

The advent of text-to-image (TTI) generation tools such as *Dall-E* (2024. Dall-E) and *MidJourney* (2024a. Midjourney), powered by generative artificial intelligence (AI), has revolutionized the way users create visual content. These tools, which translate textual prompts into images, have found applications in various domains, including art (Sivertsen et al. 2024), design (Zhou et al. 2024), education, and marketing, offering new creative possibilities and processes. This paper explores how factors such as users' education in visual design and information technology, technical proficiency with generative AI influences the use and perception of text-to-image generators.

Previous research has highlighted the importance of user background in the adoption and effective use of AI technologies (Druga et al. 2022). For instance, studies have shown that users with higher levels of education and technical literacy are more adept at leveraging AI tools for complex tasks (Han

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International Licence.
http://creativecommons.org/licenses/by-nc/4.0/

https://doi.org/10.21606/iasdr.2025.1177

et al. 2024). Additionally, there is some evidence that technical knowledge might play a crucial role in the ability to effectively use complex AI tools, with more experienced users often achieving better results and having a more positive experience (Stolpe and Hallström 2024). Furthermore, prior exposure to AI technologies can shape user expectations and satisfaction, as familiarity with AI capabilities can lead to more realistic and informed interactions (Morandini et al. 2023). Despite these insights, there remains a gap in understanding precisely how these factors collectively influence the process of generating images with text-to-image tools.

To understand how different user groups engage with TTI tools and what background factors contribute to their success or challenges, we conducted a mixed-method exploratory study on based on 25 users, using a real-world design task. Our prior work (Canossa et al. 2025) aimed at providing a more fine-grained view of performance by evaluating not only the aesthetic value of generated images, but also if they are appropriate for a realistic function in a professional setting. The results of our preliminary study found no impact of any literacy (Visual, AI or Prompt Engineering) on the general visual appeal of the generated images, participants with high AI literacy reported more understanding of the TTI process, while images created by participants with high visual literacy were rated as better fulfilling the stated visual design task as such. The focus of this paper is to investigate the impact of Visual Literacy, AI Literacy and Prompt Engineering Literacy on prompt use. More precisely we are examining how the three different types of literacy impact users' prompt usage patterns, linguistic and semantic composition, structure and morphology of prompts. Specifically, we seek to answer the following research questions (RQs):

RQ1 - The Impact of Visual Literacy, AI Literacy and Prompt Engineering Literacy on prompt use: How do the three different types of literacy (VL, AIL, PEL) impact prompt usage patterns, linguistic and semantic composition, structure and morphology of prompts?

RQ2 - AI-Assisted Creativity Process Analysis How do participants with different literacy approach the design process with TTI?

We answer RQ1 by testing the following three hypotheses:

• H1 Participants with higher visual literacy will use a larger vocabulary and will have finer-grained control of the domain.

• H2 Participants with higher visual literacy will make more references to artists and art movements as well as to style in general.

• H3 Larger vocabulary, use of expert language, and use of artistic references can predict Visual Appeal.

And we answer RQ2 by analyzing the participants' think-aloud and interview data through process analysis.

Our process analysis revealed that participants with high AI literacy report a deeper understanding of the computational creativity process. They are more willing to accept a co-creation position with TTI tools, even if this understanding has no impact on the quality of the images produced. Meanwhile, participants with high visual literacy expect a higher degree of control and are somewhat unsatisfied with the process. We found that higher prompt engineering and AI literacy does not increase participants' perceived control on the computational creativity process.

This insight shows that design education needs to focus on teaching students to translate sophisticated visual ideas into effective textual prompts and to manage expectations regarding AI control, exploring hybrid workflows where AI provides ideation. It also showed that Prompt Engineering Literacy develops dynamically, suggesting that educational programs should incorporate hands-on iterative prompting exercises. For technical education, process analysis indicated that students with high AI Literacy have a deeper understanding of computational creativity, implying a need to promote insights into AI's interpretative mechanisms to foster better AI-assisted design process.

2 Related Work

2.1 Generative AI in Design

Text-to-Image (TTI) models produce visual data from text prompts (Feuerriegel et al. 2023). TTI tools are a recent and rapidly developing field, moving towards human-AI co-creation with significant socio-technical implications (Oppenlaender 2022). While TTI models have been shown to improve efficiency in the ideation phases of visual design, there is debate whether they make design education obsolete, allowing anyone to be creative with TTI, or if AI should function to augment designers (Feuerriegel et al. 2023; Van Der Maden et al. 2024). This underscores the importance of understanding how humans, especially designers, utilize TTI tools in their creative process (Ardhianto and Nababan 2023). Research indicates benefits like aiding ideation and supporting the representation of design ideas (Feldman et al. 2017). It's argued that AI should primarily augment designers based on their expertise, rather than replacing them, and AI can help open creative perspectives (Hanafy 2023). While models for this process exist and evidence suggests AI assists various design stages, obstacles include the technology's early stage, problems with realism, and potential prompt misunderstanding (Wu et al. 2021). Despite acknowledging these advantages and limitations (Lin et al. 2024; Oppenlaender 2022; Van Der Maden et al. 2023; Wadinambiarachchi et al. 2024; Zhu et al. 2018), there is a significant lack of research on the human-AI design process itself and how AI might either eliminate or create the need for specific human competencies in the design space (Turchi et al. 2023).

2.2 Prompt Engineering and Prompt Analysis for Text-to-Image

Prompt engineering involves crafting effective prompts to achieve desired AI outputs (Jiang et al. 2022a, b; Reynolds and McDonell 2021; Yang et al. 2022). Challenges include managing task granularity, abstraction levels, request scoping, and users' potentially inaccurate mental models of the AI. Factors like lengthy instructions can negatively affect effectiveness, while word order is important. Strategies like constraining vocabulary and using examples can be employed (Liu et al. 2023; Lu et al. 2022; O'Connor and Andreas 2021).

