

The Effects of Generative AI on Design Fixation and Divergent Thinking

Samangi Wadinambiarachchi
The University of Melbourne
Melbourne, Australia
samangi.w@unimelb.edu.au

Ryan M. Kelly
RMIT University
Melbourne, Australia
ryan.kelly@rmit.edu.au

Saumya Pareek
The University of Melbourne
Melbourne, Australia
spareek@student.unimelb.edu.au

Qiushi Zhou
The University of Melbourne
Melbourne, Australia
qiushi.zhou@unimelb.edu.au

Eduardo Velloso
The University of Melbourne
Melbourne, Australia
eduardo.velloso@unimelb.edu.au

ABSTRACT

Generative AI systems have been heralded as tools for augmenting human creativity and inspiring divergent thinking, though with little empirical evidence for these claims. This paper explores the effects of exposure to AI-generated images on measures of design fixation and divergent thinking in a visual ideation task. Through a between-participants experiment (N=60), we found that support from an AI image generator during ideation leads to higher fixation on an initial example. Participants who used AI produced fewer ideas, with less variety and lower originality compared to a baseline. Our qualitative analysis suggests that the effectiveness of co-ideation with AI rests on participants' chosen approach to prompt creation and on the strategies used by participants to generate ideas in response to the AI's suggestions. We discuss opportunities for designing generative AI systems for ideation support and incorporating these AI tools into ideation workflows.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**

KEYWORDS

Design fixation, Generative-AI, Creativity support tools

ACM Reference Format:

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 (CHI '24)*, May 11–16, 2024, Honolulu, HI, USA. ACM, New York, NY, USA, 18 pages. <https://doi.org/10.1145/3613904.3642919>

1 INTRODUCTION

Consider a team of designers discussing ideas for environmentally friendly transport solutions for a city. One team member kicks off the discussion with a suggestion about electric buses. The rest of the

team then spends an hour discussing variations on this idea, all involving electric vehicles, until an intern who arrived late asks “have you considered bicycles?”. Until the intern’s suggestion, the ideas were anchored on a salient characteristic of the first proposal—an electric motor. The design literature dubs this phenomenon *design fixation*—the “blind adherence to a set of ideas or concepts limiting the output of conceptual design” [30, p. 1]. This is a common experience in any creative task, from art to engineering, and happens when exposure to one idea anchors and biases subsequent ideas, restricting exploration of the design space. Fixation happens both consciously and unconsciously, regardless of the level of experience of the practitioner [30, 76] and in all areas of creative work. The severe negative impact that design fixation has on the creative process makes it a key concern in design studies.

In the initial stages of the design process, it is common for designers to conduct precedence studies and create mood boards by compiling external stimuli as sources of inspiration to broaden their ideation space [41]. However, the exposure to previous solutions during this process can potentially be a source of design fixation. Previous studies have shown that exposure to examples of similar design solutions has mixed effects on creativity [75]. It tends to drive designers towards the example, narrowing the explored solution space [30, 38]. Further, variations in the modality [64, 68], the fidelity [13, 64], the quality [64], the diversity and novelty of the exposed stimuli, the time of exposure, and its proximity to the design problem [64] can vary the intensity of design fixation [63].

Recent developments in generative artificial intelligence (GenAI) have been heralded as the harbinger of a new paradigm of creative work, often under the guise of augmenting human creativity [22]. Publicly available AI image generators such as DALL·E¹, Artbreeder², Stable Diffusion³, and Midjourney⁴ have made it possible for designers to turn their thoughts into high-quality visuals quickly and at a low cost. The ability of these tools to generate “original” images based on user prompts potentially offers a rich source of inspiration. For example, Chiou et al. [14] have shown that when used in co-ideation tasks, AI can open up a broader conceptual space quickly and effortlessly, promoting divergent thinking [14].

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

CHI '24, May 11–16, 2024, Honolulu, HI, USA

© 2024 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0330-0/24/05.

<https://doi.org/10.1145/3613904.3642919>

¹<https://openai.com/dall-e-2>

²<https://www.artbreeder.com>

³<https://stablediffusionweb.com>

⁴<https://www.midjourney.com>

However, **there is still a lack of empirical evidence for the effect of generative AI as a source of inspiration during design tasks.** Though the specific outputs generated by these tools are novel, they are trained on existing work, blurring the lines between what is original and derivative. Further, designers could still be fixated during the ideation process despite any potential inspiration from AI.

In this paper, we aim to understand the effects of AI-generated imagery as a source of inspiration in an ideation task. We conducted a between-participants experiment in which designers took part in a visual ideation task that involved sketching ideas for a chatbot avatar. We manipulated participants' access to sources of inspiration: none, access to Google Image Search, or access to Midjourney (an image generation AI tool). Through our study, we sought answers to the following questions:

- RQ 1: How does the exposure to AI-generated images affect design fixation and divergent thinking during ideation, compared to using commonly used sources of inspiration and no inspiration support?
- RQ 2: How do different ways of interacting with AI image generators impact participants' effectiveness in an ideation task?

We evaluated the effect of inspiration sources on participants' ideation output (the sketches). In doing so, we used four divergent thinking measures from prior literature (design fixation score, fluency, variety, and originality [30, 53, 62, 76]) to assess different facets of their creative output. We found that exposure to AI-generated images induced higher design fixation in participants than in other conditions. Moreover, fluency, variety, and originality were lower in the AI-supported group compared to the baseline condition. Through our qualitative analysis, we suggest that fixation arises when creating prompts and when ideating in response to AI images. In addition, we demonstrate that using AI can result in *fixation displacement*, where the focus of fixation shifts from an exemplar onto the AI's outputs.

Our study provides an empirical contribution to the AI-powered creativity support literature by illustrating how AI-generated images influence design fixation and divergent thinking measures. It further elaborates on AI's role in providing inspiration during visual design tasks. Further, we demonstrate the importance of focusing on factors that might induce design fixation while acquiring inspiration from AI tools and propose potential strategies and directions to explore in mitigating design fixation.

2 RELATED WORK

Our research builds on studies of design fixation and on the role AI can play in supporting design ideation.

2.1 Design Fixation

Among the factors that hinder designers' creativity, "*design fixation*" is one of the most well-studied phenomena in creativity and design research. It is identified as the unconscious adherence to a set of pre-known ideas or knowledge that restricts the ideation space [30, 76]. When a person experiences design fixation, they tend to adhere to pre-conceived ideas and concepts, limiting exploration of the design space during ideation [30, 40, 47]. Design fixation narrows

designers' ability to explore the creative space between abstract ideas and potential solutions [30, 64]. Previous findings show that this is reflected heavily in their design outcomes and restrains designers from maximising their creative potential, resulting in unoriginal outputs [30].

Design fixation has been studied extensively across different fields [61], including cognitive science [9], design [5, 34], education [28], mechanical engineering [67, 74], and psychology [4, 57]. These studies have collectively shown that design fixation is more likely to occur when designers are exposed to example solutions for design tasks [30]. It has also been demonstrated that the modality, degree of abstraction of the inspiration (i.e. the fidelity), and the designer's level of expertise [1] can affect fixation intensity when exposed to external stimuli.

Fixation has typically been studied through quantitative experimental approaches in which participants are asked to solve a design problem, either with or without an example (external stimuli) [64]. For instance, Jansson and Smith's classic design fixation work [30] reported four experiments. These experiments divided participants into two groups: a treatment group (fixation group), who were given a design problem along with an example solution, and a control group, who were given the same problem with no examples to work from. They hypothesised that showing an example design would restrict the ideas of the treatment group because it would make the participants fixate on the given example. Jansson and Smith [30] found that even though both groups produced a similar number of designs, ideas in the fixation groups were more similar to the example. In a subsequent experiment, the researchers found that the flexibility and originality of the designs were limited in the fixation group and concluded that creative performance may be inhibited when an example induces design fixation. Since then, several studies have been conducted replicating or amending the method and examples [38, 64].

When looking at design fixation, it is important to distinguish different types of fixation effects [18]. Youmans and Arciszewski [76] identify three such effects. The first is *unconscious adherence* [76] to past designs without realizing. An example of this is copying the features of an example (even if the features are inappropriate to the task) [18, 30]. The second is *conscious blocking* [76], where new ideas are actively but perhaps momentarily dismissed. In this situation, a designer is aware of alternative creative paths but chooses to disregard them, perhaps due to a commitment to a current project's direction or a bias towards familiar solutions. The third type of fixation effect is *intentional resistance*, a deliberate decision against exploring new concepts. For instance, design companies engaged in research and development often prefer to explore solutions that fall within their well-established expertise, a tendency known as *local search bias* [18, 50].

Apart from trying to understand its causes, researchers have explored various strategies to overcome design fixation [64]. Such strategies include incorporating physical prototyping in the ideation activities [67], triggering frequent reminders for participants to consider all available options in a timely manner during an ideation task [42, 76], utilising design thinking and lateral thinking methods [6] such as de Bono's six thinking hats [2, 20], having short breaks or "incubation periods" during the task [58, 73], using computer-aided

design and intelligent agents [19, 27], and incorporating design by analogy [12].

Even though there is a large body of work on design fixation exploring the effects of external stimuli on creative tasks [1], studies centred around design fixation are limited within the field of HCI. Among these few studies, HCI researchers have started to examine the potential of using AI image generators as tools for supporting creativity [27, 39]. Thus, in this study, we adapt experimental methods from mechanical engineering and design research, where design fixation is framed as *unconscious adherence* and is measured by the degree to which participants directly copy features from an example stimulus. We aim to understand the influence of AI image generators on design fixation and divergent thinking, adding new empirical evidence to the HCI literature.

2.2 The Emergent Role of AI in Creativity Support

Since the early 1990s, designers have envisioned a future with intelligent design and creative aids [24]. With recent advances in Generative AI, this vision is becoming a reality. Generative AI systems can create new, plausible media [49] to aid individuals in creative tasks [31]. Generative AI models are trained on large data sets and can enable people to generate content such as images, text, audio, or video quickly and easily [35]. Currently, Generative AI tools enable users to create diverse artefacts by providing instructions in natural language called “prompts”. Generative AI systems can also synthesise diverse concepts and generate unpredictable ideas. In the case of AI image generators — the focus of our study — the output comes from the latent space of a deep learning model, arising from an iterated diffusion process that involves the model arranging pixels into a composition that makes sense to humans [66]. Because of process randomness, different results can be obtained based on the same prompt, with entirely new images each time. This differs from conventional image searches, where the search is performed by entering a query into a database to retrieve images that the search engine considers relevant. Another difference is that whereas long and specific queries might be too restrictive for an image search engine, they can benefit AI image generators.

Previous works have explored the roles that generative AI can play in the creative process [29]. For instance, AI can generate content entirely by itself with instructions from the user, or it can act as a creativity support tool, augmenting the user’s creativity [43]. AI text generators can be used as a tool to define specific problems to solve and promote convergent and divergent thinking [72] and have the potential to be used as a co-creative assistant for a designer [19, 54]. Professionals in creative industries claim that AI could be a promising tool to gather inspiration [3].

With the growing interest in AI, HCI researchers have also started to explore ways of using AI as a creativity support tool [16, 31]. Among these explorations, a growing stream of literature focuses on using generative AI to access inspiration and mitigate design fixation. Researchers speculate that generative AI will become a potential solution for inspiring designers [37, 51, 55] due to the ability of AI generators to create abstract and diverse stimuli [32].

