

Text-to-Image AI as a Catalyst for Semantic Convergence in Creative Collaborations

Citation for published version (APA):

Lin, P.-Y., Andersen, H. K. G., Schmidt, R. P. W., Schoenmakers, S., Hofmeyer, H., Pauwels, P., & IJsselsteijn, W. A. (2024). Text-to-Image AI as a Catalyst for Semantic Convergence in Creative Collaborations. In A. Vallgård, L. Jönsson, J. Fritsch, S. Fdili Alaoui, & C. A. Le Dantec (Eds.), *DIS '24: Proceedings of the 2024 ACM Designing Interactive Systems Conference* (pp. 2753-2767). ACM.
<https://doi.org/10.1145/3643834.3661543>

Document license:

CC BY-NC-SA

DOI:

[10.1145/3643834.3661543](https://doi.org/10.1145/3643834.3661543)

Document status and date:

Published: 01/07/2024

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.



Text-to-Image AI as a Catalyst for Semantic Convergence in Creative Collaborations

Pei-Ying Lin
p.y.lin@tue.nl
Eindhoven University of Technology
Eindhoven, Netherlands

Kristina Andersen
h.k.g.andersen@tue.nl
Eindhoven University of Technology
Eindhoven, Netherlands

Ralf Schmidt
r.p.w.schmidt@tue.nl
Eindhoven University of Technology
Eindhoven, Netherlands

Sanne Schoenmakers
s.schoenmakers@tue.nl
Eindhoven University of Technology
Eindhoven, Netherlands

Hèrm Hofmeyer
h.hofmeyer@tue.nl
Eindhoven University of Technology
Eindhoven, Netherlands

Pieter Pauwels
p.pauwels@tue.nl
Eindhoven University of Technology
Eindhoven, Netherlands

Wijnand IJsselsteijn
w.a.ijsselsteijn@tue.nl
Eindhoven University of Technology
Eindhoven, Netherlands

ABSTRACT

Design ideation requires both creative and divergent thinking as well as collaboration and exchange to bring forth new insights and design possibilities. In this paper, we investigate such a collaboration with the use of emerging generative AI tools, and we explore how they may help in achieving agreement and semantic convergence between designers. We conducted workshops using Text-to-Image AI with design students and found that in addition to known advantages of brainstorming, the text prompts required by the AI system engaged students in verbal articulations that aided their design process. In collaborative contexts, the generated images shifted the centre of attention for team members to reach convergence through sharing their interpretations. We conclude that text-to-image AI can be beneficial for students' design processes as a catalyst for brainstorming rather than a tool for generating design imagery and presentations.

CCS CONCEPTS

• **Human-centered computing** → **Collaborative and social computing**; **Collaborative interaction**; *Interaction design theory, concepts and paradigms*; *Empirical studies in collaborative and social computing*; *HCI theory, concepts and models*.

KEYWORDS

Research through design, digital craftsmanship, semantic convergence, ideation, human-AI collaboration, co-creation

ACM Reference Format:

Pei-Ying Lin, Kristina Andersen, Ralf Schmidt, Sanne Schoenmakers, Hèrm Hofmeyer, Pieter Pauwels, and Wijnand IJsselsteijn. 2024. Text-to-Image



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike International 4.0 License.

DIS '24, July 01–05, 2024, IT University of Copenhagen, Denmark
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0583-0/24/07
<https://doi.org/10.1145/3643834.3661543>

AI as a Catalyst for Semantic Convergence in Creative Collaborations. In *Designing Interactive Systems Conference (DIS '24)*, July 01–05, 2024, IT University of Copenhagen, Denmark. ACM, New York, NY, USA, 15 pages. <https://doi.org/10.1145/3643834.3661543>

1 INTRODUCTION AND RELATED WORKS

Design is a vast, well-studied domain with various subdomains such as Engineering Design, Industrial Design, Architecture, and User-Experience Design. In these subdomains, although their design processes differ from one another, creative ideas are key. Design can be seen as an iterative and organic process in which problems are constantly redefined, and multiple routes to possible outcomes are being explored. In Research through Design [23, 38], the creative process is focused on discovery through making [24, 29] rather than optimal solutions for rationalised problems, while in design for engineering the main focus is on problem-solving, and product development engages with the user, function, production and materials.

Most design processes do not take place as solo work but in teams and collaborations. This makes the design process complex as everyone has a different mindset (see Fig. 1) to go through an often free-flowing, organic process [46] of research, ideation, evaluation, decision-making, prototyping, and production [26–28, 30]. In such a collaborative design process, communication is critical, making creative design processes fundamentally social interactions [8, 13].

A large part of this collaborative design process consists of designers explaining what is on their mind to their fellow designers, in effect engaging in negotiations of an emerging shared semantic space. Different methods to convey non-textual complex information exists, such as brainstorming [3, 45], sketching [20, 21, 44], diagrams [11, 19, 43], and design cards [1]. The intricacy of such communication includes non-verbal, embodied knowledge, particularly when the design includes tangible elements like material properties or choreographed movements, and these elements often play a crucial role in guiding design decisions [4, 52, 58]. It has

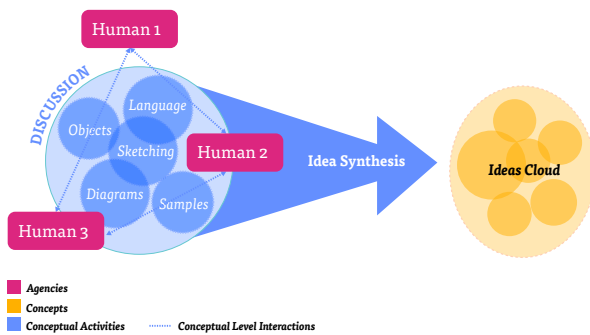


Figure 1: Diagram of a team discussion: The human collaborators (ruby pink: agencies) are brought together where they discuss ideas in a conceptual space (blue: conceptual activities). Within this conceptual space, different ways of communication take place, such as sketching, diagrams, language, and material samples. The conceptual interactions (blue dashed lines: conceptual level interactions) happen between the human collaborators. Through this process, ideas are generated, which creates an ‘ideas cloud’ (yellow: ideas cloud).

been proven that synchronising the conceptual and semantic understanding among collaborators [51] is crucial for a productive design process.