Existing work on analyzing users' prompts for TTI often uses large datasets such as the Midjourney Discord dataset (2024b. Midjourney Discord Dataset) or DiffusionDB (2024. DiffusionDB). Borrowing techniques from web query analysis, researchers have used topic modeling to develop taxonomies of specifiers used in prompts. Examples of taxonomies include Sanchez's subject, medium, influence, light, color, composition, detail, context, and Xie et al.'s subject, form, and intent (Sanchez 2023; Xie et al. 2023). Other researchers use smaller-scale user studies, such as using interviews and observations, to uncover user strategies in depth. For example, Almeda et al. identified strategies like

Iterative Prompt Exploration, Parametric Manipulation, and Semantic Explorations (Mahdavi Goloujeh et al. 2024). Goloujeh et al. noted strategies like "overview and details" (where adjectives increase efficiency), "omitted words," "reordering or rephrasing," "refining parameters," "reroll and make variations," and "descriptive sentences" (Almeda et al. 2024; Zamfirescu-Pereira et al. 2023). Finally, ethnographic approaches have also been adopted. For instance, Oppenlaender provided insights into both taxonomy (including Subject, Style modifier, Quality booster) and strategies (Define, Modify, Solidify, Vary) for iterative prompt writing (Oppenlaender 2022).

2.3 Visual, AI, and Prompt Engineering literacy

In order to operationalize users' relevant skills, we need to investigate the foundational concepts of literacy pertinent to AI-assisted visual design. We intend to challenge the notion that TTI tools can entirely level the creative playing field, as claimed by tech companies such as Midjourney. Our previous work showed that a user's background does indeed still influence the quality of outputs and the human-AI collaboration process (Canossa et al. 2025). We adopt UNESCO's definition of literacy as "the ability to identify, understand, interpret, create, communicate and compute, using printed and written materials associated with varying contexts" (2024. UNESCO). Within this framework, we define three specific types of literacy relevant to the study, and we will profile al the users in our sample according to these literacies:

Visual Literacy (VL) refers to the ability to "read and write visual language" – to successfully interpret and compose meaningful visual messages (Avgerinou and Ericson 1997). For our study, we are particularly interested in foundational visual design competency of 2D composition. It encompasses the competency of arranging visual elements such as line, shape, color, and brightness, and applying design principles like balance and symmetry. This competency is typically part of the visual arts and design education (Gardner 1970; Ruskin 1857).

AI Literacy (AIL) refers the competencies enabling individuals to critically evaluate AI technologies, communicate and collaborate effectively with AI, and use AI as a tool (Long and Magerko 2020). It is built upon digital literacy and involves aspects of knowing, understanding, using, applying, evaluating, and creating, implying that individuals trained in using or building AI tools would possess a certain level of this literacy (Ng et al. 2021).

Prompt Engineering Literacy (PEL) refers to the ability to communicate with and direct generative AI systems without requiring computer programming expertise (Maloy and Gattupalli 2024). While some suggest a degree of AI literacy is necessary for complete PEL (Knoth et al. 2024; Maloy and Gattupalli 2024), the study operationally treats it independently because generative AI tools are becoming increasingly accessible to designers without AI knowledge. We thus focus on users' practical experience with the specific syntax and vocabulary of TTI prompting.

3 Methods

To ensure a meaningful comparison, we used a purposive sampling strategy to recruit participants from technical programs (e.g., Computer Science and Data Science) at a technical university and from visual design programs at a design college in a large Northern European city. We used a mix of on-site, online, and snow-ball recruitment. We asked participants to answer a questionnaire that helped us profile them according to Visual, AI and Prompt Engineering literacies and selected only participants with relatively high technical or visual literacy. We only included students in their last year of Bachelor's or in their Masters's' study to ensure relatively high expertise in their areas of study. The procedure followed human subject research guidelines at the authors' institutions as well as relevant GDPR rules for data storage.

3.1 Procedure

After informed consent, participants were asked to provide basic demographic information such as age, gender, and educational background. Their major (visual design, CS/data science) provides a strong indicator of their literacy in that area but additionally we administered participants a short survey (11 4-point Lickert-scale items) to further measure their competency in all three types of literacy, since participants may have high literacy outside their primary degrees. We chose a 4-point Lickert-scale assessment as it forces participants to not select neutral answers. We developed these questions based on literature (Long and Magerko 2020) and by consulting experts in visual design and AI. Experts were recruited among the faculty of visual design (4 experts) and technical (2 experts) colleges. Participants were then scored and profiled according to their Visual, AI and Prompt Engineering Literacies.

To measure their Visual Literacy, the survey asked about the participants' visual design experience (e.g., "How many hours per month do you spend drawing/designing or expressing yourself visually?") and visual design knowledge ("How familiar are you with visual composition concepts such as contrast, rhythm, balance, proportion, and harmony?"). For AI Literacy, the participants were asked about their AI literacy, such as "How much knowledge of Artificial Intelligence (AI), including Machine Learning and data science, do you have?" and "How often do you use AI and machine learning-based consumer applications, such as chatbots, image or text generators, etc.?" While we specifically target participants in study programs associated with Visual and AI literacy respectively, these survey questions can help us identify participants with both. Finally, for Prompt Engineering Literacy, participants were asked about their familiarity with Text-to-Image tools such as Dall-E, MidJourney, and Stable Diffusion, how frequently they utilized those tools and how much experience they have with prompt engineering.

Next, participants were given a design task to create an image of a Brutalist medieval castle of their own design. The task brief (Fig. 1) outlined thematic requirements details and provided some inspirations. This task is used by MOOD Visuals (2024. Mood Visuals), a commercial visual studio, as part of their recruitment interviews for new designers. To maximize ecological validity, we chose to employ a real-life visual task form an internationally acknowledged visual design bureau. After a brief introduction to the Midjourney tool, each participant was given 30 minutes to complete the task. They were instructed to use the think aloud protocol as they worked on the tasks. One researcher was present to answer questions and make observations. At the end, each participant chose what they considered to be the best image as their final design.