One of the early examples in HCI for incorporating AI to mitigate design fixation was the *Creative Sketching Partner (CSP)* [19, 32], an AI-based creative assistant that generates inspiration for creative tasks. Through multiple studies, Davis et al. [19] suggest that the CSP helped participants in ideation and in overcoming design fixation. Hoggenmueller et al. have also explored how generative text-to-image tools can support overcoming design fixation experienced in the field of Human-Robot Interaction [27]. They conducted a first-person design exploration and reflection using “CreativeAI Postcards” inspired by Lupi and Posavec’s “Dear Data book” method to ideate and visualize robotic artefacts. They noted that AI-generated images have the potential to inspire new robot aesthetics and functionality and also claimed that the designer’s AI-co-creativity can help to eliminate biases and expand limited imagination. In a different case, Lewis [39] reflects that a digital assistance tool like “ChatGPT” helped her by acting as an art teacher and providing instructions. Lewis points out that it is challenging to distinguish between inspiration and copying when utilizing generative AI and reflects on concerns such as “*transparency of attribution*”, “*ethical considerations*”, and the clarity of the “*creation process*”. Rafner et al. [48] conducted an in-the-wild study to examine the effects of AI-assisted image generation on creative problem-solving tasks, aiming to investigate the effects of generative AI on problem identification and problem construction. They developed a human-AI co-creative technology that combines a GAN and stable diffusion model to support AI-assisted image generation. They found that this intervention enabled participants to facilitate idea expansion and prompt engineering, suggesting that AI can “aid users in generating new ideas and refining their initial problem representations” [48].

As the domain of AI-powered creativity support is still in its infancy, the available literature provides only a nascent understanding of the effect of AI on creativity and design fixation. Our work extends the literature by using established techniques from design fixation research to better understand how AI image generators affect design fixation during a visual design task.

3 METHOD

We conducted a between-participants experiment to understand how AI-generated imagery affects designers’ divergent thinking during visual ideation after being exposed to an example design. We compared this scenario to the use of online image search and to no inspiration support. The independent variable was the Inspiration Stimulus: none (BASELINE), Google Image Search (IMAGE SEARCH), or Generative AI (GENAI). The dependent variables were the DESIGN FIXATION score (the number of features in each sketch in common with the example), FLUENCY (the number of sketches produced), VARIETY (the number of different types of sketches produced), and ORIGINALITY (how infrequently other participants devised the same type of sketch). We conducted the experiment in a controlled laboratory setting following a mixed-method approach. All participants gave informed written consent to participate after reading a plain language statement describing the procedure. The study received ethics approval from our university.

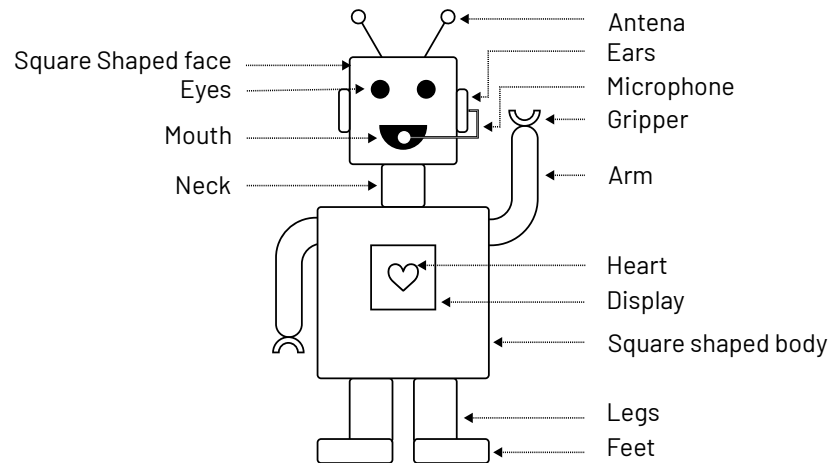


Figure 1: The example with the 14 salient features we monitored. Note: The example was given to the participants without the callouts.

3.1 Study Design and Materials

The experimental task consisted of a visual ideation activity in which participants were asked to devise as many ideas as possible for a new chatbot avatar by sketching them on paper. The written design brief given to participants was:

“Your task is to design a character we plan to use as an avatar for a chatbot. This chatbot is kind, loving, caring, and intelligent. It can assist you in solving your problems and is always there for you to talk to whenever you need to. So, imagine that you are conversing with this chatbot in real life and then come up with as many sketches as possible. Remember, you can annotate the sketch if you need to explain more about your design. And please always number each sketch you draw in the order you come up with them.”

This written design brief included an example of an avatar with the figure caption “Example chatbot avatar (for reference only)”. The example avatar is shown in Figure 1.

Further, we provided verbal instructions for the participants, asking them to produce as many different ideas as they could during the experiment. For participants in the Image search and Gen AI conditions, we additionally informed them that they could use the digital tool (either Google Image Search or Midjourney, depending on the condition) to gather inspiration for their work. The full study protocol can be found in supplementary material.

Similar to previous work [30], we started the task by showing participants an example avatar to induce design fixation. We drew inspiration from Ward’s creature invention task [36, 70], which asked participants to imagine and create animals that lived on a different planet. The authors of this paper created the example

chatbot avatar after several design iterations. We created the avatar so that it had 14 salient features, which we used to quantitatively assess design fixation (see Figure 1). We considered the presence of these features in participants’ ideas to be evidence of design fixation, following standard practice in the literature [30]. In the experimental task, participants were given 20 minutes to sketch their ideas for addressing the brief. We chose this time limit because it is the median time given to participants in previous design fixation studies [64] and because we aimed to cap each experimental session at one hour to avoid fatigue. We provided participants with pencils, pens, felt pens, and coloured pencils, along with blank A4 sheets to sketch their ideas. A timer was placed outside their peripheral view for them to keep track of time.

The experiment included a single between-participants independent variable—the INSPIRATION STIMULUS available during the task—with three levels:

- **BASELINE:** no inspiration support.
- **IMAGE SEARCH:** Participants had access to Google Images⁵ during the task, accessed through a web browser in incognito mode to avoid the browser history influencing results.
- **GENAI:** Participants had access to the paid version of Midjourney V4, an AI image generation tool, through a private Discord server running the Midjourney bot (which was required to enter prompts and view outputs from the model). Midjourney V4 was the default model when our study was conducted (May 2023)⁶. Participants interacted with Midjourney through textual prompts that the model used to generate sets of four images per prompt.

⁵<https://images.google.com/>

⁶<https://www.midjourney.com>

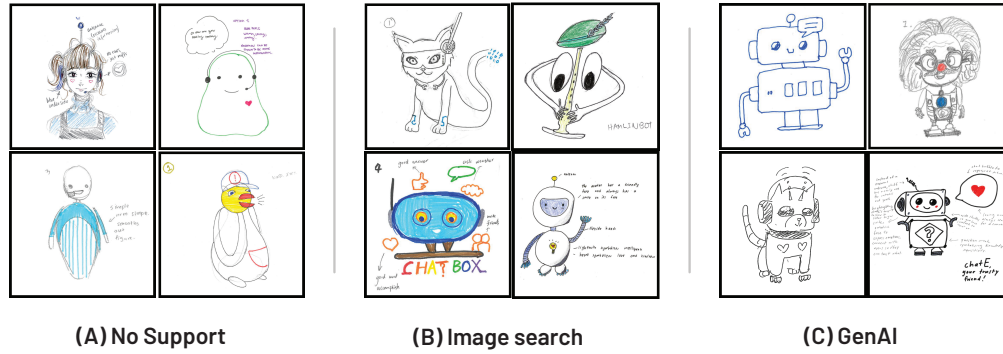


Figure 2: Examples of sketches created by participants in each experimental condition. (A) No support condition, (B) Image search condition, (C) GenAI condition

We assessed participants’ creative output using four standard measures from the design fixation literature: design fixation, fluency, variety, and originality, which we describe as follows:

DESIGN FIXATION is the unintentional conformity towards existing ideas or concepts that limits exploration of the ideation space [30, 76]. Researchers use the degree of copying as a method to quantify design fixation [30, 45]. Therefore, we operationalise design fixation as an objective property of each sketch based on the presence or absence of features available in the example. Following the approach used in design fixation literature [45], two raters blind to the experiment’s aims counted the presence of features from the example avatar in the sketches created by the participants. We validated the ratings by computing the inter-rater reliability and computed the design fixation score (DFS) as follows:

$$\text{Design fixation score} = \frac{\text{Number of features repeated from the example}}{\text{Number of fixing features in the example}} \quad (1)$$

FLUENCY refers to the number of ideas produced by the participants [25, 62]. We operationalise it by counting the number of sketches produced by each participant within the available time (20 minutes).

$$\text{Fluency} = \text{Number of sketches produced by the participant} \quad (2)$$

VARIETY measures the coverage of the solution space explored during the idea-generation process [53]. It aims to capture the extent of the design space covered during ideation. If the majority of ideas are similar, it indicates less variety. To compute variety, we assigned a numerical identifier to all the sketches (N=277), imported them into a Miro⁷ (an online collaborative whiteboard), and displayed them in randomised order. Two raters (blind to the conditions) iteratively and inductively grouped similar sketches into mutually exclusive clusters. This activity considered several factors: appearance, embodiment, appendages, shape, and accessories. The process resulted in 83 clusters. Each participant received a VARIETY score based on the number of clusters their sketches were classified into. We subtract 1 from the number of clusters so that if all of a

participant’s sketches belong to the same cluster, their score is 0, and if they have sketches in every cluster, their score is 1.

$$\text{Variety} = \frac{\text{Number of clusters that a participant's sketches belong to} - 1}{\text{Number of clusters} - 1} \quad (3)$$

ORIGINALITY (also called Novelty [23, 53]) refers to the uniqueness of a particular sketch within the total pool of sketches made by participants [25, 30]. It measures how unusual and unexpected a given idea is. Intuitively, the more people have the same idea, the less original it is. We computed an idea’s originality by counting the number of other participants who had an idea belonging to the same cluster, dividing it by the total number of other participants, and computing its complement to 1 (to normalise the value between 0 and 1). In other words, it is the proportion of *other* participants who did *not* have the same idea. This score is 0 when every participant had an idea in the same cluster and 1 if only a single participant had an idea in that cluster.

$$\text{Originality} = 1 - \frac{\text{Number of other participants with ideas in the cluster}}{\text{Number of other participants}} \quad (4)$$

3.2 Participants

We recruited 60 participants through digital student notice boards, mailing lists of university student clubs, and word of mouth. Participants expressed their interest through a digital signup form. Participants self-described their prior experience in visual design (measured in years/months). We did not specify this experience should only be professional design experience. We screened participants based on our eligibility criteria and invited those who were 18 years or older and had experience in visual design via email. Further, to avoid dependent relationships, we ensured that none of the participants had a direct connection with the primary researcher running the study. Participants had a mean age of 25.8 years (18–49, SD = 5.4). They included undergraduate, master’s and PhD students from diverse domains such as arts, business, computer science & IT, design, engineering, and science. Each condition had an equal number of participants and was gender-balanced, with 10 women and 10 men per condition (gender was self-described by participants).

⁷miro.com

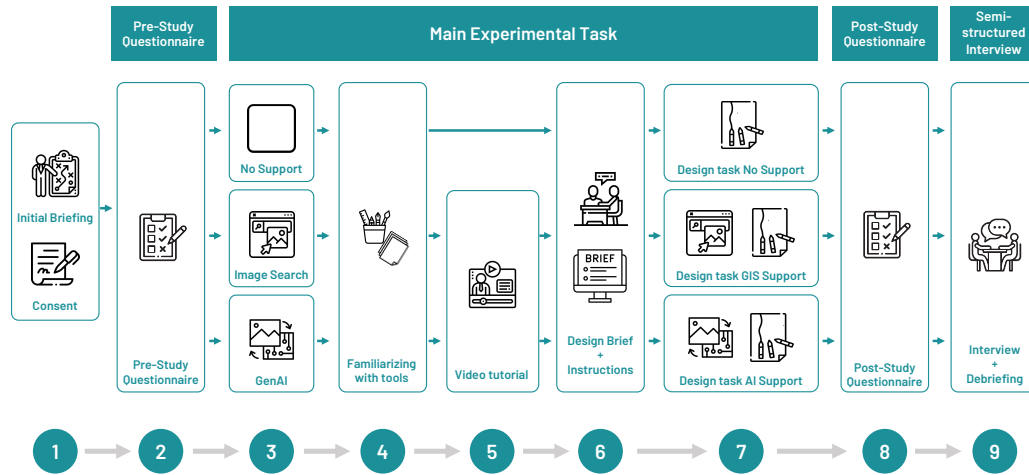


Figure 3: The overall experiment flow 1: Initial briefing and participant consent, 2: Pre-study questionnaire, 3-7: Main experimental task, 8: Post-study questionnaire, 9: Semi-structured interview and debriefing

3.3 Procedure

Participants booked a time to participate individually based on their availability. The study was carried out in a quiet research laboratory. Upon arrival, participants read a plain language statement describing the study and consented to participate (Figure 3-1).