To make a collaboration flow freely, design methods such as sketching, diagrams, and mood boards are used to create lightweight, fast, abstract communication. As a trade-off, detailed information is often absent. The recent emergence of text-to-image AI in the public domain has introduced the possibility of generating images filled with details from relatively simple verbal prompts. Text-to-image AI models work by taking text prompts and delivering imagery through language-based models. As text-to-image AI tools such as Dall-E [47], Stable Diffusion [49], Midjourney [41], and Adobe Firefly [33] become integrated into the creative industry, they are increasingly used in the design process for ideation, communication, and production - often to flesh out ideas with details that would otherwise be time-consuming to produce [2, 12, 31, 32, 42, 50, 54]. While the possibilities of using text-to-image AI for idea generation is well researched [12, 37, 39, 40], less attention has been paid to how this might change the nature of team collaborations in design.

In this paper, we look at how designers might be able to adapt text-to-image AI in the ideation process for semantic convergence, with a particular focus on novice designers in an educational context. Previous studies have found that experienced designers are capable of ‘forming abstract conceptualisation’ and ‘accessing larger chunks of information’ [17], whereas novice designers tend to narrow their options and focus on non-actionable, or more mundane directions for prototyping, or even sometimes fail to come to any result [5, 14, 17, 36].

We propose that text-to-image AI could be used to help facilitate the emergence and convergence of ideas through the generation of detailed images making use of the latent semantic-image space. In addition to this, text-to-image AI may also allow novice drawers

to create explanatory imagery for discussion and brainstorms. For instance, when entering a text prompt as simple as ‘a face’ to a text-to-image AI model, it generates a female face with shades of skin tones, reflections in pupils, textured background, shapes of collar indicating the clothing style and a particular atmosphere. Such detail of the image could provide inspiration and material for the designer to work on [7], but at the same time provides ambiguity [25] that invites interpretations. At the same time, we note that this example also clearly illustrates the deep problems inherent to these systems in terms of how they may enforce cultural bias and creative assumptions [53].

In questioning how text-to-image AI might fuel the ideation process, we look to broader perspectives on the position AI holds in design. These perspectives include text-to-image AI as a tool or facilitator of a creative process [12, 37, 54], AI as a collaborator [48], and post-humanism AI-human-material co-creation [58]. Each of these has a different framing for the position on interaction between AI and humans.

In this paper, we aim to begin the unpacking of how a creative process may be integrated with text-to-image AI, identify the positioning of text-to-image AI in such a process, and explore the possible engagements and outcomes. We are looking particularly at text-to-image AI for it provides a connection between text and image generations, other types of generative AI are out of the scope of this paper.

We speculate that a text-to-image AI engagement could unfold as shown in Figure 2: When a text-to-image AI joins the collaboration as a tool, it executes the intention of the human participants. When a text-to-image AI joins the collaboration as a collaborator, it engages in generating ideas. In the context of post-humanism where AI, humans, and materials are seen together as an assemblage [18], the text-to-image AI might occupy the space that connects and facilitates between the humans and the material, that AI and the other parties form a constituency [58] which influences the final design goal.

We hypothesise that the text-to-image AI collaborations as non-human participants would bring drastic change to the creative process. In this paper, we conceptualise this by studying design ideation processes in the context of students designing ‘wearables’ within a semester-long project assignment in order to observe emerging patterns of creative collaboration with the integration of text-to-image AI. The assignment challenges students to design something related to ‘wearables’, which are tangible and physical, and engage real-life materials. This criterion demands that students not see results generated by text-to-image AI as the final product but as a process within a longer and larger creative process. This allows us to observe the nuances of text-to-image AI engagement within the creative process.

We have identified three emerging traits that are beneficial to the ideation design process for inexperienced designers with text-to-image AI: *Enhancing verbal articulation* during prompt formulation, *semantic convergence between collaborators* as the result of interpreting the ambiguous generated images, and the formation of a *secondary discussion space* that redirects the social tension happening amongst the collaborators towards an agreement into constructively building on each other’s interpretations of what the generated image of text-to-image AI had proposed. We suggest that