Lastly, the participant partook in a semi-structured interview about their overall experience and how they felt their technical or visual background affected their approach. They were also asked to rate their final image, on a 5-point Likert scale, about how close they felt the result was to what they initially had imagined. We refer to this evaluation parameter as Controllability, which is used in connection with the expert evaluation (see below) for the design outcome analysis.

3.2 Data Analysis

Our two research questions investigated the impact of the three forms of literacy on participants' prompt use and the creative process with TTI. We analysed the collected data through 1) lexical/linguistic analysis of the participants' prompts, and 2) qualitative process analysis based on participants' think aloud and interview data.

3.2.1 Lexical/linguistic Prompt Analysis.

To answer RQ1 we followed the Constructivist grounded theory approach: we began with an existing taxonomy derived from Sanchez's established prompt specifiers (Sanchez 2023) and iteratively refined it through constant comparison and analysis (see "Specifiers use" below). The data was initially divided by participants' educational institution: One researcher coded participants from the design college, while the other coded participants from the technical university. Throughout this initial coding phase, repeated check-ins and discussions took place to reach a common understanding. After the initial round of coding, both researchers reviewed all coded profiles together to ensure continuity and consistency across the data set. This process resulted in eight parameters: prompt iterations, unique tokens, unique non-stop tokens, prompt variance, prompt variety, adjective use, advanced use, specifiers use. The dataset was comprised of the user prompts from each participant. Automated word analysis was performed using the spaCy package (Explosion.ai 2024), while automatic semantic analysis was performed using Universal Sentence Encoder (Cer et al. 2018).

Figure 1. The design brief used in the user study.

Prompt iterations: the number of iterations users employ in generating pictures has been flagged by several researchers as important information when analysing prompting practices. For example, Almeda et al. identified three structures for exploring prompts: Iterative Prompt Exploration, Parametric Manipulation and Semantic Explorations (Almeda et al. 2024). Our choice to account for the number of iterations is based on the identified structure “iterative prompt exploration".

Unique tokens: the number of unique words (excluding punctuation) per user. O’Connor et al. suggests that long token contexts are important for efficient performance of current transformer language models (O’Connor and Andreas 2021), therefore we chose to analyse both unique and unique non-stop tokens.

Unique non-stop tokens: the number of unique words (excluding punctuation) less spaCy’s default list of “stop words”, i.e. smaller words like “a”, ”the”, and ”and”. Researchers have found that users with high prompt literacy are less likely to form full sentences and thus less likely to use such stop words.

Adjective use: the number of adjectives used in prompts, to identify prompts written in descriptive or aesthetically appreciative language. It is well established that adjectives (such as “blue", “slanted", or “delicious") play a large role in aesthetic appreciation (McNally and Stojanovic 2017). Goloujeh et al. identified five structures for writing prompts: their “overview and details” explains how the use of adjectives increases the efficiency of current language models (Mahdavi Goloujeh et al. 2024).

Figure 2. Examples of AI-generated designs by Participants 1, 3, 10, 13, 18, and 22, respectively.

Prompt variance: the semantic difference between each prompt, using the Universal Sentence Encoder’s similarity function (Cer et al. 2018). A user making only small changes between prompts will receive a high similarity score, and vice versa. Goloujeh et al. also identified seven strategies for refining prompts (Mahdavi Goloujeh et al. 2024), their “omitted words” and “reordering or rephrasing” strategies led us to include prompt variance as a factor in our analysis.

Advanced use: To account for advanced use of the TTI tool we observed two variables that documented correct and incorrect uses of the “Vary Region” function. This function was used by participants with the intention of changing particular elements of the previously generated image with a new prompt. Incorrect uses of the Vary Region function included prompts resulting in undesired generations or no applied changes to the image. Almeda et al. identified “Parametric Manipulation” as a useful structure for exploring prompts (Almeda et al. 2024). At the same time Goloujeh et al. point at “refining parameters” and “reroll and make variations” as recurrent strategies for refining prompts (Mahdavi Goloujeh et al. 2024). This led us to log advanced use of Midjourney functions such as “Vary Region".

Prompt variety: how many unique specifiers are contained in each prompt iteration. Sanchez set out to understand the semantic structure of prompts by establishing a prompt specifier taxonomy (Sanchez 2023). Almeda et al. with the “semantic exploration” strategy also hints at specifiers as tool for understanding the semantic structure of prompts (Almeda et al. 2024). Goloujeh et al. propose “descriptive sentences” as a structure to account for prompts’ semantic level (Mahdavi Goloujeh et al. 2024). We started from the specifier taxonomy proposed by (Sanchez 2023) and expanded it with a grounded theory analysis of all prompts of the 25 participants.

Specifiers use: To examine the semantic content of prompts, we utilised a Constructivist grounded theory approach (Tie et al. 2019). We started with an existing taxonomy of prompt specifiers established by Sanchez (Sanchez 2023). The initial categories from Sanchez’s taxonomy included: Subject, Medium (Subcategories: Photography and Cinema, Painting, Rendering, and Illustration), Influence (Subcategories: Artists, Art Genres and Movements, Artwork, and Art Repositories), Light, Colour, Composition, Detail, Context (Subcategories: Era, Weather, and Emotions).

Through a process of constant comparative analysis and continuous discussion during initial coding, we iteratively refined and expanded this framework. This led to the addition of new categories and subcategories to capture more specific nuances in the prompt descriptions. These new additions included: Historic Reference (e.g., "King Arthur", as a subcategory of Subject) Geographic Region (e.g., "Northern Ireland", as a subcategory of Subject) Details of the Subject (e.g., "no towers"), Texture (e.g., "rough"), Shape Language (e.g., "round") Finally, the developed coding framework was further refined by comparing and aligning these categories to Xie et al.’s (Xie et al. 2023) broader categorization of TTI prompts into Subject, Form, and Intent.