The experiment had four stages: pre-study questionnaire, main experimental task, post-study questionnaire, and semi-structured interview. Each session lasted 45–60 min in total. In the **pre-study questionnaire**, we collected participants' basic demographic information, their experience with similar design tasks (measured in years/months), and their familiarity with AI image generators (a yes/no question, and participants were asked to list any systems they had used if they answered yes). (Figure 3-2). The main objective of this questionnaire was to understand and control for any variables that might confound the results. After completing the questionnaire, the participants were randomly assigned to one of the three conditions and were assigned a unique ID generated by the computer (3 random digits) (Figure 3-3).

In the **main experimental task** (Figure 3, steps 3-7), participants in all conditions received the same design brief, which asked them to design an avatar for a chatbot in 20 minutes, as described in Section 3.1. We started by allowing participants to familiarise themselves with the available materials. Then, the participants assigned to the image search and AI-supported groups received an introduction to the tool they would use during the design task (Figure 3-5). These tools were available for them to use on an Apple MacBook M1 Pro laptop. The tool introduction included a video tutorial created by the research team. This video tutorial explained how to use the tool. After the video tutorial, we allowed participants to ask questions and clarify any doubts.

We provided task instructions to participants both verbally and as a written brief. The written brief included an example of a chatbot avatar, which served as a stimulus to induce design fixation (Figure 3-6). Participants were given 20 minutes to complete the design task (Figure 3-7). We limited the design task to 20 minutes to minimise the possibility of fatigue and because previous studies considered it an ideal duration for maintaining focus for producing ideas with both quality and quantity [64, 74]. Once participants indicated they were ready to start, the researcher started the screen recording with participants' consent (in IMAGE SEARCH and GENAI conditions), switched on the timer and left them alone to work in the room, allowing them to work independently.

After the design task, the researcher entered the room and asked the participant to fill in the **post-study questionnaire** (Figure 3-8). As the post-study questionnaire, we administered the NASA-TLX [26] to ensure that all conditions induced an equivalent workload. To analyze the NASA-TLX, we used a one-way ANOVA; the effect of the independent variable "condition" on the NASA-TLX overall score was not statistically significant ($F(2, 57) = 1, p = 0.37$). Therefore, we did not conduct post-hoc tests.

Then, the researcher conducted a **semi-structured interview**. Each semi-structured interview lasted 15–20 min. Through the semi-structured interview, we aimed to get insights into the participant's background and their past experience in creating logos and avatars. We also probed for possible feelings of design fixation during the experiment and how it was affected by their previous knowledge, experience and process. In addition, we asked questions to understand how the stimuli (or lack thereof) affected their ideation process. To conclude the study, we debriefed the participants about the purpose of the research. We thanked each participant with a \$20 gift voucher.

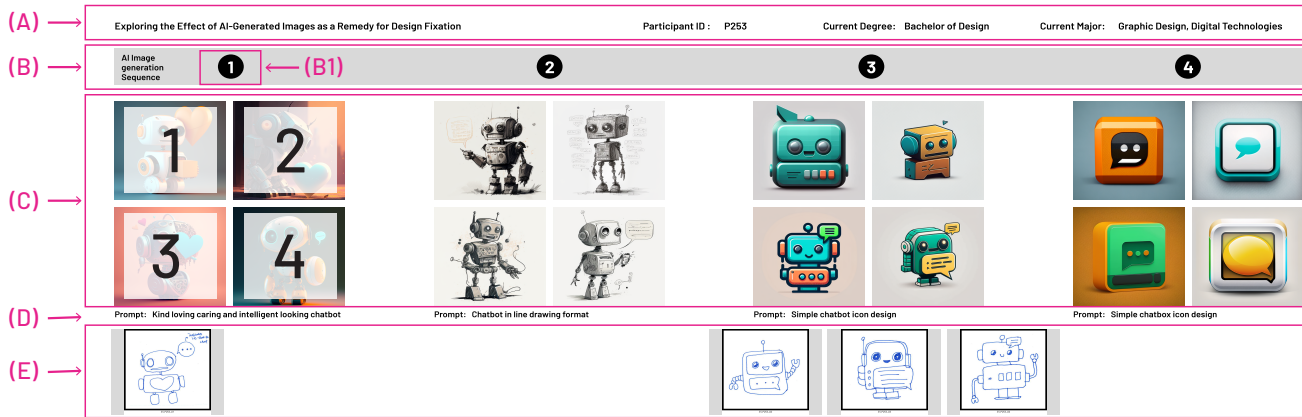


Figure 4: An example of visual sequence board (A): Participant information and meta data, (B): AI image generation sequence, (B1): Image generation number, (C): AI-generated images in the order 1-2-3-4, (D): Prompt used for each generation, (E): Participant sketch sequence.

3.4 Data Preparation

We scanned all the sketches participants created and assigned them a unique identifier. Two independent evaluators rated the sketches to compute the design fixation score, variety, and originality measures. These evaluators were researchers from the human-computer interaction domain with experience in teaching and evaluating design.

We extracted all prompts and images from the Midjourney gallery where the logs were saved (not visible to the participants), compiled the sketches and arranged them in the sequence in which they were created as a visual sequence board. Underneath the AI-generated images, we added the sketches of the participants (Figure 4).

3.5 Data Analysis

We used a mixed-method approach for our analysis. For quantitative analysis of design fixation and divergent thinking, we built Bayesian statistical models to quantify relationships between our dependent and independent variables (see subsection 4.1). We employ Bayesian statistical methods to analyze our results, opting for this approach due to its added flexibility, capability to quantify uncertainty, better handling of small samples, and greater potential for future extensibility. For a comprehensive rationale advocating the use of Bayesian methods over traditional frequentist statistics in the field of Human-Computer Interaction (HCI), see Kay et al. [33]. Readers who may not be familiar with these methods can find a beginner-friendly introduction in McElreath [44] and can see examples of their practical application in HCI in Schmettow [52]. In this manner, we shift the focus away from p-values and dichotomous significance testing, directing our discussion towards causal modelling and parameter estimation.

For qualitative analysis of participants’ interview data, we used Braun and Clarke’s 6-phase approach to reflexive thematic analysis [7, 60]. The analysis was inductive, i.e. data-driven, based on transcripts of the interviews. Each phase of the analysis was progressed using NVivo12⁸ for coding procedures, theme development and

⁸<https://lumivero.com/products/nvivo>

naming. The analysis aimed to understand potential causes of design fixation during the experiment and participants’ approaches to creating sketches in each condition. In this paper’s findings, we use interview quotes to illustrate participants’ approaches to prompt creation and their stated approaches to ideation based on AI images. This enables us to probe plausible explanations for observed differences between experimental conditions and explore why particular kinds of sketches were created in response to AI-generated images.

4 RESULTS

4.1 Statistical Analysis

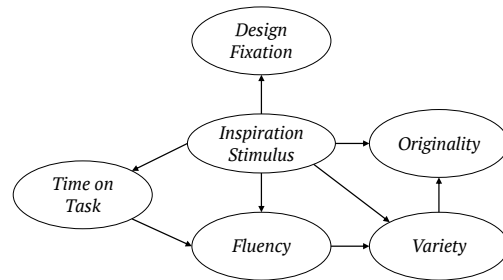


Figure 5: Theorised causal directed acyclic graph.

We summarise our theoretical claims as a directed acyclical graph (DAG) in Figure 5. We argue that the INSPIRATION STIMULUS affects users’ DESIGN FIXATION, FLUENCY, VARIETY, and ORIGINALITY. The choice of inspiration stimulus affects how much time is spent on the sketching task (as opposed to seeking inspiration), which, in turn, affects the number of sketches produced (fluency). Higher fluency is also likely to lead to higher variety—as producing more sketches also increases the likelihood that they will cover more

ground during ideation. A greater variety of sketches, in turn, will likely lead to more original ideas.

We used the brms package [8] to fit our models. This package facilitates the implementation of Bayesian multilevel models in R, leveraging the Stan probabilistic programming language [11]. To ensure the reliability of our Bayesian Markov Chain Monte Carlo (MCMC) sampling process, we assessed convergence and stability through two metrics: R-hat, which ideally should be less than 1.01 [65], and the Effective Sample Size (ESS), which should ideally exceed 1000 [8]. All of our model estimates met these criteria. We built our models based on the original count data in the direct measurements but report normalised values as described in Section 3.1 in our plots for easier comparisons with future work.

In our reporting of model results, we present the posterior means of parameter estimates, their corresponding standard deviations, and the boundaries of the 89% compatibility interval, often referred to as the credible interval. The choice of an 89% compatibility interval aligns with the recommendation by McElreath [44] to mitigate potential confusion with the frequentist 95% confidence interval, as the two intervals have distinct interpretations. The compatibility interval specifies the range of values within which there is an 89% probability that the true value lies. We report hypothesis test results using Bayes Factors, which compares the likelihood of the observed data under the proposed model over the null. We interpret these values following Wagenmakers et al. [69], considering values above one as supporting a given hypothesis, values under 3 offering anecdotal evidence; under 10, substantial evidence; under 30, strong evidence; under 100, very strong evidence; and above 100, extreme evidence. We note that p-values are not used in Bayesian statistics, and no claims about “statistical significance” should be derived from our results.

4.2 Design Fixation

To model DESIGN FIXATION, we consider the number of salient features in participants’ sketches also found in the example avatar provided at the beginning of the experiment. We model this data as a binomial distribution with $N = 14$ (the maximum number of features) and a probit link. We use weakly informative, regularising priors for the model parameters (drawn from a normal distribution with mean zero and standard deviation of 2). We model the random effects of participants and images as being drawn from a normal distribution with mean zero and standard deviation computed from the data through partial pooling.

The model suggests that the effect of GENAI has a 100% probability of leading to higher DESIGN FIXATION (mean = .32, 89% CI [.11, .55]), and a Bayes Factor of 124 suggests extreme support for the hypothesis that inspiration from GENAI leads to higher DESIGN FIXATION than the baseline. The effect of IMAGE SEARCH on DESIGN FIXATION was also detrimental, but less so (mean = .27, 89% CI [.05, .49]), with a 98% probability of this effect leading to higher DESIGN FIXATION. A Bayes Factor of 42.48 suggests very strong evidence for the hypothesis of a higher Design Fixation than the NO SUPPORT baseline. In summary, our model suggests that, **on average, both stimuli led to more features in common with the example avatar, and GENAI led to even more design fixation than IMAGE SEARCH.**

Table 1: Summary of the binomial model for design fixation: $\text{DFS}|\text{trials}(14) \sim \text{Stimulus} + (1|\text{Participant ID}) + (1|\text{Image ID})$. We provide the posterior means of parameter estimates (Est.), posterior standard deviations of these estimates (SD), and the bounds of their 89% compatibility interval. We note that this is not the same as the frequentist confidence interval but a percentile of the posterior distribution. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00.

Parameter	Est. (SD)	89% CI
Intercept	-.71 (.10)	[-.86, .56]
Image Search	.27 (0.14)	[.05, .48]
GenAI	.32 (0.13)	[.11, .54]

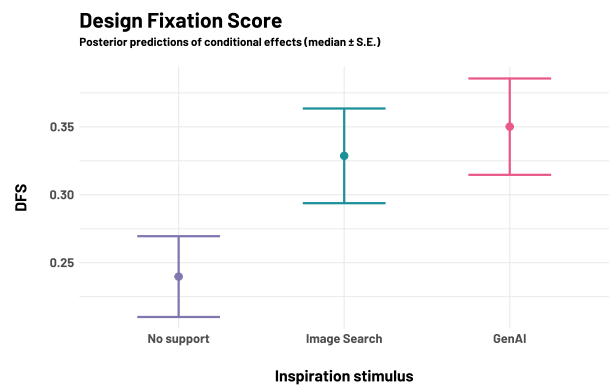


Figure 6: Model posterior predictions for DESIGN FIXATION scores. Error bars represent the standard error of the estimates. Scores correspond to the percentage of salient features in the example found in participants’ sketches (higher is worse).