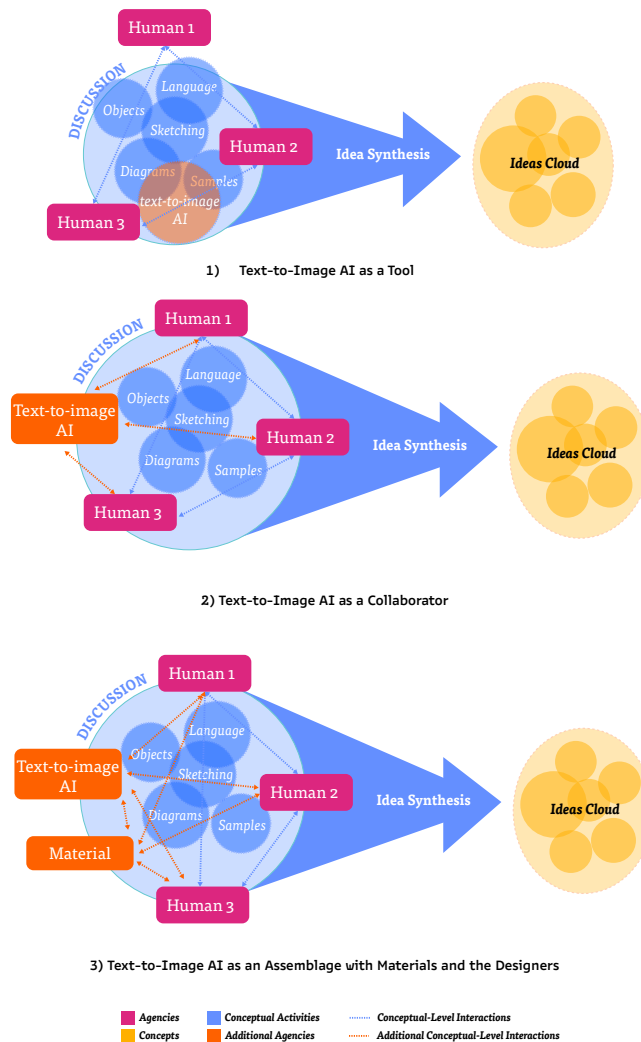


Figure 2: Diagrams of possible forms of how text-to-image AI can occupy different positions in team discussion. 1) Based on the basic model of team discussion, the text-to-image AI joins as a tool that is sub-ordinate to the discussion and contributes as merely a tool for communication that conveys the human participants' ideas. Text-to-image AI does not interact conceptually with human agencies. Instead, the text-to-image AI generates what is asked by the human agencies. 2) When text-to-image AI join the discussion as a collaborator, it gains agency. There are conceptual-level interactions between the AI and the human agencies. 3) When using the non-human framework where the material is also considered an agency in the assemblage, the material, text-to-image AI, and humans will have conceptual-level interactions with each other that drive the idea synthesis.

researchers and educators who work with machine learning models for creative collaborations should look not only at the function the machine learning models provide but also at the epistemological exercise machine learning models induce.

2 METHODS

2.1 Context

The study was conducted as a short workshop in a semester-long project-oriented course in an industrial design department. Some

students work on projects in teams of 3-4 students, and some students work on the project individually, depending on their years. Our study (text-to-image AI workshop) took place towards the end of the first month of the course while students were still in an ideation phase. The study was framed as an opportunity to use text-to-image AI for 30 minutes in support of their ideation. The students were free to choose to join, and the participation did not count towards the course credit. The majority of the course students have signed up to participate, but not all. The course has a duration of 4 months, and the students are asked to produce a project

related to wearable or soft materials. The students had to find their own project context and assign their own design briefs, making the design assignment very open-ended at the beginning of the course. The course is mandatory for undergraduate and graduate students in the industrial design department. The structure of the course comprises lectures, group discussions, production, coaching sessions, a mid-term presentation, and a final-term presentation. A coach is assigned to each project.

2.2 Participants

The workshop was conducted twice with different groups of students in two semesters until repeating patterns were observed. In total, fourteen projects were invited to participate in the text-to-image ideation workshop, each project had a 30-minute session. The workshops were conducted in two formats: team sessions and individual sessions. There were 5 sessions of team projects, which consisted of 18 student participants in their 2nd-year and 3rd-year of bachelor study, or their first year of master study. The remaining 9 sessions were individual projects by 3rd-year bachelor students and 2nd-year master students. The total number of participants was 27.

2.3 Setup of the workshop

The workshop lasted 30 minutes and consisted of a 5-minute pre-workshop questionnaire and a 25-minute interactive session. Participants were asked to fill out a post-workshop questionnaire in their own time. Before the workshop, they were asked to submit a short text description of their project and sign a consent form. The interactive session was conducted in a room with a computer connected to a large screen showing a text-to-image AI interface. We used Midjourney version 3 and version 4 on the Discord [34] platform between September 2022 to March 2023. Discord platform was used because Midjourney can only be accessed as a chatbot on the Discord platform, either through the official Midjourney channel, personal message, or channels that have the Midjourney chatbot. Up until Midjourney versions 3 and 4, the text prompts are sent to the Midjourney chatbot on Discord, where the Midjourney chatbot will reply with its generated images as messages. The workshop was centred on verbalising their project to communicate with the text-to-image AI model, a task also known as *prompt engineering* [40]. The facilitator was present to assist with inputting text and handling the Midjourney interface. The participants would sit around the screen with the facilitator in front of the computer. The participants were asked to ‘think aloud’ throughout [6] and iterations and multiple prompts running in parallel were allowed. The images generated in the session were saved separately with timestamps. The screen was recorded along with a sound recording of the conversation in the room. A semi-structured interview was conducted at the end of the workshop. After the session, each participant filled out a post-workshop questionnaire where they were asked to describe their project again and reflect on the overall experience.

2.4 Analysis and Sample Size

To confirm the validity of the research, the workshops were conducted until repeating patterns were observed across semesters and

across both the bachelor student and the master student teams. A pre-workshop questionnaire was used to confirm whether the participant had previous experience with any text-to-image AI, along with questions to see if they have a clear vision for the project and have reached an agreement with other members of the team if they are in a team’s context. A semi-structured interview was conducted at the end of the workshop, while similar questions were asked again in the post-workshop questionnaire for data triangulation [22]. The post-workshop questionnaire asks questions regarding 1) if they have a clearer vision for the project, 2) if the Midjourney generated images could represent the project, 3) if there is anything within the image that is inspiring, 4) describe what happened and how they feel in the workshop in the beginning, during the workshop, and after the workshop, 5) did the vision of the project change after the workshop, 6) was there anything their collaborator and the workshop facilitator did that was inspiring, 7) other interesting thing that hasn’t been asked, 8) was Midjourney helpful? and why? 9) other remarks. Question 1 and 5 was used to compare with the pre-workshop questionnaire. Question 6 was used to check if the facilitator had unconsciously influenced the process. The result confirms that none of the participants found the facilitator inspiring or guiding the process. Thematic analysis was done mainly using the answers for questions 4, 5, 6, and 7 and the answers from the semi-structured interview.

A thematic analysis following the steps proposed by Braun & Clarke [9] was done using the results of the post-workshop questionnaire. Further analysis was done combining the workshop process along with screen recording and conversation transcripts according to the identified themes. The recordings were transcribed with timestamps. The interactions were analysed along with the screen recording and the AI-generated images. To evaluate the status of the project, we used the metrics of material, form, and function to reflect the status of the projects. The ‘form’ looks at whether the project has decided on its form, such as whether it will be a garment, a toolkit, or other possible setups. The ‘function’ looks at whether the project has clearly defined its context and usage. The metrics were used to measure the participants’ responses in the pre-workshop questionnaire about their vision of the project.