3.2.2 Qualitative Process Analysis.

To answer RQ2 we used Grounded Theory to conduct process analysis based on the interview data and think-aloud data. We aimed to explore the user design processes, and how they might differ between individuals with different types of literacy. The process analysis methodology is based on Grounded Theory as presented by Glaser and Strauss (Glaser and Strauss 1967) and Adele E. Clarke’s situational analysis methodology, a further reworking of Grounded Theory (Clarke 2003). Clarke’s situational maps offered a scaffold to map out the data from initial coding and identify categories and properties within the data. After initial coding, the user study and interview data were combed through for substance, and six participants’ user studies and interviews were chosen, three from each school, as best possible participant representatives in the further analysis and as support for the subsequent presentation of results. Based on these six interviews, and with inspiration from the process of Clarke's situational maps, a mapping of interesting points in the data was created, to visualize better the relevant data for the results of the process analysis (Fig. 3).

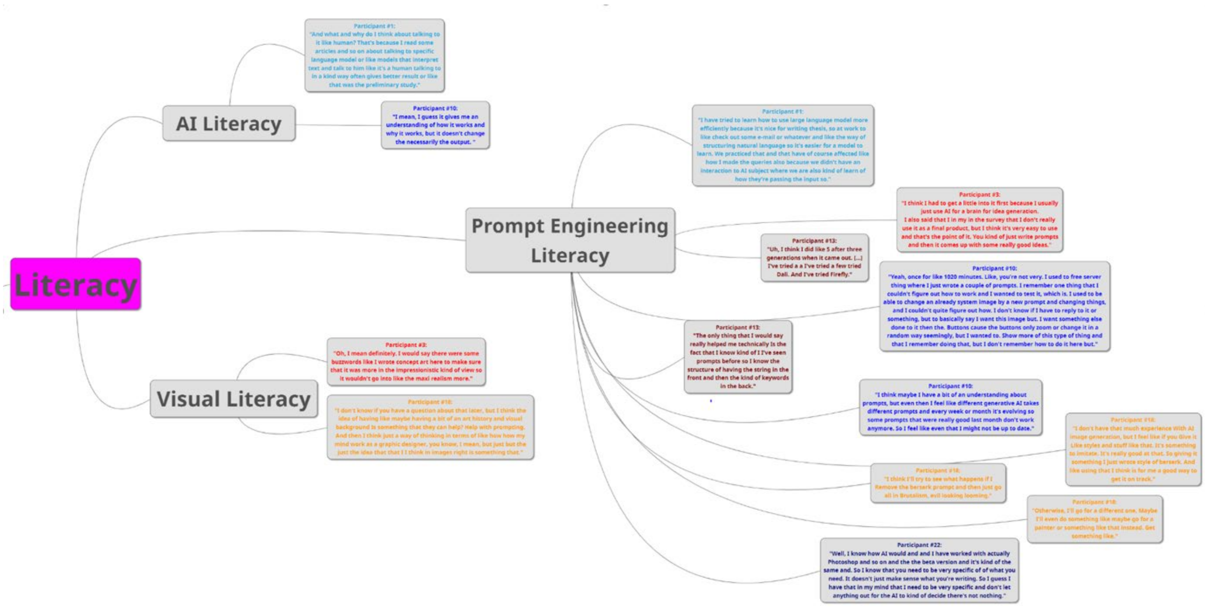

Figure 3. Snippet of Mindmap for Process Analysis, based on the process of Situational Maps in Clarke (2011).

4 Results

25 participants (11 females and 14 males) completely the study (Fig. 2). Among them, 14 were from a design college and 11 from a technical university. Their Visual, AI and Prompt Engineering literacies were not assumed because of their educational background but were based on their scores on the surveys they completed. They successfully submitted 25 final designs. Their age range is between 22 and 37 (mean=25.32, standard deviation=3.10).

4.1 RQ1: The Impact of Visual Literacy, AI Literacy and Prompt Engineering Literacy on prompt use

<u>Hypothesis 1: Participants with visual literacy will use a larger vocabulary and will have finer-grained control of the domain.</u>

To account for larger vocabulary and fine-grained control, Spearman's nonparametric correlation was computed to assess the relationship between these pairs: Visual Literacy and Unique Tokens, Visual Literacy and Unique non-stop Tokens, Visual Literacy and Adjective Use, Visual Literacy and Prompt Variety, Visual Literacy and Prompt Variance.

- There is a moderate positive correlation between Visual Literacy and Unique non-stop Tokens (ρ = .42, p = .033) and it is statistically significant.
- There is a strong positive correlation between Visual Literacy and Adjective Use (ρ = .61, p = .001), which is statistically significant.
- There's a moderate positive correlation between Visual Literacy and Prompt Variety (ρ = .49, p = .012), which is statistically significant.

This demonstrates that participants with higher visual literacy use a larger vocabulary, especially regarding the use of unique non-stop tokens. There is strong evidence that these participants tend to use more adjectives and have greater prompt variety, suggesting more detailed and varied descriptions. These findings suggest that while visual literacy may not be strongly related to the overall number of unique tokens, it is associated with a larger vocabulary and more nuanced and varied use of language in prompts, particularly in terms of adjective use and the variety of specifiers employed.

<u>Hypothesis 2: Participants with visual literacy will make more references to artists and art movements as well as to style in general.</u>

In order to account for style, genres, artists and art movements, Spearman's nonparametric correlation was computed to assess the relationship between Visual Literacy and the specifiers: "Influence", "Medium", "Detail", "Composition", "Colour" and "Light".

- There's a moderate positive correlation between Visual Literacy and Influence (ρ = .44, p = 0.027), which is statistically significant.
- There's a moderate positive correlation between Visual Literacy and Medium (ρ = .50, p = .010), which is statistically significant.
- The correlations between Visual Literacy and Detail (ρ = .30, p = .138), Composition (ρ = .13, p = .535), Colour (ρ = .26, p = .192), and Light (ρ = .34, p = .095) are weak to moderate and not statistically significant.