4.3 Fluency

To model FLUENCY, we consider the number of sketches produced by each participant. Our causal model considers two effects of the stimulus on FLUENCY: a direct effect and an effect mediated by TIME ON TASK (ToT). We model these effects through two models, with and without OFFSET(LOG(ToT)) as a covariate. This approach was taken to account for varying time on task as participants of both GenAI and Image search utilised different times to sketch. In both cases, we model the expected value for each response as a negative binomial distribution with a log link. This models the sketch count based on the mean and the shape parameter, both of which depend on the inspiration stimulus. We opted for this model instead of a Poisson model due to its ability to model overdispersion. We use weakly informative, regularising priors for the model parameters (drawn from a normal distribution with mean zero and standard deviation of 2).

In the total effects model, both IMAGE SEARCH and GENAI demonstrate detrimental effects on FLUENCY. Specifically, for GENAI, there is a notable negative impact on Fluency with an estimate of -.21

Table 2: Summary of the negative binomial models for FLUENCY—Direct effects: Fluency \sim + offset(log(Time on Task)) and Total Effects: Fluency \sim Stimulus. We provide the posterior means of parameter estimates (Est.), posterior standard deviations of these estimates (SD), and the bounds of their 89% compatibility interval. We note that this is not the same as the frequentist confidence interval but a percentile of the posterior distribution. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00.

Parameter	Direct Effects		Total Effects	
	Est.(SD)	89% CI	Est.(SD)	89% CI
Intercept	-1.48 (.19)	[-1.79, -1.17]	1.50 (.19)	[1.20, 1.80]
Image Search	-.22 (.25)	[-.63, .17]	-.49 (.25)	[-.88, -.10]
GenAI	.10 (.31)	[-.39, .58]	-.21 (.30)	[-.66, .26]
Intercept _{shape}	.75 (.45)	[.04, 1.47]	.77 (.46)	[.04, 1.49]
Image Search _{shape}	1.58 (1.15)	[-.01, 3.61]	1.66 (1.14)	[.06, 3.74]
GenAI _{shape}	-.51 (.61)	[-1.49, .47]	-.40 (.65)	[-1.41, .66]

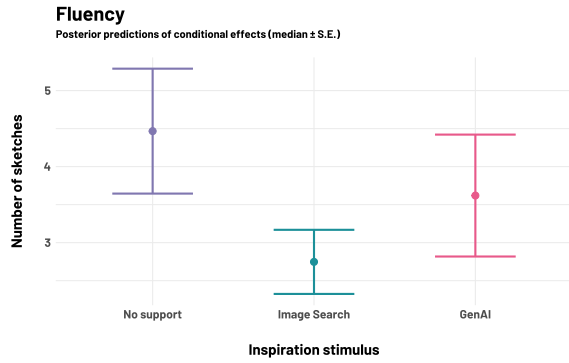


Figure 7: Model posterior predictions for FLUENCY (number of sketches generated by each participant). Error bars represent the standard error of the estimates.

(89% CI [-.7, .28]) and a Bayes factor of 3.20, suggesting a 76% posterior probability of a negative effect. This is a substantial indication of its negative total effect on Fluency. The effect of IMAGE SEARCH is more pronounced (mean = -.49, 89% CI [-.9, -.10]) with a Bayes factor of 46.62, indicating a 98% posterior probability of a negative effect, strongly supporting its detrimental influence on Fluency.

In the direct effects model, which accounts for the time-on-task, the impact of IMAGE SEARCH on FLUENCY is minimal (mean = -0.22, 89% CI [-0.64, 0.19]) with a Bayes factor of 4.20, indicating an 81% posterior probability of a negative effect. In contrast, GENAI shows a relatively neutral direct effect on FLUENCY (mean = .10, 89% CI [-.41, .60]) with a Bayes factor of .61, implying only a 38% posterior probability of a negative effect.

In summary, **neither IMAGE SEARCH nor GENAI enhanced Fluency compared to the baseline, with both generally resulting in lower Fluency.** The effect of IMAGE SEARCH on Fluency is minimal when controlling for total output time, indicating a less direct impact. However, GENAI does not exhibit a considerable direct negative effect on FLUENCY, highlighting its influence is not strongly dependent on the total time available for sketching.

4.4 Variety

To model VARIETY, we consider the number of clusters a participant’s sketches belong to minus one to account for the fact that variety only begins with the second sketch. Our causal model considers two effects of the stimulus on VARIETY: a direct effect and an effect mediated by FLUENCY. We model these effects through two models, with and without FLUENCY as a covariate. In both cases, the expected value for each response is based on a negative binomial model with a log link. We use weakly informative, regularising priors for the model parameters (drawn from a normal distribution with mean zero and standard deviation of 2 for coefficients and a gamma distribution with parameters set to 0.01 for the shape).

Table 3: Summary of the negative binomial models for VARIETY—Direct effects: Variety \sim Stimulus + Fluency and Total Effects: Variety \sim Stimulus. We provide the posterior means of parameter estimates (Est.), posterior standard deviations of these estimates (SD), and the bounds of their 89% compatibility interval. We note that this is not the same as the frequentist confidence interval but a percentile of the posterior distribution. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00.

Parameter	Direct Effects		Total Effects	
	Est. (SD)	89% CI	Est. (SD)	89% CI
Intercept	.00 (0.23)	[-.38, .36]	1.01 (.19)	[.71, 1.30]
Image Search	.02 (0.25)	[-.38, .42]	-.39 (.28)	[-.84, -.06]
GenAI	-.15 (0.24)	[-.53, .23]	-.29 (-.28)	[-.74, .17]
Fluency	.14 (0.02)	[.11, .18]		

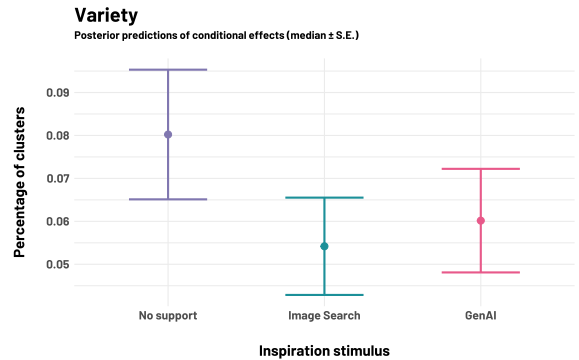


Figure 8: Model posterior predictions for VARIETY (percentage of clusters in which the participant has sketches). Error bars represent the standard error of the estimates.

The total effects model shows a detrimental effect of both stimuli on VARIETY. The model suggests that the effect of GENAI has only an 86% probability of being negative (mean = -.29, 89% CI [-.75, .16]), and a Bayes Factor of 5.93 provides substantial evidence for the hypothesis that it yields a negative total effect on the variety of output. The effect of IMAGE SEARCH was even more negative (mean = -.40, [-.87, -.07]), with a Bayes Factor of 11.38 providing strong support for the hypothesis that it has a negative effect.

The model including FLUENCY as a covariate, models the direct effect of the stimulus on the VARIETY of the output. Comparing

the two models, we see that after accounting for the number of sketches that the participant produced, Google Image Search did not have much of an effect on VARIETY (mean = .02 [-.39, .43]). However, GenAI still had a negative effect (mean = -.15 [-.55, .24]), with a Bayes Factor of 2.64 suggesting anecdotal evidence for this effect being negative. In summary, **neither IMAGE SEARCH nor GENAI provided meaningful support over the baseline in terms of enhancing the variety of the output, yielding, on average, lower variety than the baseline. The effect of IMAGE SEARCH was fully mediated by FLUENCY, but GENAI also had an additional negative direct effect on VARIETY.**

4.5 Originality

To model ORIGINALITY, we consider the number of other participants with sketches in the same cluster as each sketch. As in the case of VARIETY, our causal model considers two effects of the stimulus on ORIGINALITY: a direct effect and an effect mediated by VARIETY. We model these effects through two models, with and without VARIETY as a covariate.

Table 4: Summary of the linear regression model for ORIGINALITY—Direct Effects: $\text{Originality} \sim \text{Stimulus} + \text{Variety} + (1|\text{Participant ID})$ and Total Effects: $\text{Originality} \sim \text{Stimulus} + (1|\text{Participant ID})$. We provide the posterior means of parameter estimates (Est.), posterior standard deviations of these estimates (SD), and the bounds of their 89% compatibility interval. We note that this is not the same as the frequentist confidence interval but a percentile of the posterior distribution. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00.

Parameter	Direct Effects		Total Effects	
	Est. (SD)	89% CI	Est. (SD)	89% CI
Intercept	.83 (0.02)	[.81, .86]	.86 (.01)	[.84, .88]
Image Search	-.01 (0.02)	[-.03, .02]	-.01 (.02)	[-.04, .01]
GenAI	-.03 (0.02)	[-.05, -.01]	-.03 (.02)	[-.06, -.01]
Variety	.01 (<0.01)	[.00, .01]		

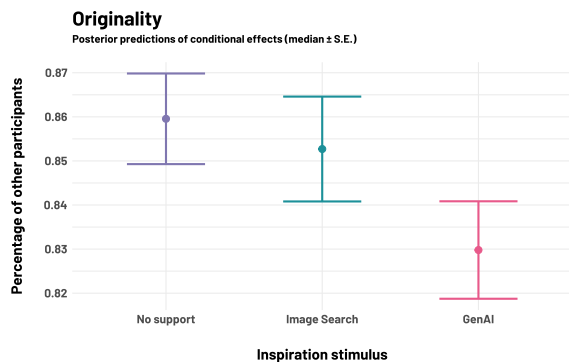


Figure 9: Model posterior predictions for originality (percentage of other participants who did not have an idea in the same cluster of an idea, averaged per participant). Error bars represent the standard error of the estimates.

The expected value for each response is based on a linear regression model. This models the originality score based on the inspiration stimulus (and variety score in the direct effects model), as well as a random effect of the participant. We use weakly informative, regularising priors for the model parameters (drawn from a normal distribution with mean zero and standard deviation of 2 for coefficients). We modelled our random effects as being drawn from a normal distribution with mean zero and standard deviation computed from the data through partial pooling.

The total effects model suggests that both stimuli had small negative effects on ORIGINALITY. The model suggests that the effect of GENAI has a 97% probability of being negative (mean = -.03, 89% CI [-.06, .00]), and a Bayes Factor of 35. provides extreme evidence for the hypothesis that it yields a negative effect on the variety of output. The effect of IMAGE SEARCH was slightly less negative but also rather small (mean = -0.01, [-.04, .01]), but a Bayes Factor of 3.9 provides only anecdotal evidence against the hypothesis that it has a positive effect. Adding VARIETY did not change the model in any meaningful way, suggesting that VARIETY does not mediate an effect on ORIGINALITY. In summary, **neither IMAGE SEARCH nor GENAI provided a considerable aid in terms of developing ORIGINALITY of the output, offering, on average, lower originality than the baseline, but these effects were negligible.**

4.6 Why did ideating with Generative AI cause design fixation?

The results from our statistical models suggest that support from Generative AI led to higher design fixation. To understand why this occurred, this section draws on our interview data, the prompts created by participants, the AI-generated images, and the participants' sketches. We first explore the content of participants' prompts as one potential cause of design fixation. This encapsulates how participants claimed to develop the prompts and how they were influenced by the design brief or the example design. Next, we quantitatively explore the similarity between participants' sketches and the AI images used to inform that sketch in terms of design fixation. We then discuss the types of AI-generated images returned by participants' prompts and the sketches created based on them, using a case-study-based approach to illustrate our claims.

Overall, our analysis indicates that participants frequently relied on prompts containing keywords copied directly from the design brief or used prompts inspired by the example design. These prompts resulted in AI-generated images that were conceptually similar to the example design in 44% of cases and which frequently contained fixating features that were present in the example design. Further, while not all sketches exhibit high similarity to the example we provided, ideating based on AI images can lead to *fixation displacement*, where participants simply fixate on the images generated by the AI and copy what they see. This can occur irrespective of whether the participant imitates the example design or whether they attempt to explore other areas of the conceptual space.

Table 5: Frequency of words used in the prompt by the participants in GenAI condition.