3 FINDINGS

Amongst the five teams, four arrived having already decided on the basic context of their design project. They were able to agree on three to four keywords or a specific form of garment to aim for according to the pre-workshop questionnaire. One team had extremely different descriptions, with one member thinking the project would be shoes while the others thought symbiotic was the main keyword. All teams were still unable to specify the exact form of the design. The nine individual projects had more variation in terms of progress, with four participants having decided on the material, form and function of their design project, and the other five participants still uncertain, which is reflected in the pre-workshop questionnaire.

The pre-workshop questionnaire shows that out of the 27 participants, only 11 had prior experience with text-to-image AI. The post-workshop questionnaire reveals that all of the 27 participants found the text-to-image AI workshop helpful. However, only two

of the participants found that the text-to-image AI was able to help them generate an image to represent their project, while the 25 others found that the text-to-image AI instead gave them more options and ideas to work from. The participants generally found that Midjourney, the text-to-image AI used, provided them with new ideas and inspiration on shapes, materials, and aesthetics.

From the post-workshop questionnaire, two participants from different teams mentioned that 'Midjourney helps them to see what others (teammates) see' and another two participants from two other teams mentioned that 'AI helps to align the group (team) vision'.

Thematic analysis [9] of the post-workshop questionnaire along with the semi-structured interview identifies two themes: 'idea alignment / understanding the others' and 'verbal articulation' of text prompts. These themes were taken into account to do further analysis of the recorded screen capture videos and discussion audio recordings.

3.1 Multiple Layers of Negotiations and Alignment

The interactive session started with participants trying to formulate the project in a prompt sentence. Midjourney delivers four images for each prompt per iteration. Once the grid of four images is generated, participants express their preferences: 'what I meant', 'intriguing', 'interesting' or 'nice', leading to the decision of which image to choose for further iterations. Because Midjourney allows requests to run in parallel and takes some time to generate images, some participants would try to enter different prompts before seeing the results of the first prompt. Participants in all sessions tried multiple different words and variations to get an image close enough to their vision. Additionally, many would rephrase the prompt after seeing the outcomes. In other words, while the initial input is a set of words, the participants tend to evaluate those words based on the images they generate. In that sense, the images are not just describing a verbal prompt, but the prompt itself is developed through image evaluation. Such recursive interaction with the system makes the topic shift, from what the project is about, to commenting on the generated images and whether the images fit with the expectations of the prompt. At the context level, the participants tend to discuss the implication of the elements in the image as inspiration for future project directions. When unexpected images appear, participants tend to see them as 'broadening possibilities' and 'inspiring', and they make use of some of these outcomes to revise their prompts and iterate along with the generated variations.

Furthermore, because the workshop took place in the context of a course that requires students to produce physical projects, the participants looked for useful elements of fabrication from the Text-to-Image AI. The images generated by Midjourney are two-dimensional, but often ambiguous with an implicit hint of materiality. In order to make the digital images useful for the design project, participants would enter into a dynamic process of negotiation and alignment. Such negotiations would happen at two different levels: between the participants and the prompt, and between the participants themselves which we will discuss in the following.

3.2 Idea Alignment between Human and the Text-to-Image AI – an Exercise in Verbal Articulation

The first level of negotiation is with the prompt that occurs between the AI and the participant both in the team and individual setting. Both participants in the team session and individual session expressed if the text-to-image AI has generated an image that aligns with their idea. Under the premise that the AI has been trained using a large amount of existing text-image annotations, the participants assume the generated image is a targeted interpretation of the text prompt instead of something random. If the result is too different from what the participants expect, they identify the word that potentially 'went wrong', and think of whether the 'wrong word' was used (prompt text has mistakes), or the AI made a mistake. The participants then started to substitute the words to look for 'inspiring results' that were not completely aligned with the vision but along the line of generated images and prompts.

One example of negotiation and alignment within prompt articulation can be seen in the following example, where a team was 'creating a special fashionable rain outfit that transforms in reaction to the weather' according to their initial description. The iterations of prompts are seen in Figure 3.

In this example, 'raincoat' becomes 'rainwear'; colour is added to see how it would fit into the concept; the idea of interaction with the environment is first described as 'reacting to the environment' and later becomes the more active 'transforming the environment'.

As a result of this process, the concept is moved from making an outfit for rain into something that is wearable during rain, and the relationship between the artefact and the environment is changed from the passive status of reacting *towards* the environment to the active status of *transforming* the environment.

The idea alignment was also observed in the individual context, where the sole participant tries to align their idea with the text-to-image AI. Out of 14 sessions, 10 sessions have had this phenomenon, both in the team settings and the individual settings. One selected example is an individual project of a final year bachelor student. Originally, the participant was exploring how recycled denim could be used for different contexts and how to make value from the recycled material (Figure 4). In the process, the participant started to explore the concept of high fashion as opposed to recycled waste. The first three prompts are somewhat vague and general, and the imagery was considered less interesting by the participant. From the fourth prompt, the concept of 'product' started to narrow down from garments - wearable - garment - statue - to installation. And from the fifth prompt, 'recycled textile/fibres' became more descriptive as 'denim' and 'blue woven', and 'cotton polyester blends'. For the participant, the significant moment came when they changed the word 'garment' to 'statue': 'Because I feel like a garment is maybe too limited and it comes back to a jacket or a dress, whereas a statue is more free form.' This is an example of participants adopting deliberately incorrect words in order to circumvent the more obvious outcomes of the AI system. Despite these more complex outcomes, the participant considered the results more actionable: 'It's cool because this could be made from recycled yarn, but it could also be made from old fabrics without recycling it.' It is an example of a participant trying to add details for the prompt to get