The hypothesis that participants with higher visual literacy make more references to art genres, styles, and movements is supported by the data. However, there's no strong evidence that they make more references to specific artists. Participants with higher visual literacy tend to make more references to medium, suggesting a focus on the technical aspects of art. The relationship between visual literacy and specific style elements (detail, composition, color, light) is not statistically significant, though there are weak to moderate positive correlations. These findings suggest that while visual literacy is associated with more references to art genres, styles, movements, and media, it doesn't necessarily correlate with more references to specific artists or detailed style elements. The users' focus seems to be more on broader art categories and technical aspects rather than individual artists or specific stylistic details.

<u>Hypothesis 3: Larger vocabulary, use of expert language, and use of artistic references can predict Visual Appeal.</u>

We established that participants with visual literacy make use of expert language (larger vocabulary and finer-grained control of the domain). We aim to assess the impact of visual literacy and expert language (unique tokens, adjective use, prompt variance, prompt variety, and "influence" specifiers used) on visual appeal leveraging Multiple Regression Analysis (least Squares Fit) (Rencher and Christensen 2012).

Model Fit: The R-squared (RSq) value is 0.39, which means that only about 39% of the variance in the Visual Appeal score is explained by the predictors in the model. The overall model p-value is .569, which is not statistically significant at the conventional .05 level.

Individual Predictors: None of the predictors, including Visual Literacy, show statistically significant effects. The p-values for all predictors are above the 0.05 threshold: Unique Tokens ($p = .050$) is closest to significance but still above the threshold. Unique non-stop Tokens ($p = .062$) is also close to significance but not statistically significant. Other predictors like Prompt Variety ($p = .129$), Influence ($p = .219$), Adjectives use ($p = .468$), and Visual Literacy ($p = .604$) are far from significance.

Visual Literacy Impact: Visual Literacy is not a significant predictor in this model, with a p-value of .604 and a low logworth value of .219, indicating that it does not strongly predict Visual Appeal score.

4.2 RQ2: Computational Creativity Process Analysis

This section presents the results of our process analysis, based on six user studies and interviews, representative of the two educational institutions and with different types of literacy. Through this analysis, we aim to uncover key insights into how user competencies shape human-AI collaboration in creative contexts.

Impact of Different Types of Literacy on AI-assisted Design. The study revealed distinct impacts of different literacy on the computational creativity process. Participants with Visual Literacy (VL) demonstrated a sophisticated design vocabulary, using specific visual concepts and styles to guide the AI, often creatively expanding on the design brief. However, these users frequently expressed frustration with Midjourney's limitations in executing their precise visual ideas and their perceived lack of granular control over the output. Conversely, individuals with AI Literacy (AIL) exhibited a deeper understanding of AI interaction, influencing their prompting strategies and expectations, but this understanding did not always clearly translate to the final output's quality. Prompt Engineering Literacy (PEL) emerged as a crucial skill that developed dynamically during the task, even in those with little initial AI experience, manifesting in the ability to iteratively refine prompts, adapt to AI responses, and learn from the feedback loop. While AI-literate users often showed faster improvement in PEL, visually literate users sometimes struggled to translate their visual ideas into effective prompts.

Task Approach and Ideation Process. Task Approach and Ideation Process Participants' approaches to the design brief and their ideation strategies varied significantly based on their literacy backgrounds. Those with high AI Literacy tended to adhere closely to the brief's exact wording, treating it as a set of precise specifications. In contrast, visually literate participants often creatively expanded upon the brief, adding their own interpretations and additional elements. Regarding ideation, participants from technical universities, often less confident in visual design, relied more heavily on Midjourney for creative ideas, sometimes finding the AI provided concepts they hadn't initially considered. Conversely, design college participants typically approached the task with clearer initial ideas but sometimes struggled to translate these pre-conceived notions into effective prompts for the AI. Awareness of the AI's capabilities also differed: AI-literate users were more attuned to potential AI-generated errors, actively looking for them, while visually literate users focused on the tool's limitations in realizing their specific visual visions, leading to frustration when their exact ideas couldn't be matched.

User Perceptions and Comparisons. User Perceptions and Comparisons Participants frequently compared their experience with Midjourney to other AI tools and traditional design processes, providing insights into its strengths and limitations. AI-literate participants often highlighted features they felt were missing compared to other generative AI tools, expressing a desire for more granular control. Visually literate participants, on the other hand, wished for features that would allow them to leverage their existing skills, such as providing a sketch to start from, indicating a desire for more control over the creative process beyond pure text-based interactions. When comparing to traditional processes, participants with less visual design experience often favored Midjourney for its ability to generate tailored results quickly, even if it meant waiting for outputs, contrasting with the often less relevant results from traditional image searches. Visually literate participants viewed Midjourney as a powerful ideation tool but saw it as a starting point rather than a complete solution, intending to add further details themselves. Across all groups, the speed and efficiency of Midjourney were consistently appreciated as a fast way of visualization and ideation. These varied perceptions suggest that optimal use of TTI tools may necessitate a blend of visual, AI, and prompt engineering literacy.

5 Discussion

Our previous study (Canossa et al. 2025) indicated that only Visual Literacy significantly impacted Task Fulfilment, while it surprisingly showed no effect on Visual Appeal. Moreover, neither AI Literacy nor Prompt Engineering Literacy correlated with participants' perceived control over the computational creativity process. The Visual Appeal of generated images appeared unaffected by any of the literacy types or prompt features, likely because TTI tools are trained on vast datasets that already incorporate aesthetically pleasing elements.