Word	Length	Count	Weighted percentage	Similar words	Included in the brief
robot	5	38	10.80%	robot, robots	No
kind	4	27	7.67%	kind	Yes
chatbot	7	24	6.82%	chatbot	Yes
intelligent	11	20	5.68%	intelligent	Yes
cute	4	20	5.68%	cute	No
caring	6	19	5.40%	caring	Yes
loving	6	17	4.83%	love, loving	Yes

4.6.1 “I just took the words from the brief”: Fixation from prompts based on the brief and example design. One plausible source of design fixation in our experiment is the prompts that participants used to generate AI images. That is, if prompts include terms that are closely related to the example design or which draw from the design brief, then they might conceivably give rise to AI-generated images that are similar.

To investigate this possibility, we first analysed the prompts that participants used for generating images. Participants created a total of 117 prompts, with a mean of 5.85 prompts per participant (median = 5.5 range = 2–15). The length of each prompt ranged from 1 to 26 words (mean = 3.5 words). To explore the content of the prompts, we conducted a simple word frequency analysis using the automated word counting feature in NVivo12. This feature enables us to identify the total number and frequency of unique words that appear in the prompts. Table 5 shows a summary of the most frequent words appearing in participants’ prompts.

This analysis revealed that participants frequently created prompts by using keywords copied from the design brief. In total, 52 prompts (44%) contained at least one word that appears in the brief. Examples included *kind*, which appeared in 27 different prompts; *chatbot*, which appeared in 24 prompts; *intelligent*, which appeared in 20 prompts; and *caring*, which appeared in 19 prompts. During the interviews, participants reported using this approach due to a feeling of being ‘stuck’ when trying to develop a prompt. Others attempted to generate ideas that met the requirements of the design brief. GenAI-P253, for example, described adapting content from the brief into his prompt and told us that the process he followed was to “read the brief, take the descriptions that they had, and make sure that I was meeting those descriptions.”

A second approach involved participants using keywords that were themed around the example design. In total, 57% (67/117) of the prompts contained the word *robot*, *chatbot* or *chatbox* (a homophone of chatbots).

This suggests that participants often translated what they saw into a prompt before ideating based on the results. In addition, 78 prompts (66.6%) included terms related to robots alongside words from the design brief. For example, P437’s very first prompt was ‘*kind loving caring robot*’ whereas GenAI-P253 entered ‘*cute kind chatbox character*’.

Data from the interviews also supports the notion that participants created prompts which were fixated on the example design. GenAI-P605, for example, described their process as starting with ‘robots’ and then trying to factor in other aspects of the design

brief. He said that he, “searched up intelligent robots. But all those robots that I saw in the [AI-generated images], they looked intelligent, but they didn’t look kind or caring. [I thought], how can I make it both caring and intelligent?”. This participant created three distinct prompts: *Intelligent robot*, *Caring robot*, and *Baymax* (referring to an inflatable computerised robot from a Disney movie) in an attempt to come up with alternative ideas.

However, it is worth noting participants created 39 prompts that did not contain words from the brief or phrases related to robots. These prompts evince participants’ attempts to explore different possibilities within the conceptual space of a ‘kind and loving’ character. GenAI-P166, for example, recounted how they started the task by reading the brief and thinking about what to draw. This led them to the idea of ‘family’, which they then translated into three distinct but related prompts: *family*, *mom*, and *Mom - young*. They subsequently drew a sketch of a woman’s face as their only design after seeing the images Midjourney returned from these prompts.

Taken together, these cases illustrate how creating prompts based on the brief and the example design may be an initial stimulus for fixation. A successful strategy to overcome this problem was to try to think ‘beyond’ the brief and the example. This latter quality is what may be needed from AI systems that truly support designers in avoiding fixation.

4.6.2 “I would just kind of copy it and then tweak”: AI-generated images as a cause of fixation. A second putative cause of design fixation in our experiment is the AI imagery that participants saw. That is, if the AI images were not meaningfully different to the example design, then this may have encouraged fixation because participants did not consider (or simply were not exposed to) other possible alternatives. This explanation is plausible given that 66.6% of all prompts contained terms related to robots or words copied from the design brief. Prompts containing these terms might be expected to produce images similar to the example design, in turn leading to fixated sketches.

To explore the relationship between fixation in participants’ sketches and the AI images, we first computed the correlation between the design fixation score of participants’ sketches (previously calculated by two independent raters, see Section 3.1) and the design fixation score of the most recent set of AI-generated images immediately preceding each sketch. We selected Spearman’s rank correlation (a non-parametric test) as the data did not satisfy normality assumptions.

For this analysis, we begin with the total set of sketches produced by participants in the AI condition (92 in total). We found that 10 of these sketches were created prior to entering any prompts into Midjourney; we, therefore, removed these sketches from consideration as there are no equivalent AI images to compare them against. This left us with 82 sketches, which we plotted against the relevant AI-generated images seen immediately before drawing the sketch.

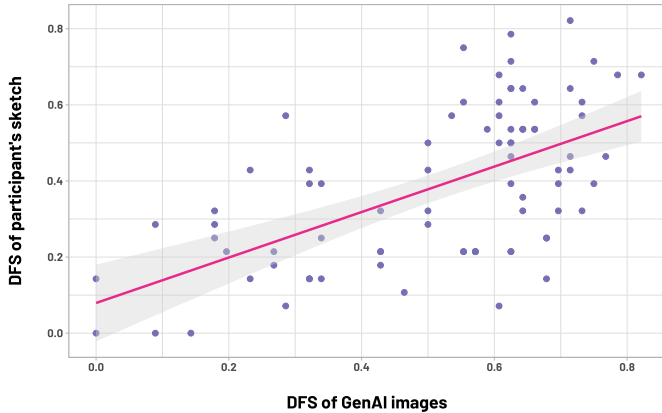


Figure 10: Scatterplot illustrating the correlation between the design fixation score of GenAI images appearing immediately before a participant’s sketch and the design fixation score of the sketch associated with the same set of GenAI images. DFS = design fixation score.

Figure 10 illustrates the correlation. We observed a moderate positive correlation between the design fixation score of each sketch and the design fixation score of the AI images immediately preceding that sketch ($\rho = 0.56$). This provides quantitative support for the idea that **AI-generated images that contained features of the example avatar led to sketches with higher design fixation scores.**

Next, we qualitatively investigated what kinds of images were generated by Midjourney in response to participants’ prompts and whether the resulting sketches were fixated on these images. We began with a simple visual inspection of the AI images to probe whether Midjourney’s outputs were meaningfully different to the example design.

This inspection revealed that 44% (206/468) of the AI-generated images portrayed humanoid robots that were conceptually similar to the example avatar. In turn, these images were qualitatively similar to the sketches participants made in response to them, indicating a tendency among participants to imitate — or even directly copy — what they saw.

Figure 11 shows an example of this phenomenon. The figure shows the AI images seen by one participant in Midjourney over time. It then positions the participant’s sketches according to the most recently issued group of AI images before the sketch was drawn. In this example, it is immediately evident that the majority of sketches are superficially similar to the images returned by Midjourney. Likewise, these sketches are similar to the example design we provided (i.e. a cutesy robot) and typically contain the same

salient features (legs, arms, and so on). The presence of these features and their inclusion in the subsequent sketches is one plausible explanation for why the AI support did not encourage participants to ‘break free’ of fixation. It appears to have merely reinforced the existing problem.

This phenomenon arose irrespective of whether participants ideated on the fly or considered multiple ideas before creating a sketch. Figure 12 illustrates a second case where the participant is once again fixated on the idea of a robot. In this instance, the participant delays sketching until after issuing multiple prompts and seeing several rounds of AI images. It can be seen that the single sketch the participant created is of a robot-type character, evidencing fixation. The sequence also highlights how the prompt plays a role in this effect, with the participant attempting to vary their initial ‘chatbot’ prompt by adding keywords such as ‘intelligent’ or ‘kind’, but resulting in thematically similar returns each time.

Cases such as these are illustrative of how fixation may have occurred, with participants repeatedly generating ideas that were similar to the example design and imitating the ideas within them. The interview data supported this latter idea. When asked about their approach to ideation, participants described the AI as a “source of inspiration” but openly admitted they sometimes copied what they observed. For example, GenAI-P253 claimed that using AI “helped a lot of the inspiration for a lot of the designs that very much I just put down what I wanted it to give me, and I would just kind of copy it and then tweak it a little bit for the designs.”

Overall, these cases highlight the risk of AI simply reinforcing the phenomenon of fixation on an initial example. In turn, they raise the question of how AI systems might be usefully designed to encourage shifts away from this effect.

4.6.3 The notion of “Fixation Displacement”. In addition to investigating how fixated sketches resulted from fixated images, our inspection of the sketches in relation to AI images revealed an additional phenomenon not well-captured by the correlational analysis. That is, there is evidence of what we describe as *fixation displacement*, where the participant creates sketches with little relation to the original design but which are very clearly fixated on the AI imagery. Here, the sketches produced are both objectively and subjectively different to the example design but demonstrate a high degree of fixation with the AI images.

Figure 13 illustrates an example of fixation displacement in action. Here, the participant entered the prompt ‘goddess’ as a way of beginning their ideation. This prompt has little connection to the design brief or the idea of a robot avatar. The participant then produced a sketch of a woman’s face, which shares a small number of features with the example robot avatar (eyes, mouth, ears) but which is qualitatively different. Then, they proceed to iterate on this idea, resulting in three sketches that are similar in appearance and which bear a close resemblance to what the participant is seeing in Midjourney.

Crucially, this phenomenon is not captured in our earlier scatterplot (Figure 10) because the images and sketches have only a few features in common with the example design. This means they would be rated as quantitatively ‘low’ on design fixation. In our experiment, fixation is operationalised in terms of similarity to the

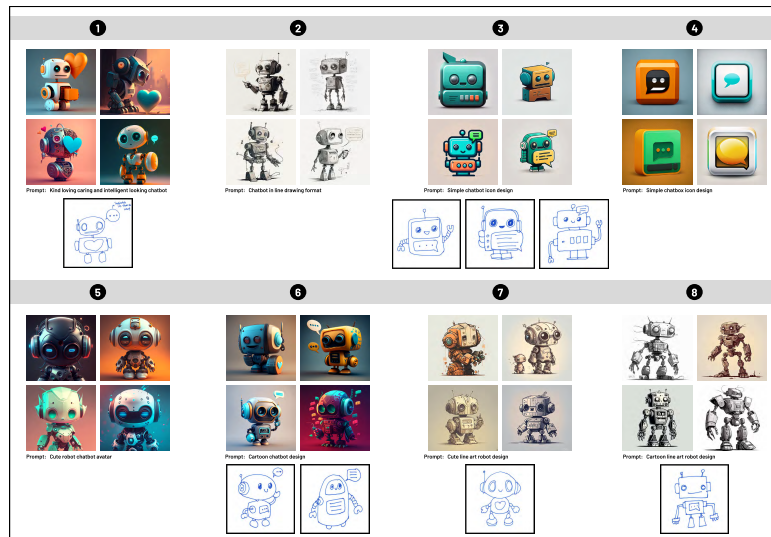


Figure 11: An example of a participant producing sketches based on AI images that are similar to the example design, evidencing fixation when co-ideating with AI.

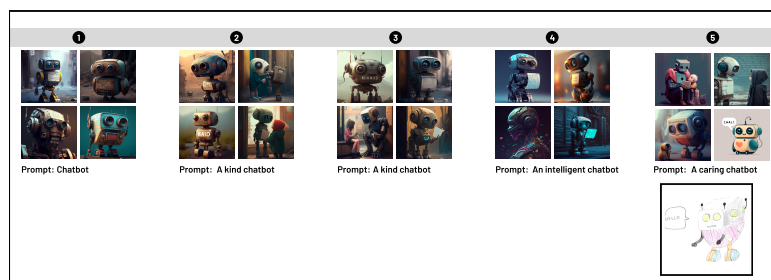


Figure 12: A second example of a fixated sketch created after several rounds of prompting and image generation. In this case, the prompt is also fixated on the idea of a chatbot, creating similar returns from Midjourney each time.