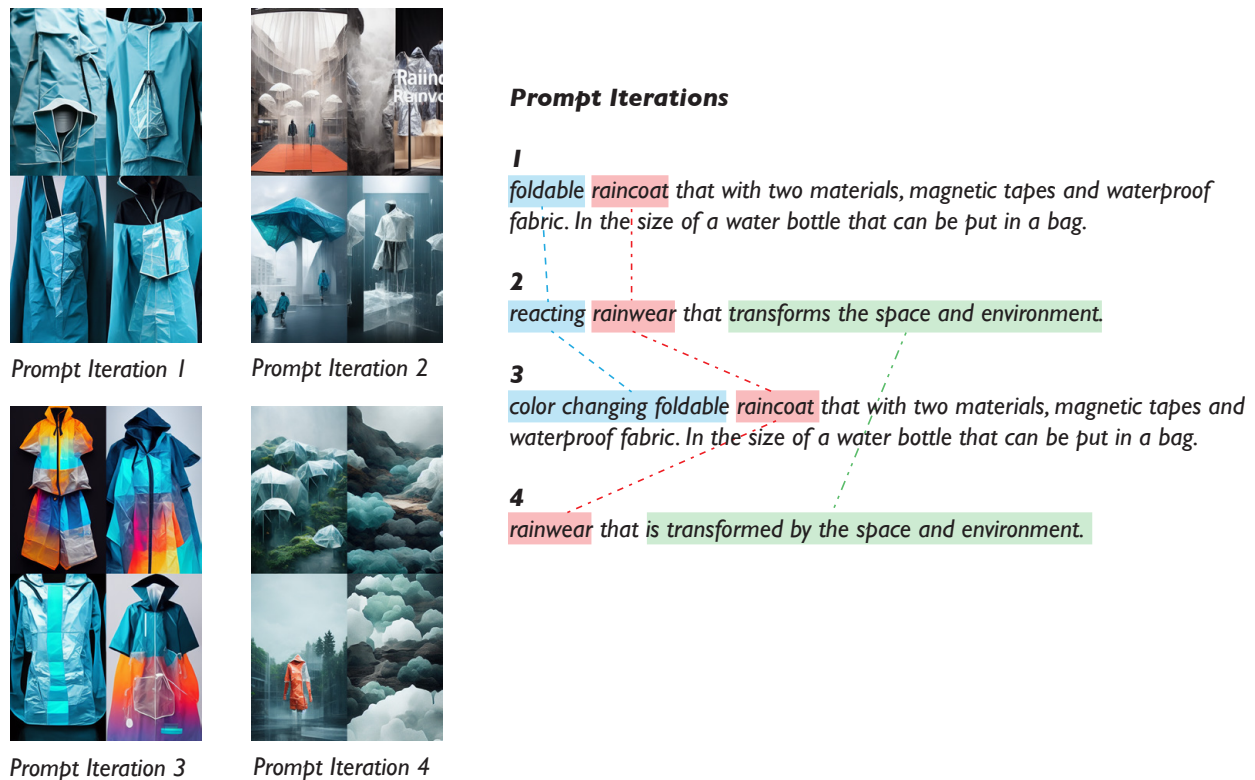


Figure 3: Prompt iterations of the team 'creating a special fashionable rain outfit that transforms in reaction to the weather'. Three flows of evolving concepts were identified and marked with blue, red, and green as the prompt iterated.

useful results from the AI. The participant later decided to develop material explorations in the style of the last generated images.

The process of trial-and-error between the prompt and the generated images is described by the participants as a process of 'what language to talk to with the AI'. In this process, they go through strategies such as *reframing* the sentence, *word substitution* ('garment' to 'installation'), and *becoming more explicit*. Participants say: 'I started to understand better how to use certain words and more conceptual terms to manipulate Midjourney into simulating more abstract/explorative imagery', and it is 'interesting to figure out what parts of the image were which part of the sentence. And how to see what part would make it go the "wrong way"'. Some participants treat the text-to-image AI as more than a tool: 'I think it's like another person and you tell your idea using words and the other person thinks about a picture in his or her mind which is unique to him/her' and 'Midjourney understood me'. In this way, the system is considered an external conversation partner, where the participants seek the 'correct way to communicate the idea with the AI'.

Despite the overwhelmingly positive feedback, all participants, including participants with individual projects and team projects, reported that the text-to-image exploration did not define the final

vision of their projects, but instead facilitated the possibilities within their idea to become clearer. The *executability* of the ideas provided by text-to-image AI was generally considered low. In other words, by broadening the scope of the vision, text-to-image AI suggested outcomes that would be much harder for the participants to execute.

3.3 Change of Social Interaction and Attention and Search for a Common Ground

The integration of generative AI in the team collaboration process has brought changes to the social interaction dynamic for which the generated images became the main focus of discussion with the purpose of alignment within the group. The social aspect within team collaboration settings is apparent where each team member has to interact with others through discussions and negotiations [13] which is different from individual workshop settings. As the imagery comes closer to a shared vision, participants stop evaluating whether the images are the 'correct' interpretations of the prompt and move towards dissecting and analysing them in detail. Since their projects were required to make something, the participants had to make a connection between the generated images and the physical world. The ambiguity of a generated image towards