5.1 Insights from Prompt Use Analysis

Prompt analysis provided several key insights into how participants' literacy types influenced their prompt usage patterns, linguistic and semantic composition, structure, and morphology. Notably, participants with higher Visual Literacy demonstrated a larger vocabulary, especially using more unique non-stop tokens ($p = .42$, $p = .033$) and significantly more adjectives ($p = .61$, $p = .001$) in their prompts. They also exhibited greater prompt variety ($p = .49$, $p = .012$), indicating a more detailed and nuanced use of language. This suggests that Visual Literacy leads to richer and more varied descriptions when crafting prompts for TTI tools. Furthermore, prompt analysis showed that visually literate participants tended to make more references to art genres, styles, and movements (Influence: $p = .44$, $p = 0.027$) and to the medium ($p = .50$, $p = .010$). This indicates their focus on broader art categories and technical aspects in their prompting language, rather than solely on specific artists or granular style elements. While it was hypothesized that a larger vocabulary, use of expert language, and artistic references could predict how aesthetically pleasing the generated images were, the prompt analysis, leveraging multiple regression, found no statistically significant relationship between these prompt features and Visual Appeal.

5.2 Insights from Process Analysis

The qualitative data from our process analysis largely supports and provides context for our quantitative findings. It reinforces the observation that participants with higher visual literacy were better at fulfilling task requirements, as evidenced by their more sophisticated design vocabulary and creative interpretation of the brief. The qualitative data also aligns with our finding that AI literacy doesn't significantly impact visual appeal or task fulfillment, but it offers additional insights into how participants with different backgrounds approached the task and interpreted AI outputs. Specifically, AI literacy provides a deeper understanding of the computational creativity process, enabling users to embrace a co-creation role with TTI tools. Users with lower scores in AI literacy expect a higher degree of control on the computational creativity process and are somewhat unsatisfied with the results. Although AI literacy did not impact the quality of the images produced, it did impact the users' understanding of the process, making them more willing to accept a co-creation position with TTI tools. This highlights the importance of AI literacy for HCI researchers and practitioners, who should ensure that the tools they develop are easy to understand and use, regardless of the user's level of AI knowledge.

While our previous quantitative analysis showed no significant impact of prompt engineering literacy on VA, TF or C, the qualitative data suggests that this skill developed rapidly during the task for many participants, potentially explaining its lack of measurable impact. This dynamic skill development presents an area for future investigation. The study indicates that prompt engineering literacy is a skill that can be quickly developed during interaction with TTI tools. HCI practitioners should consider designing interfaces and tutorials that support users in developing this skill, thereby enhancing the quality of their creative outputs. Overall, the qualitative insights provide valuable context that enriches our understanding of the quantitative results, offering a more nuanced picture of the human-AI computational creativity process in visual design tasks.

5.3 Implications for Education

Our prompt analysis also yielded important information for educational programs. It revealed that participants with higher visual literacy tended to use a larger vocabulary of unique non-stop tokens and significantly more adjectives, suggesting a more detailed and nuanced use of language in their prompts. They also demonstrated greater prompt variety and made more references to art genres, styles, and movements, as well as to the medium. While prompt analysis did not find a significant correlation between these linguistic features and general visual appeal, it provides crucial insights for design education on how expert visual language is structured and used to articulate creative intent for TTI tools. Understanding these prompt usage patterns, linguistic and semantic composition, and structure can inform curriculum development. For technical education, prompt analysis underscores the importance of developing AI models that can effectively interpret and leverage such rich, domain-specific vocabularies and references. By studying how visually literate users naturally construct their prompts, technical programs can better equip students to design TTI systems that are more responsive to creative input and understand nuanced textual descriptions.

Process analysis, through examining participants' think-aloud data and interviews, revealed significant insights into how different types of literacy influence the AI-assisted design process, impacting educational recommendations for both design and technical fields. For design education, the analysis showed that participants with high visual literacy often held sophisticated visual concepts but struggled to translate these ideas into effective textual prompts, indicating a need to teach methods for "textualizing" visual concepts and developing a rich expressive visual vocabulary. Process analysis also highlighted that visually literate users expected a higher degree of control over the output and were sometimes unsatisfied, feeling constrained by the tools. This suggests design programs should set realistic expectations for AI control and explore hybrid workflows that integrate AI outputs with traditional design tools for detailed work. For technical education, process analysis indicated that participants with high AI literacy had a deeper understanding of the computational creativity process and were more willing to accept a co-creation role with TTI tools. This implies technical programs should promote understanding of AI's interpretative mechanisms to help future developers build more robust and user-centric TTI tools. Both fields can benefit from fostering a co-creation mindset in AI development. Process analysis also noted that Prompt Engineering Literacy emerged as a crucial skill that developed dynamically through interaction, even for those with little initial AI experience. This means educational programs can integrate hands-on iterative prompting exercises and design tools that facilitate rapid development of this literacy.

6 Limitations and Future Work

The main limitation of the current work is its sample size: due to the exploratory nature of the study and the array of methods deployed to analyze the data including resource-intensive qualitative research such as process analysis, it was not feasible to collect and process data from more than a few dozen participants.

The use of Spearman's correlation and regression analysis provided valuable insights into the relationships between visual literacy, language use, and evaluation measures. However, these methods also have limitations, particularly when dealing with non-linear relationships or potential confounding variables not included in the study. Future research could benefit from employing more advanced statistical techniques, such as path analysis or structural equation modeling, which allow for a more nuanced understanding of the relationships between multiple variables.

7 Conclusion

This paper extended existing work on Text-to-Image (TTI) tools by focusing specifically on how Visual, AI, and Prompt Engineering Literacy influence user strategies and prompt composition during co-creation with AI. Through a multi-faceted approach involving participant profiling, a visual design task, prompt analysis, and qualitative interviews, the research demonstrated that visual literacy significantly impacts prompt usage patterns, leading to a larger and more nuanced vocabulary, including more unique non-stop tokens and adjectives, as well as greater prompt variety and references to art genres and media. While visual literacy influenced prompt richness, no statistically significant relationship was found between prompt features and the aesthetic appeal of the generated images. The study also revealed that visually literate users often experienced frustration with the perceived lack of granular control over the TTI tool, whereas AI-literate users demonstrated a deeper understanding of the AI's capabilities and were more accepting of a co-creation role. Prompt Engineering Literacy emerged as a crucial skill that developed dynamically through interaction. Ultimately, the findings highlight that the optimal use of TTI tools likely requires a blend of visual, AI, and prompt engineering literacy, offering key implications for educational programs in both design and technical fields to enhance human-AI collaboration.