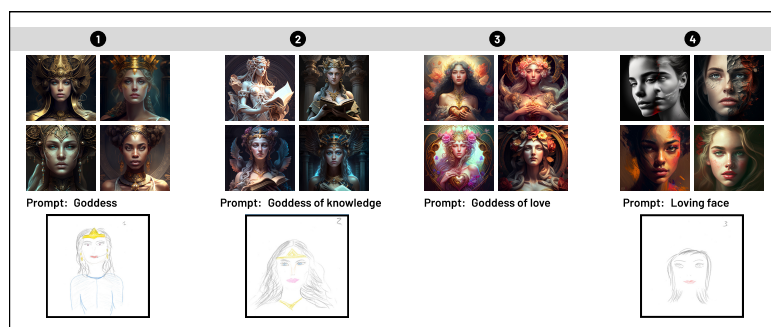


Figure 13: An example of fixation displacement, this is where the participant has shifted their sketches away from the example robot avatar but has now become fixated on the idea of a woman’s face via the AI images.

example design, where similarity is assessed by the presence or absence of features from the robot avatar. Conceptually, however, fixation refers to “blind adherence to a set of ideas or concepts” [30]. This general phenomenon is clearly depicted by the images and sketches in Figure 13, highlighting a new and novel risk of

employing AI in ideation. That is, design fixation towards an initial example may not be overcome by using AI but may simply be displaced onto the examples that the AI provides. If one operationalises fixation in terms of deviation from an initial example, one might argue that such an outcome is opposite or even desired.

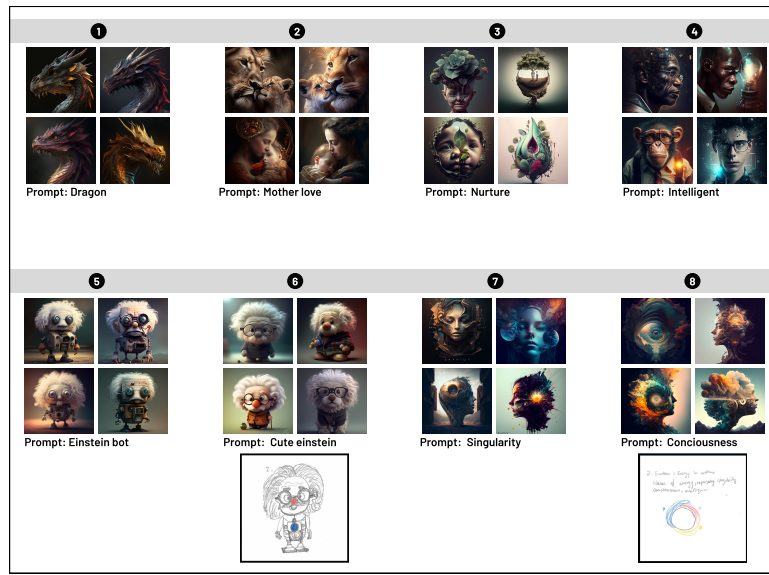


Figure 14: An example of a participant progressing through different ideas and arriving at a final sketch that bears no resemblance to the example robot avatar.

But if one operationalises fixation in terms of blind adherence to an idea, then this outcome is questionable.

What one may wish to see from AI-based ideation might be more akin to the process seen in Figure 14. Here, the participant-generated 8 groups of images from Midjourney, beginning with prompts (such as ‘dragon’) that have no relationship with the example design but which might inspire useful ideas and further exploration of the conceptual space. By the fourth prompt, the participant latched onto the idea of intelligence, which is then used to produce an Einstein-themed robot after prompt 6. However, the participant abandoned this idea and moved to an abstract design which bears no resemblance to the exemplar. While this still evidences some degree of fixation displacement, given the sketches the participant produced, it represents a significant conceptual shift from the example design. That is, the participant has considered a range of alternatives and has produced a seemingly useful design that bears no resemblance to the given example. This is perhaps more indicative of what we would consider to be effective AI-supported ideation.

5 DISCUSSION

This study aimed to identify the effects of using an AI image generator as inspiration support for an ideation task. Our quantitative analysis revealed that using AI-generated images had a detrimental effect on participants’ ideation performance. Therefore, we aimed to uncover the cause of this effect through our qualitative analysis. We identified that AI caused more design fixation in participants and hindered the variety, originality and fluency of ideas compared to the baseline condition. Further, we observed a moderate positive correlation between the design fixation score of participants’ sketches and the design fixation score of AI-generated images, which suggests that AI has a potential influence in determining

the outcome of the ideas. Further, we observed that AI induced a fixation displacement in participants where, even if they shifted their focus away from the initial example, they became fixated on the AI-generated images instead. In this section, we reflect on our learnings and discuss potential opportunities for developing generative AI to better facilitate ideation tasks, and propose strategies for improving divergent thinking during ideation. In doing so, we look at different phases of the ideation task performed by participants in detail (see figure 15).

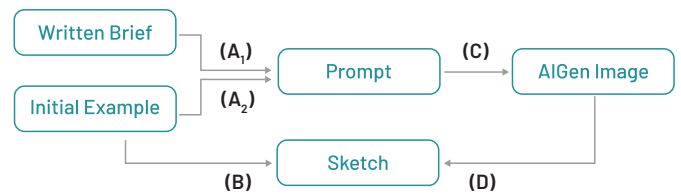


Figure 15: The overall ideation workflow of participants in the AI-supported condition, (A₁): Written brief to the prompt, (A₂): Initial example to the prompt, (B): Initial example to the sketch, (C): Prompt to AI-generated images, (D): AI-generated images to the sketch

5.1 How to avoid fixation on the design brief when determining prompts for Generative AI?

Through our study, we identified that the prompt acted as a potential source of design fixation. While participants used different strategies to devise their prompts (Figure 15-A), the results suggest that most of them used keywords from the brief (Figure 15-A₁) and

built upon the idea of a ‘robot’ from the initial example (Figure15-A₂) when devising their prompts. Participants claimed that they tried to avoid copying the initial example later during sketching (Figure15-B), but this was not reflected in the data given that the high design fixation score was calculated based on the ratio of replicated features from the initial example. Therefore, we assume that poor prompt design led to the generated images sharing similar features with the example. This suggests that design fixation happened when the participants created the prompts (Figure15-A). Participants tended to produce prompts that were semantically similar to the words given in the design brief and examples, while exhibiting strategies of repeating the same steps when creating prompts. Prior work indicates that participants can become fixated on exposed examples [10, 30]. Further, studies have shown that participants tend to fixate on self-generated ideas and concepts compared to initial examples [38]. Our study aligns with these findings, suggesting that some participants were fixated on the design brief, the example avatar, or self-generated ideas when determining keywords for the prompt.

Thus, paying attention to different prompting strategies might help mitigate this first potential occurrence of fixation when co-ideating with AI. Youmans et al. [76] summarise different strategies to mitigate design fixation based on the cause of occurrence. One strategy to snap participants out of fixation is to have timely warnings to consider alternative options [76]. Therefore, creativity support systems based on generative AI could not only turn prompts into images but also scaffold users’ abilities to craft better prompts for ideation. AI systems could support users in creatively interpreting the brief, push users’ thinking into alternative directions, or mix arbitrary ideas into the prompts. This functionality could be enabled through other generative AI techniques, such as large language models.

Cheng et al. [13] have found that showing low-fidelity, abstract, and partially completed ideas led participants to become more divergent in their thinking and reduce fixation. Users of AI systems should consider using prompts that generate low-fidelity abstract or partial images when interacting with AI because images with these qualities might alleviate design fixation and encourage divergent thinking in an ideation task. Having a predetermined prompt structure or template that describes ways to make the images more abstract and less refined might lower the risk of fixation.

5.2 How can AI generate images that better support ideation?

The images generated by the AI system in this study were high in fidelity, visual detail, and quality; they appeared to be rich in shape, form, texture, colour, composition, and visual expressiveness (see Figure 11-14). Though this showcases impressive functionality, it might have amplified conformity towards the generated image, causing fixation displacement. This aligns with prior studies, which have shown that complete and strong examples carry the potential to cause fixation [10, 13, 15, 59]. Previous work has found that introducing some incubation time can help dissolve concrete exemplars into more abstract concepts, lowering fixation [58, 73] and supporting the emergence of novel ideas [21, 56]. Though this was not possible in our study, given the short time available for the task,

it is a process that users of AI systems can incorporate into their ideation process. Further, when developing generative AI to support ideation, it may be useful to introduce mechanisms to lower the fidelity and the richness in detail of the output. Another direction is to show partially completed or blurred outputs, which might be beneficial for introducing ambiguity and pushing ideas in new directions [13]. Recent works by Davis et al. [19] and Williford et al. [71] provide initial evidence suggesting that these mechanisms might be plausible approaches to be embedded in generative AI.

5.3 How to translate AI images into design ideas?

Through our visual comparisons, we observed similarities between the sketches produced by participants and the AI-generated images, suggesting that participants had imitated and, in some instances, directly copied elements from the images generated by Midjourney. Further, we identified that regardless of whether participants ideated on the fly or considered multiple ideas before creating a sketch, they gravitated towards features of the images which Midjourney generated, leading to fixation. Copying elements from an example is the easiest way to result in fixated outputs [10, 30], and our findings were consistent with this.

To successfully act as sources of inspiration, generative AI tools must encourage more strategies that are more effective than copying. Previous work has shown that techniques like visual analogy—identifying abstract correspondences between the images being generated and the solution being sought—can improve the ideation effectiveness of designers of all levels, including novices [12]. However, Casakin and Goldschmidt highlight that even though novices already have an inherent understanding of how visual analogy works, they must be shown how to do it well and how it can support problem-solving in design activities [12]. Scaffolding these skills is a promising role for AI-based creativity support tools.

We observed lower fluency in the GenAI and Image search conditions. Because the time given to complete the task was the same in all conditions, there was an inherent trade-off between spending time producing ideas vs. seeking inspiration. Interacting with both the AI image generator and the web image search led to less time spent sketching. These results tally with the findings of Vishwanathan and Linsey [67], who found that though physical prototyping techniques that required more effort led to higher quality ideas in an engineering design task, they also increased design fixation and lowered fluency. They hypothesise that this is due to a “sunk-cost effect”—the higher the effort spent in a given direction, the harder it is to move into a different one. Participants who spent more time refining prompts and interacting with the AI also had worse ideation performance. Users of generative AI systems should be careful and deliberate in their approach when seeking inspiration from external stimuli like AI image generators to mitigate the risk of design fixation. Crilly [17] suggests that empowering designers to recognise and reflect upon fixating episodes might be beneficial in developing a less fixating co-ideation workflow with AI. Further, Neroni and Crilly [46] state that uncovering participants’ fixation tendencies, which they call “demonstrated vulnerability”, is an effective approach that can further strengthen participants’ ability to overcome fixation. In summary, when developing generative AI for

co-ideation tasks, there is a rich opportunity for designing interactions with intelligent agents that not only generate stimuli but also encourage better ideation behaviours. Triggering timely reminders, suggesting new idea directions, preempting fixation, varying the abstraction of the visual outcomes, and facilitating visual analogical reasoning are all promising directions for future work.

5.4 Limitations

We acknowledge several limitations in our study. First, at the time of writing, generative AI tools are still nascent technologies. Interaction paradigms are emerging, and users are still learning to leverage their potential. As such, our results describe a picture of somewhat naïve use of these tools. It will be interesting to see how these results evolve as users become accustomed to generative AI tools and incorporate them into their practice.

Next, in this study, we gave all groups of participants the same amount of time to complete the task, but we observed lower fluency in the GENAI condition. We note that when the experiment took place, the AI system did not produce results instantaneously, which potentially delayed participants in that condition. However, we also note that participants in the IMAGE SEARCH condition, who *did* get their results instantaneously, also exhibited lower fluency than the baseline. This could be due to the exposure to a large number of images with endless scrolling, which added another layer of decision making to pick the ideas that suit them. Therefore, in a real-world setting, it is important to consider the trade-off between spending time on the task (e.g. by sketching) vs. seeking inspiration (e.g. by interacting with a creativity-support tool).