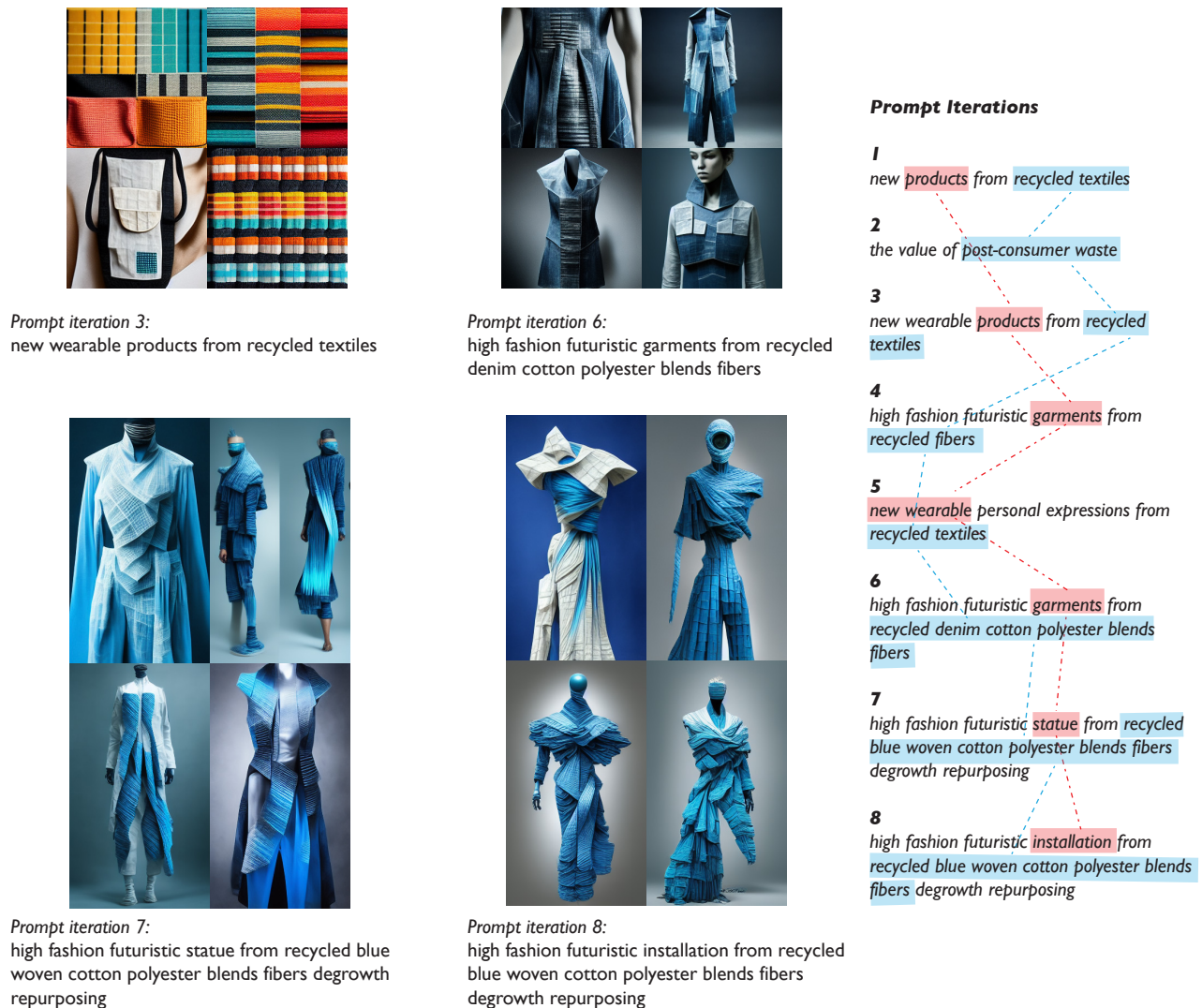


Figure 4: Prompt iterations of the recycle denim project. Two flows of evolving concepts were identified and marked with blue and red.

materiality creates a space for the participants to discuss their interpretations of what they see. This interaction was more obvious in the team sessions where multiple participants were looking at the same image while having different ideas in mind. Through discussing their interpretation of the same element in the image, the participants also go through a process of negotiation and alignment on the semantics of the image. They would comment on how each element looks like in their mind, and then worked towards *confirming each other's interpretations*, moving towards *an agreement on interpretations* that 'oh yes, I see what you mean. That is also nice.' Whenever partial agreement was reached, the conversations would move on to understanding the *contexts of the agreed interpretations*. This process is described by the participants as brainstorming.

An example of such brainstorming can be seen in a team that explores 'Creating a symbiotic relationship between wearables and its [sic] wearer. Wearable plants. Living vs. dead materials.' according to the team's original description. (Figure 5) The pre-workshop questionnaire shows that they had ideas for growing plants on clothes but were unclear about how. They also did not provide sketches when describing their ideas indicating that their ideas were still based on words and themes rather than being visual. Being certain of their words, they only tested 3 different prompts but spent a long time interpreting elements of the generated images.

In the interpretation process, as shown in Figure 6 and Figure 7, the participants spent a long time talking about what they see in the images, how each element represents the material, context and