References

2024. DiffusionDB. https://github.com/poloclub/diffusiondb
2024. Dall-E. https://dalle3.ai/
2024a. Midjourney. https://www.midjourney.com/home
2024b. Midjourney Discord Dataset.
https://bridges.monash.edu/articles/dataset/Midjourney_2023_Dataset/25038404?file=44176019
2024. Mood Visuals. https://www.moodvisuals.com/
2024. UNESCO Literacy definition. https://uis.unesco.org/node/3079547

Shm Garanganao Almeda, J.D. Zamfirescu-Pereira, Kyu Won Kim, Pradeep Mani Rathnam, and Bjoern Hartmann. 2024. Prompting for Discovery: Flexible Sense-Making for AI Art-Making with Dreamsheets. 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 160, 17 pages. https://doi.org/10.1145/3613904.3642858

Peter Ardhianto and Ryan S. Nababan. 2023. Artificial Intelligence Approach in Visual Design Ideation Process. 112–117.

Maria Avgerinou and John Ericson. 1997. A Review of the Concept of Visual Literacy. British Journal of Educational Technology 28, 4 (1997), 280–291.

Alessandro Canossa, Lisa Toender, Lucas Fellner, Jesper Juul, Willem Van der Maden, Jichen Zhu. 2025. Algorithmic Creativity: How Visual and AI Literacy Impact the Use of Text-to-Image Tools in Design Tasks. International Conference on Human-Computer Interaction. Springer Nature. Switzerland

Daniel Cer, Yinfei Yang, Sheng yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St. John, Noah Constant, Mario Guajardo-Cespedes, Steve Yuan, Chris Tar, Yun-Hsuan Sung, Brian Strope, and Ray Kurzweil. 2018. Universal Sentence Encoder. https://arxiv.org/abs/1803.11175

Adele E. Clarke. 2003. Situational Analyses: Grounded Theory Mapping After the Postmodern Turn. Symbolic Interaction 26, 4 (2003), 553–576.

Stefania Druga, Nancy Otero, and Amy J. Ko. 2022. The Landscape of Teaching Resources for AI Education. In Proceedings of the 27th ACM Conference on Innovation and Technology in Computer Science Education Vol. 1 (Dublin, Ireland) (ITiCSE ’22). Association for Computing Machinery, New York, NY, 96-102.

Stefan Feuerriegel, Jochen Hartmann, Christian Janiesch, and Patrick Zschech. 2023. Generative AI. SSRN Electronic Journal (01 2023).

Helen Gardner. 1970. *Art through the Ages*. Harcourt Brace Jovanovich.

B.G. Glaser and A.L. Strauss. 1967. The Discovery of Grounded Theory: Strategies for Qualitative Research. Aldine.

Ariel Han, Xiaofei Zhou, Zhenyao Cai, Shenshen Han, Richard Ko, Seth Corrigan, and Kylie A Peppler. 2024. Teachers, Parents, and Students’ perspectives on Integrating Generative AI into Elementary Literacy Education. 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 678, 17 pages. https://doi.org/10.1145/3613904.3642438

Nervana Osama Hanafy. 2023. Artificial intelligence’s effects on design process creativity: "A study on used A.I. Text-to-Image in architecture". Journal of Building Engineering 80 (2023), 107999. https://doi.org/10.1016/j.jobe.2023.107999

Ellen Jiang, Kristen Olson, Edwin Toh, Alejandra Molina, Aaron Donsbach, Michael Terry, and Carrie J Cai. 2022a. PromptMaker: Prompt-based Prototyping with Large Language Models. 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 35, 8 pages. https://doi.org/10.1145/3491101.3503564

Ellen Jiang, Edwin Toh, Alejandra Molina, Kristen Olson, Claire Kayacik, Aaron Donsbach, Carrie J Cai, and Michael Terry. 2022b. Discovering the Syntax and Strategies of Natural Language Programming with Generative Language Models. 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 386, 19 pages. https://doi.org/10.1145/3491102.3501870

Nils Knoth, Antonia Tolzin, Andreas Janson, and Jan Marco Leimeister. 2024. AI literacy and its implications for prompt engineering strategies. Computers and Education: Artificial Intelligence 6 (2024), 100225. https://doi.org/10.1016/j.caeai.2024.100225

Pei-Ying Lin, Kristina Andersen, Ralf Schmidt, Sanne Schoenmakers, Hèrm Hofmeyer, Pieter Pauwels, and Wijnand Ijsselsteijn. 2024. Text-to-Image AI as a Catalyst for Semantic Convergence in Creative Collaborations. In Proceedings of the 2024 ACM Designing Interactive Systems Conference (Copenhagen, Denmark) (DIS ’24). Association for Computing Machinery, New York, NY, USA, 2753–2767. https://doi.org/10.1145/3643834.3661543