We acknowledge that because we limited the task time to 20 minutes based on previous work, we restrict the scope of our insights to short-term usage of these tools in a rapid ideation task. In the real world, people may spend longer reflecting on the outputs of AI, and incubation time along with iteration of sketched ideas may produce results different from those of our experiment. Further, we screened participants for prior skills in visual design, but few had professional industry experience. Though our sample was balanced across conditions, we make no claims about how these effects might be affected by expertise. Therefore, with this study, we can only provide initial insights into how a novice designer might approach a design task, and to generalize these claims, we need further investigation.

In this study, we operationalised design fixation by looking for a restricted set of salient features from the example in participants' sketches. Though the choice of focusing on denotative elements of the design aims to facilitate operationalization, we acknowledge that there are also connotative aspects that were left outside the scope of our analysis, including art style, emotional expression, and cultural references. Finally, our study only evaluated the potential of generative AI tools for ideation support through the specific example of image generators in a visual ideation task. It remains to be seen how these effects translate to other modalities, such as text, video, audio, and music generation.

6 CONCLUSION

Through this study, we contribute empirical evidence to the discussion of the potential of generative AI to augment human creativity.

Our study revealed that using an AI image generator as a source of inspiration by novice designers led to higher design fixation on an initial example and lower fluency, variety, and originality of ideas compared to using a conventional image search or no inspiration support. We suggest that fixation can happen in how the brief and the example influence the prompt given to the AI system, how the system translates it into images, and how the images inspire participants' ideas. All of these offer rich opportunities for re-design. Our work suggests that, at least in the current context of AI tool usage, given a fixed amount of time for a visual ideation task, this time is better spent sketching than seeking inspiration through AI. Our work suggests that generative AI tools aimed at supporting co-ideation should not only focus on generating stimuli but also on encouraging more effective ideation behaviours. We believe that incorporating well-thought-out methods and strategies into user practices and developing generative AI tools that can reduce common obstacles, such as design fixation and other creativity blockers, can maximise its potential to speed up the creative process and improve the quality of innovative design output.

ACKNOWLEDGMENTS

This research is supported by the Rowden White Scholarship and the Melbourne Research Scholarship offered by the University of Melbourne. We also would like to thank Christian Davey at the Melbourne Statistical Consulting Platform for their support.

REFERENCES

- [1] Leyla Alipour, Mohsen Faizi, Asghar Mohammad Moradi, and Gholamreza Akrami. 2018. A review of design fixation: research directions and key factors. *International Journal of Design Creativity and Innovation* 6, 2 (2018), 22–35. <https://doi.org/10.1080/21650349.2017.1320232>
- [2] Carina Andersson, Yvonne Eriksson, Lasse Frank, and Bill Nicholl. 2012. Design fixations among information design students: What has been seen cannot be unseen. In *DS 74: Proceedings of the 14th International Conference on Engineering & Product Design Education (E&PDE12) Design Education for Future Wellbeing*. Design Society, Antwerp, Belgium. https://www.designsociety.org/download-publication/33183/design_fixations_among_information_design_students_what_has_been_seen_cannot_be_unseen
- [3] Audi MediaCenter. 2022. Reinventing the wheel? “FelGAN” inspires new rim designs with AI | Audi MediaCenter. <https://www.audi-mediacycenter.com/en/press-releases/reinventing-the-wheel-felgan-inspires-new-rim-designs-with-ai-15097>
- [4] B.G. Bellows, J.F. Higgins, M.A. Smith, and R.J. Youmans. 2012. The Effects of Individual Differences in Working Memory Capacity and Design Environment on Design Fixation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 56 (9 2012), 1977–1981. <https://doi.org/10.1177/1071181312561293>
- [5] B.G. Bellows, J.F. Higgins, and R.J. Youmans. 2013. An individual differences approach to design fixation: Comparing laboratory and field research. In *Design, User Experience, and Usability. Design Philosophy, Methods, and Tools. DUXU 2013. Lecture Notes in Computer Science*. Springer, Berlin Heidelberg, 13–21. <https://doi.org/10.1007/978-3-642-39229-0>
- [6] Iouri Belski and Ianina Belski. 2015. Application of TRIZ in improving the creativity of engineering experts. *Procedia Engineering* 131 (2015), 792–797. <https://doi.org/10.1016/j.proeng.2015.12.379>
- [7] Virginia Braun and Victoria Clarke. 2022. *Thematic analysis: a practical guide*. SAGE Publications Inc., Thousand Oaks, California, United States.
- [8] Paul-Christian Bürkner. 2017. brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software* 80, 1 (2017), 1–28. <https://doi.org/10.18637/jss.v080.i01>
- [9] J. Cao, W. Zhao, and X. Guo. 2021. Utilizing EEG to Explore Design Fixation during Creative Idea Generation. *Computational Intelligence and Neuroscience* 2021 (2021). <https://doi.org/10.1155/2021/6619598>
- [10] C. Cardoso, P. Badke-Schaub, and A. Luz. 2009. Design fixation on non-verbal stimuli: The influence of simple vs rich pictorial information on design problem-solving. In *Proceedings of the ASME Design Engineering Technical Conference*. ASME, San Diego, California, USA., 995–1002. <https://doi.org/10.1115/DETC2009-86826>

- [11] Bob Carpenter, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus A Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. Stan: A probabilistic programming language. *Journal of statistical software* 76 (2017).
- [12] Hernan Casakin and Gabriela Goldschmidt. 1999. Expertise and the use of visual analogy: implications for design education. *Design Studies* 20, 2 (3 1999), 153–175. [https://doi.org/10.1016/S0142-694X\(98\)00032-5](https://doi.org/10.1016/S0142-694X(98)00032-5)
- [13] Peiyao Cheng, Ruth Mugge, and Jan P.L. Schoormans. 2014. A new strategy to reduce design fixation: Presenting partial photographs to designers. *Design Studies* 35, 4 (2014), 374–391. <https://doi.org/10.1016/J.DESTUD.2014.02.004>
- [14] Li-Yuan Chiou, Peng-Kai Hung, Rung-Huei Liang, and Chun-Teng Wang. 2023. Designing with AI: An Exploration of Co-Ideation with Image Generators. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. ACM, New York, NY, USA, 1941–1954. <https://doi.org/10.1145/3563657.3596001>
- [15] Evangelia G Chrysikou and Robert W Weisberg. 2005. Following the Wrong Footsteps: Fixation Effects of Pictorial Examples in a Design Problem-Solving Task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 5 (2005), 1134–1148.
- [16] John Joon Young Chung. 2022. Artistic User Expressions in AI-powered Creativity Support Tools. In *UIST 2022 Adjunct - Adjunct Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. Association for Computing Machinery, Inc, New York, NY, USA, 1–4. <https://doi.org/10.1145/3526114.3558531>
- [17] Nathan Crilly. 2015. Fixation and creativity in concept development: The attitudes and practices of expert designers. *Design Studies* 38 (5 2015), 54–91. <https://doi.org/10.1016/J.DESTUD.2015.01.002>
- [18] Nathan Crilly and Carlos Cardoso. 2017. Where next for research on fixation, inspiration and creativity in design? *Design Studies* 50 (5 2017), 1–38. <https://doi.org/10.1016/J.DESTUD.2017.02.001>
- [19] N. Davis, S. Siddiqui, P. Karimi, M.L. Maher, and K. Grace. 2019. Creative sketching partner: A co-creative sketching tool to inspire design creativity. In *Proceedings of the 10th International Conference on Computational Creativity, ICCO 2019*. Association for Computational Creativity, North Carolina, 358–359.
- [20] Edward de Bono. 2008. *Six Thinking Hats* (revised edition ed.). Penguin, United Kingdom.
- [21] Saurabh Deo, Aimane Blej, Senni Kirjavainen, and Katja Holttä-Otto. 2021. Idea Generation Mechanisms: Comparing the Influence of Classification, Combination, Building on Others, and Stimulation Mechanisms on Ideation Effectiveness. *Journal of Mechanical Design, Transactions of the ASME* 143, 12 (12 2021), 1 – 46. <https://doi.org/10.1115/1.4051239/1109505>
- [22] Tojin T. Eapen, Daniel J. Finkenstadt, Josh Folk, and Lokesh Venkataswamy. 2023. How Generative AI Can Augment Human Creativity. <https://hbr.org/2023/07/how-generative-ai-can-augment-human-creativity>
- [23] Lorenzo Fiorineschi and Federico Rotini. 2023. Uses of the novelty metrics proposed by Shah et al.: what emerges from the literature? *Design Science* 9 (5 2023), e11. <https://doi.org/10.1017/DSJ.2023.9>
- [24] John Gero, A T Purcell, J S Gero, H M Edwards, and E Matka. 1994. Design fixation and intelligent design aids. In *Artificial Intelligence in Design '94*. Springer, Dordrecht, 483–495. <https://doi.org/10.1007/978-94-011-0928-4>
- [25] Joy Paul Guilford. 1956. The structure of intellect. *Psychological bulletin* 53, 4 (1956), 267–293. <https://doi.org/10.1037/h0040755>
- [26] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology* 52, C (1 1988), 139–183. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- [27] Marius Hoggenmueller, Maria Luce Lupetti, and Willem Van Der Maden. 2023. Creative AI for HRI Design Explorations. In *HRI '23: Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. Association for Computing Machinery, New York, NY, USA, 40–50. <https://doi.org/10.1145/3568294.3580035>
- [28] Thomas Howard, Anja Maier, Balder Onarheim, and Morten. Friis-Olivarius. 2013. Overcoming design fixation through education and creativity methods. In *Proceedings of the International Conference on Engineering Design, ICED, Vol. 7 DS75-07*. The Design Society, Seoul Korea, 139–148. https://www.designsociety.org/download-publication/34578/overcoming_design_fixation_through_education_and_creativity_methods
- [29] Angel Hsing Chi Hwang. 2022. Too Late to be Creative? AI-Empowered Tools in Creative Processes. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–9. <https://doi.org/10.1145/3491101.3503549>
- [30] David G. Jansson and Steven M. Smith. 1991. Design fixation. *Design Studies* 12, 1 (1 1991), 3–11. [https://doi.org/10.1016/0142-694X\(91\)90003-F](https://doi.org/10.1016/0142-694X(91)90003-F)
- [31] John Joon, Young Chung, Shiqing He, and Eytan Adar. 2021. The Intersection of Users, Roles, Interactions, and Technologies in Creativity Support Tools; The Intersection of Users, Roles, Interactions, and Technologies in Creativity Support Tools. In *Designing Interactive Systems Conference 2021*. ACM, New York, NY, USA, 1817 – 1833. <https://doi.org/10.1145/3461778>
- [32] P. Karimi, J. Rezwana, S. Siddiqui, M.L. Maher, and N. Dehbozorgi. 2020. Creative sketching partner: An analysis of human-AI co-creativity. In *International Conference on Intelligent User Interfaces, Proceedings IUI*. Association for Computing Machinery, New York, NY, USA, 221–230. <https://doi.org/10.1145/3377325.3377522>
- [33] Matthew Kay, Gregory L. Nelson, and Eric B. Heckler. 2016. Researcher-Centered Design of Statistics: Why Bayesian Statistics Better Fit the Culture and Incentives of HCI. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (CHI '16). Association for Computing Machinery, New York, NY, USA, 4521–4532. <https://doi.org/10.1145/2858036.2858465>
- [34] Jieun Kim, Hokyoun Ryu, and Hyeonah Kim. 2013. To Be Biased or Not to Be: Choosing between Design Fixation and Design Intentionality. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13)*. Association for Computing Machinery, New York, NY, USA, 349–354. <https://doi.org/10.1145/2468356.2468418>
- [35] Janin Koch, Nicolas Taffin, Michel Beaudouin-Lafon, Markku Laine, Andrés Lucero, and Wendy E. MacKay. 2020. ImageSense: An Intelligent Collaborative Ideation Tool to Support Diverse Human-Computer Partnerships. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (5 2020), 27. <https://doi.org/10.1145/3392850>
- [36] Aaron Kozbelt and Yana Durmysheva. 2007. Understanding Creativity Judgments of Invented Alien Creatures: The Roles of Invariants and Other Predictors*. *The Journal of Creative Behavior* 41, 4 (12 2007), 223–248. <https://doi.org/10.1002/J.2162-6057.2007.TB01072.X>
- [37] Bart Lamiroy and Emmanuelle Potier. 2022. Lamuse: Leveraging Artificial Intelligence for Sparking Inspiration. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 13221 LNCS (2022), 148–161. https://doi.org/10.1007/978-3-031-03789-4_10/FIGURES/6
- [38] Keelin Leahy, Shanna R. Daly, Seda McKilligan, and Colleen M. Seifert. 2020. Design fixation from initial examples: Provided versus self-generated ideas. *Journal of Mechanical Design, Transactions of the ASME* 142, 10 (10 2020), 101402. <https://doi.org/10.1115/1.4046446/1074761>
- [39] Makayla Lewis. 2023. AIxArtist: A First-Person Tale of Interacting with Artificial Intelligence to Escape Creative Block.
- [40] J S Linsey, I Tseng, K Fu, J Cagan, K L Wood, and C Schunn. 2010. A Study of Design Fixation, Its Mitigation and Perception in Engineering Design Faculty. *Journal of Mechanical Design (JMD)* 132, 4 (4 2010), 041003. <https://doi.org/10.1115/1.4001110>
- [41] Andrés Lucero. 2012. Framing, Aligning, Paradoxing, Abstracting, and Directing: How Design Mood Boards Work. In *Proceedings of the Designing Interactive Systems Conference*. Association for Computing Machinery, New York, NY, USA, 438–447.
- [42] Abraham S. Luchins. 1942. Mechanization in problem solving: The effect of Einstellung. *Psychological Monographs* 54, 6 (1942), i–95. <https://doi.org/10.1037/h0093502>
- [43] Marian Mazzone and Ahmed Elgammal. 2019. Art, Creativity, and the Potential of Artificial Intelligence. *Arts 2019, Vol. 8, Page 26* 8, 1 (2 2019), 26. <https://doi.org/10.3390/ARTS8010026>
- [44] Richard McElreath. 2020. *Statistical rethinking: A Bayesian course with examples in R and Stan (2e)*. Chapman and Hall/CRC.
- [45] Diana P. Moreno, Lucienne T. Blessing, Maria C. Yang, Alberto A. Hernández, and Kristin L. Wood. 2016. Overcoming design fixation: Design by analogy studies and nonintuitive findings. *AI EDAM* 30, 2 (5 2016), 185–199. <https://doi.org/10.1017/S0890060416000068>
- [46] Maria Adriana Neroni and Nathan Crilly. 2021. How to Guard Against Fixation? Demonstrating Individual Vulnerability is More Effective Than Warning of General Risk. *The Journal of Creative Behavior* 55, 2 (6 2021), 447–463. <https://doi.org/10.1002/JOCB.465>
- [47] A Terry Purcell and John S Gero. 1996. Design and other types of fixation. *Design studies* 17 (1996), 363–383. [https://doi.org/10.1016/S0142-694X\(96\)00023-3](https://doi.org/10.1016/S0142-694X(96)00023-3)
- [48] Janet Rafner, Blanka Zana, Peter Dalsgaard, Michael Mose Biskjaer, and Jacob Sherson. 2023. Picture This: AI-Assisted Image Generation as a Resource for Problem Construction in Creative Problem-Solving. In *Proceedings of the 15th Conference on Creativity and Cognition*. Association for Computing Machinery (ACM), New York, NY, USA, 262–268. <https://doi.org/10.1145/3591196.3596823>
- [49] Christian Remy, Lindsay Macdonald Vermeulen, Jonas Frich, Michael Mose Biskjaer, and Peter Dalsgaard. 2020. Evaluating creativity support tools in HCI research. In *DIS 2020 - Proceedings of the 2020 ACM Designing Interactive Systems Conference*. Association for Computing Machinery, Inc, New York, NY, USA, 457–476. <https://doi.org/10.1145/3357236.3395474>
- [50] Lori Rosenkopf and Atul Nerkar. 2001. Beyond local search: boundary-spanning, exploration, and impact in the optical disk industry. *Strategic Management Journal* 22, 4 (4 2001), 287–306. <https://doi.org/10.1002/SMJ.160>
- [51] Othman Sbai, Mohamed Elhoseiny, Antoine Bordes, Yann LeCun, and Camille Couprie. 2019. DesIGN: Design inspiration from generative networks. *Computer Vision – ECCV 2018 Workshops, ECCV 2018. Lecture Notes in Computer Science*(1) 11131 (2019), 0–0. <https://doi.org/10.1007/978-3-030-11015-4>
- [52] Martin Schmettow. 2021. *New statistics for design researchers*. Springer.