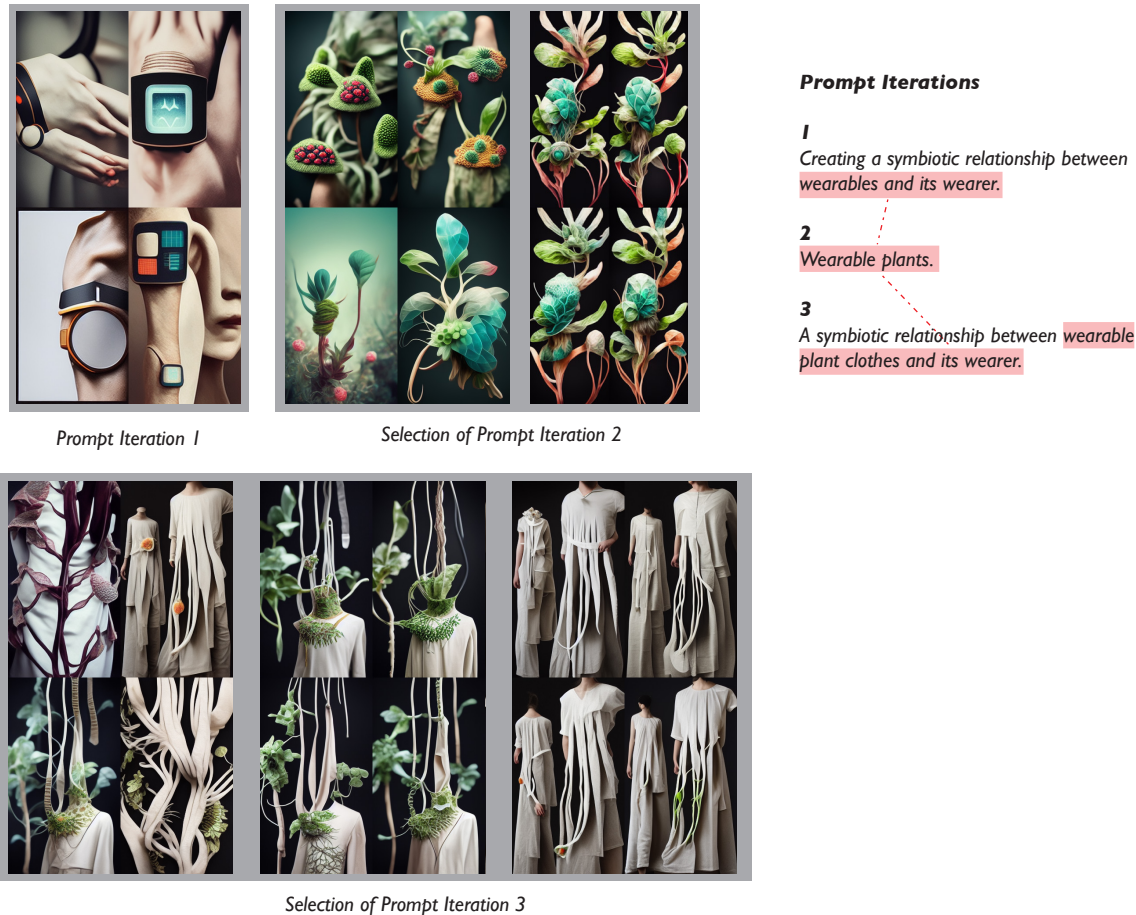


Figure 5: Prompt iterations of team ‘symbiotic relationship between human and plants’. The marked red texts illustrate how the concept of ‘wearable’ was explored in the text.

relationship. While one participant is speaking, the others would agree and add their own interpretations in a brainstorming fashion.

For teams who were having difficulties reaching a common vision, the text-to-image AI provided a neutral playground for interpretation, allowing them to see the potential for idea combinations. One team could not come to an agreement on their project before the workshop, with two members wanting to work on sneakers, and one not convinced. These participants decided to combine variations of their keywords in their prompts: ‘sneaker’, ‘modular systems’, ‘biomimicry’, and ‘plastic’. As a result, they identified interesting elements such as organically shaped plastic structures, which allied them to see the potential of integrating the idea of sneakers and organic shapes, as can be seen from the interview afterwards:

Q: Are you convinced by doing the sneaker?

A: ‘I do think that if you want to work with harder materials, sneakers are better than garments.’

Q: How do you feel about seeing all these images with your different ideas put together?

B: ‘I feel inspired.’

A: ‘It is really cool.’

A: ‘There are some things from each image that I really like and that I think we can combine.’

A: ‘I really like the first thing that we did with your idea (participant B) with the colored ones.’

A: ‘It was a kind of melty, flowy feeling to them. That’s really cool.’

B: ‘Yeah, it’s really cool.’

Their prompt iterates from ‘a colorful sneaker made out of plastic and fungi’, onto ‘modular sneakers based on biomimicry’, then ‘a colorful sneaker made out of plastic and fungi’. When Midjourney combined the idea of ‘biomimicry’ with ‘sneaker’ and generated shapes of non-functional sneakers, they had an in-depth discussion on shapes and materials as shown in Figures 8, 9 and 10. They started to agree with each other’s interpretation of the imagery



- A:** The thingies, the green things, it looks like it's already made of like a textile.
B: Yeah, yeah. It does kind of.
C: Yeah, yeah, I get what you mean. Though it does look a bit like pistils as well.
C: It looks like these small cactus plants or flowers.
A: It also kind of reminds me of little raspberries or something.
B: And you see like the top leaf at the top, it looks like sage to me.
A: Yeah, but they look a bit like they were actually pruned.
B: Yeah, exactly.
C: I still wonder how they were worn though, how they're supposed to. Because I see a bit of skin, I presume.
B: Yeah, I feel like it's a thigh, not a leg.
A: Like from the skirt sort of situation.
C: Oh!
B: I just saw that.
A: But I could also, you can see there's so many things. Kind of weird how much the little bit of skin does on that image.

Figure 6: Discussion of team 'symbiotic relationship between human and plants' with a generated image from the prompt 'Wearable plants'



- B:** You see like some twigs or bark in the post print.
A: Oh, yeah.
B: You also see like the greenery and the plant that just started to come out.
A: And the neckline.
C: It looks like it's sort of a host.
C: It's hosting a plant life.
A: Yeah.
A: This is really nice.
B: And you also see still the textiles.

Figure 7: Discussion of team 'symbiotic relationship between human and plants' with a generated image from the prompt 'A symbolic relationship between wearable plant clothes and its wearer'

without refusing each other's idea of 'sneaker' or 'biomimicry'. They also build on each other's interpretations such as in Figure 8 where Participant B talks about the shapes that remind him of 'the sand on the beach' and Participant A went further about 'those little waves you get on dunes and stuff'.

The emergence of convergent thinking was also confirmed in the post-workshop questionnaire. One participant mentioned that the text-to-image AI helped 'better understanding some other group member's (teammates') ideas and other forms of wearables' and that 'I felt first not sure about our alignment of visions and towards the end I think we could better understand one another's ideas'. Another participant wrote 'We discussed prompts with the group (team). To find them I visualized my idea of what the product would be and thought of words that would be associated with the idea.' One participant said 'Yes, I felt more isolated in terms of my ideas before' when asked if Midjourney had helped with the group (team) vision,

and during the recorded session this participant also expressed that they now see the possibility of integrating with other group members' (teammates') ideas.