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
Duri Long and Brian Magerko. 2020. What is AI Literacy? Competencies and Design Considerations. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI ’20). Association for Computing Machinery, New York, NY, USA, 1–16. https://doi.org/10.1145/3313831.3376727
Yao Lu, Max Bartolo, Alastair Moore, Sebastian Riedel, and Pontus Stenetorp. 2022. Fantastically Ordered Prompts and Where to Find Them: Overcoming Few-Shot Prompt Order Sensitivity. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Smaranda Muresan, Preslav Nakov, and Aline Villavicencio (Eds.). Association for Computational Linguistics, Dublin, Ireland, 8086–8098. https://doi.org/10.18653/v1/2022.acl-long.556
Atefeh Mahdavi Goloujeh, Anne Sullivan, and Brian Magerko. 2024. Is It AI or Is It Me? Understanding Users’ Prompt Journey with Text-to-Image Generative AI Tools. 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 183, 13 pages. https://doi.org/10.1145/3613904.3642861
Robert Maloy and Sai Gattupalli. 2024. Prompt Literacy. EdTechnica (2024). https://doi.org/10.59668/371.14442
Louise McNally and Isidora Stojanovic. 2017. Aesthetic adjectives. Oxford University Press, Oxford, UK, 17–37.
Sofia Morandini, Federico Fraboni, Marco De Angelis, Gabriele Puzzo, Davide Giusino, Luca Pietrantoni, et al. 2023. The impact of artificial intelligence on workers’ skills: Upskilling and reskilling in organisations. Informing Science 26 (2023), 39–68.
Davy Tsz Kit Ng, Jac Ka Lok Leung, Samuel Kai Wah Chu, and Maggie Shen Qiao. 2021. Conceptualizing AI literacy: An exploratory review. Computers and Education: Artificial Intelligence 2 (2021), 100041. https://doi.org/10.1016/j.caeai.2021.100041
Joe O’Connor and Jacob Andreas. 2021. What Context Features Can Transformer Language Models Use? https://arxiv.org/abs/2106.08367
Jonas Oppenlaender. 2022. The Creativity of Text-to-Image Generation. In Proceedings of the 25th International Academic Mindtrek Conference (Tampere, Finland) (Academic Mindtrek ’22). Association for Computing Machinery, New York, NY, USA, 192–202. https://doi.org/10.1145/3569219.3569352
A.C. Rencher and W.F. Christensen. 2012. Methods of Multivariate Analysis. Wiley.
Laria Reynolds and Kyle McDonell. 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 (Yokohama, Japan) (CHI EA ’21). Association for Computing Machinery, New York, NY, USA, Article 314, 7 pages.
John Ruskin. 1857. The Elements of Drawing. Smith, Elder & Co.
Téo Sanchez. 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 (Virtual Event, USA) (C&C ’23). Association for Computing Machinery, New York, NY, USA, 43–61.
Sivertsen, C., Salimbeni, G., Løvlie, A.S., Benford, S.D. and Zhu, J., 2024, May. Machine learning processes as sources of ambiguity: Insights from ai art. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (pp. 1-14).
Karin Stolpe and Jonas Hallström. 2024. Artificial intelligence literacy for technology education. Computers and Education Open 6 (2024), 100159.
Ylona Chun Tie, Melanie Birks, and Karen Francis. 2019. Grounded theory research: A design framework for novice researchers. SAGE Open Medicine 7 (2019), 2050312118822927.
Tommaso Turchi, Silvio Carta, Luciano Ambrosini, and Alessio Malizia. 2023. Human-AI Co-Creation: Evaluating the Impact of Large-scale Text-to-Image Generative Models on the Creative Process. https://doi.org/10.1007/978-3-031-34433-6_3
Willem Van Der Maden, Evert Van Beek, Brett A Halperin, Petra Jääskeläinen, Eunsu Kang, Peter Kun, James Derek Lomas, Timothy Merritt, Joseph Lindley, Michael Muller, et al. 2024. Death of the Design Researcher? Creating Knowledge Resources for Designers Using Generative AI. In Companion, Publication of the 2024 ACM Designing Interactive Systems Conference. 396–400.
Willem Van Der Maden, Evert Van Beek, Iohanna Nicenboim, Vera Van Der Burg, Peter Kun, James Derek Lomas, and Eunsu Kang. 2023. Towards a Design (Research) Framework with Generative AI. In Companion Publication of the 2023 ACM Designing Interactive Systems Conference. 107–109.

Samangi Wadinambiarachchi, Ryan M Kelly, Saumya Pareek, Qiushi Zhou, and Eduardo Velloso. 2024. The Effects of Generative AI on Design Fixation and Divergent Thinking. In Proceedings of the CHI Conference on Human Factors in Computing Systems. 1–18.

Zhuohao Wu, Danwen Ji, Kaiwen Yu, Xianxu Zeng, Dingming Wu, and Mohammad Shidujaman. 2021. AI Creativity and the Human-AI Co-creation Model. In Human-Computer Interaction. Theory, Methods and Tools: Thematic Area, HCI 2021, Held as Part of the 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part I. Springer-Verlag, Berlin, Heidelberg, 171–190. https://doi.org/10.1007/978-3-030-78462-1_13

Yutong Xie, Zhaoying Pan, Jinge Ma, Luo Jie, and Qiaozhu Mei. 2023. A Prompt Log Analysis of Text-to-Image Generation Systems. In Proceedings of the ACM Web Conference 2023 (WWW '23). ACM, 3892–3902. https://doi.org/10.1145/3543507.3587430

Zhou Hongwei, Jichen Zhu, Michael Mateas, and Noah Wardrip-Fruin. "The Eyes, the Hands and the Brain: What can Text-to-Image Models Offer for Game Design and Visual Creativity?" In Proceedings of the 19th International Conference on the Foundations of Digital Games, pp. 1-13. 2024

Daijin Yang, Yanpeng Zhou, Zhiyuan Zhang, Toby Jia-Jun Li, and LC Ray. 2022. AI as an Active Writer: Interaction Strategies with Generated Text in Human-AI Collaborative Fiction Writing 56-65. In IUI Workshops. https://api.semanticscholar.org/CorpusID:248301902

J.D. Zamfirescu-Pereira, Shm Garanganao Almeda, Kyu Won Kim, and Bjoern Hartmann. 2023. Towards Image Design Space Exploration in Spreadsheets with LLM Formulae. In Adjunct Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (San Francisco, CA, USA) (UIST '23 Adjunct). Association for Computing Machinery, New York, NY, USA, Article 82, 3 pages. https://doi.org/10.1145/3586182.3615790

Jichen Zhu, Antonios Liapis, Sebastian Risi, Rafael Bidarra, and G Michael Youngblood. 2018. Explainable AI for designers: A human-centered perspective on mixed-initiative co-creation. In 2018 IEEE conference on computational intelligence and games (CIG). IEEE, 1–8.