- [53] Jami J. Shah, Noe Vargas-Hernandez, and Steve M. Smith. 2003. Metrics for measuring ideation effectiveness. *Design Studies* 24, 2 (3 2003), 111–134. [https://doi.org/10.1016/S0142-694X\(02\)00034-0](https://doi.org/10.1016/S0142-694X(02)00034-0)
- [54] Joon Gi Shin, Janin Koch, Andrés Lucero, Peter Dalsgaard, and Wendy E. MacKay. 2023. Integrating AI in Human-Human Collaborative Ideation. In *Conference on Human Factors in Computing Systems - Proceedings*. Association for Computing Machinery, New York, NY, USA, 1–5. <https://doi.org/10.1145/3544549.3573802>
- [55] Dilpreet Singh, Nina Rajcic, Simon Colton, and Jon McCormack. 2019. Camera obscurer: Generative art for design inspiration. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 11453 LNCS (2019), 51–68. https://doi.org/10.1007/978-3-030-16667-0_4/TABLES/3
- [56] Ut Na Sio and Thomas C. Ormerod. 2009. Does Incubation Enhance Problem Solving? A Meta-Analytic Review. *Psychological Bulletin* 135, 1 (1 2009), 94–120. <https://doi.org/10.1037/A0014212>
- [57] Melissa A.B. Smith, Robert J. Youmans, Brooke G. Bellows, and Matthew S. Peterson. 2013. Shifting the focus: An objective look at design fixation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 8012 LNCS, PART 1 (2013), 144–151. https://doi.org/10.1007/978-3-642-39229-0_17/COVER
- [58] Steven M. Smith and Julie Linsey. 2011. A Three-Pronged Approach for Overcoming Design Fixation. *The Journal of Creative Behavior* 45, 2 (6 2011), 83–91. <https://doi.org/10.1002/J.2162-6057.2011.TB01087.X>
- [59] Steven M Smith, Thomas B Ward, and Jays Schumacher. 1993. Constraining effects of examples a creative generation task. In *Memory & Cognition* 21, 6 (1993), 837–845.
- [60] Gareth Terry and Nikki Hayfield. 2021. *Essentials of Thematic Analysis*. American Psychological Association, Washington, DC, USA. <https://uwe-repository.worktribe.com/output/7240960>
- [61] L.A. Vasconcelos, M.A. Neroni, C. Cardoso, and N. Crilly. 2018. Idea representation and elaboration in design inspiration and fixation experiments. *International Journal of Design Creativity and Innovation* 6, 1-2 (2018), 93–113. <https://doi.org/10.1080/21650349.2017.1362360>
- [62] L.A. Vasconcelos, M.A. Neroni, and N. Crilly. 2016. Fluency results in design fixation experiments: An additional explanation. In *4th International Conference on Design Creativity, ICDC 2016*. The Design Society, Atlanta, GA, USA, 1–8.
- [63] Luis A Vasconcelos, Carlos C Cardoso, Chih-Chun Chen, and Nathan Crilly. 2017. Inspiration and Fixation: The Influences of Example Designs and System Properties in Idea Generation. *Journal of Mechanical Design* 139, 3 (2017), 031101. <https://doi.org/10.1115/1.4035540>
- [64] Luis A. Vasconcelos and Nathan Crilly. 2016. Inspiration and fixation: Questions, methods, findings, and challenges. *Design Studies* 42 (1 2016), 1–32. <https://doi.org/10.1016/J.DESTUD.2015.11.001>
- [65] Aki Vehtari, Andrew Gelman, Daniel Simpson, Bob Carpenter, and Paul-Christian Bürkner. 2021. Rank-Normalization, Folding, and Localization: An Improved \hat{R} for Assessing Convergence of MCMC (with Discussion). *Bayesian Analysis* 16, 2 (2021), 667 – 718. <https://doi.org/10.1214/20-BA1221>
- [66] Mathias Peter Verheijden and Mathias Funk. 2023. Collaborative Diffusion: Boosting Designerly Co-Creation with Generative AI. In *Conference on Human Factors in Computing Systems - Proceedings*. Association for Computing Machinery, New York, NY, USA, 1–8. <https://doi.org/10.1145/3544549.3585680>
- [67] Vimal Viswanathan and Julie Linsey. 2012. Design Fixation in Physical Modeling: An Investigation on the Role of Sunk Cost. *Proceedings of the ASME Design Engineering Technical Conference* 9 (6 2012), 119–130. <https://doi.org/10.1115/DETC2011-47862>
- [68] V. Viswanathan, M. Tomko, and J. Linsey. 2016. A study on the effects of example familiarity and modality on design fixation. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AIEDAM* 30, 2 (2016), 171–184. <https://doi.org/10.1017/S0890060416000056>
- [69] Eric-Jan Wagenmakers, Ruud Wetzels, Denny Borsboom, and Han LJ Van Der Maas. 2011. Why psychologists must change the way they analyze their data: the case of psi: comment on Bem (2011). (2011).
- [70] T. B. Ward. 1994. Structured Imagination: the Role of Category Structure in Exemplar Generation. *Cognitive Psychology* 27, 1 (8 1994), 1–40. <https://doi.org/10.1006/COGP.1994.1010>
- [71] Blake Williford, Samantha Ray, Jung In Koh, Josh Cherian, Paul Tael, and Tracy Hammond. 2023. Exploring Creativity Support for Concept Art Ideation. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–7. <https://doi.org/10.1145/3544549.3585684>
- [72] Roosa Wingström, Johanna Hautala, and Riina Lundman. 2022. Redefining Creativity in the Era of AI? Perspectives of Computer Scientists and New Media Artists. *Creativity Research Journal* (2022), 1–17. <https://doi.org/10.1080/10400419.2022.2107850>
- [73] Robert J. Youmans. 2011. Design Fixation in the Wild: Design Environments and Their Influence on Fixation. *The Journal of Creative Behavior* 45, 2 (6 2011), 101–107. <https://doi.org/10.1002/J.2162-6057.2011.TB01089.X>
- [74] Robert J. Youmans. 2011. The effects of physical prototyping and group work on the reduction of design fixation. *Design Studies* 32, 2 (3 2011), 115–138. <https://doi.org/10.1016/J.DESTUD.2010.08.001>
- [75] Robert J. Youmans and Tomasz Arciszewski. 2014. Design Fixation: A Cloak of Many Colors. In *Design Computing and Cognition '12*. Springer, Dordrecht, 115–129. <https://doi.org/10.1007/978-94-017-9112-0>
- [76] Robert J. Youmans and Thomaz Arciszewski. 2014. Design fixation: Classifications and modern methods of prevention. *AI EDAM* 28, 2 (2014), 129–137. <https://doi.org/10.1017/S0890060414000043>

APPENDIX

A THE MEAN SCORES AND STANDARD ERRORS OF THE NASA TASK LOAD INDEX (NASA-TLX) SCALES

Table A.1: The mean scores and standard error for each NASA TLX scale (mental demand, physical demand, temporal demand, performance demand, effort demand, frustration demand) in the three conditions: No support, Image Search and GenAI

NASA TLX	Mean scores & Std. Error		
	No Support	Image Search	GenAI
Mental Demand	4.1 (0.3)	4.0 (0.4)	4.2 (0.4)
Physical Demand	2.2 (0.3)	2.6 (0.3)	2.6 (0.3)
Temporal Demand	4.3 (0.4)	3.7 (0.4)	4.4 (0.4)
Performance	4.1 (0.3)	5.2 (0.3)	4.7 (0.3)
Effort	4.5 (0.3)	4.0 (0.4)	4.0 (0.2)
Frustration	3.1 (0.4)	2.0 (0.3)	2.2 (0.3)