4 DISCUSSION

4.1 An AI-assisted Sketching Process

Sketching can be understood as 'rough drawing' [16] and has been used to communicate ideas and perform discussions that are beyond verbal in practice [57] and beyond hand-drawing [56]. Whereas the text-to-image AI in these experiments takes the semantic meaning of words to generate detailed image representations of those meanings. As such it can be seen to be performing a version of the task of *sketching* for the participants by highlighting the potential semantic meanings of the *chief features* of the prompt. In doing so, the text-to-image AI provides direct connections to the latent space



A: It looks like plants in there.
B: It looks like broccoli.
Facilitator: Like what?
A: Broccoli.
A: I do not see it.
A: Just because it's green doesn't mean it's broccoli.
B: I'm just hungry.
B: I see rice, and then broccoli, and then paprika.
A: What's the blue stuff, according to you?
B: I see ferns.
A: Ferns.
A: Yeah, I guess it is, as well.
A: It's kind of like a mossy type structure, especially in the heel there.
A: It's kind of like moss.
B: Oh, yeah, yeah.

Figure 8: Discussion of team 'sneaker' on a selected generated image from the prompt 'a colorful sneaker made out of plastic and fungi' where the participants were discussing how the green part reminds them of 'broccoli', 'ferns', and 'moss', which they were freely expressing their own interpretations.



A: Those are really cool, as well.
A: I really like this one.
B: Cheese strings.
A: Cheese strings!
C: Yeah, pasta.
A: Oh, yeah.
A: It kind of does look like fresh pasta.
C: Yeah, like when you have pasta, when it cools down, it sticks together, right?
A: Like this.
C: Yeah.

Figure 9: Discussion of team 'sneaker' on a selected generated image from the prompt 'a colorful sneaker made out of plastic and fungi', where the participants find the structure looks like 'cheese strings' and 'pasta', which they have come to an agreement for the interpretations.

it was trained on. In other words, the AI has an inbuilt connection with a much larger reference set, or maybe more specifically, a reference system that is close to but different from the one of the designer. In this way, the AI can be seen to contribute to the sketch as an external partner. Generative text-to-image AI also provides a kind of digital sketching. Previous research has found that analogue sketching is preferable to digital sketching because digital sketching takes time to refine the details [10]. However, text-to-image AI provides ease in generating images with exquisite detail and could greatly reduce the time cost.

4.2 Engagement in Verbal Articulation

The machine learning model provides its own interpretation of the prompt texts articulated by the participants and makes a translation to imagery according to the data it was trained on. However, these results are often strange and different from the participant's perspective. This difference encourages reflection on the choice of

words and what these words might mean, especially when the result is unexpected. The surprises provoke the participant to search for different words and explore the possible connotations between the words and the images they generate. In this way, participants tend to experiment with the text-to-image AI in a 'sketch-like' manner, where the prompts can be revised easily, and images are generated within minutes. This back-and-forth cognitive process engages translation between words and images and pushes the participants to explore. As a result, 25 out of 27 participants found text-to-image AI inspiring and provided more possibilities that they could explore. Previous findings by Tholander and Jonsson [54] illustrate that experienced designers have to find a *syntax* to talk to the AI, and that such a process helps the users to 'fine-tune' their design. Our finding supports such arguments, as well as its advantage in the design education context, where our student participants went through a co-evolutionary semi-guided idea exploration through their articulation of the text prompts.



A: It looks like a carcass.
B: Yeah.
B: The front looks like a shell or an exactly like bone material.
B: Shoes.
B: The shapes are really damn nice.
A: Yeah.
A: Very organic as well.
B: It's like the sand on the beach.
C: Oh Yeah, that as well.
A: Oh, yeah.
A: Those little waves that you get on dunes and stuff. In deserts, you also have those waves on top of the sand. A kind of a pattern, I get that, yeah.

Figure 10: Discussion of team ‘sneaker’ on a selected generated image from the prompt ‘modular sneakers based on biomimicry’, where the participants construct their interpretation on top of the others’ interpretations.

In addition to the semi-guided idea exploration interactions between the participants and AI, the ambiguity of generated images from the prompt redirects the attention of the participants towards the generated images instead of the project. This encourages the participant to distance from the project momentarily and reflect, which aligns with what Cross [17] have summarised that experts are more capable of ‘mentally stand back from the specifics of the accumulated examples’.

4.3 Building a Common Verbal-Visual Vocabulary while Brainstorming

The ambiguity that is innate within the AI-generated contents has been previously assumed as a negative characteristic by Tholander and Jonsson [54] because of ‘the system’s lack of context’ but this lack of clarity might in fact be providing an open space for participants in teams to interpret, negotiate, and aligning with each other. The generated images provide the participants with a shared reservoir of visual, textural, and structural elements, which allows them the opportunity to align with each other. They also provide visual artefacts for each individual to express their *ways of seeing* through annotation and discussions of what can be seen as an *objective* object generated by a perceived ‘third party’. Because of this and the fact that the imagery holds certain ambiguity (see example of how the participants interpret a piece of skin in Figure 6), participants felt free to express their interpretations of the elements, how they perceive them, and what they find interesting. The ambiguity also opens a space that allows them to elaborate on one another’s ideas. The unexpected results make the text-to-image AI a good partner for brainstorming by bringing related but un-thought-of elements. We also witnessed a process of grounding [15] where all participants linger on a common image for interpretations and agreements, while also acknowledging each others’ understanding. Furthermore, previous research has discovered that verbal-visual translation plays a strong role in ideation [55], whereas in our research the verbal-visual translation also happens as the text-to-image AI requires participants to enter texts in order to generate

images. Such a process allows the participants to engage in deeper discussions by exercising verbal-visual translation.

4.4 Shift of Attention to create the First and Secondary Discussion Space

In observing the team dynamics (Figure 11), we can see that introducing text-to-image AI (Figure 12) draws the team’s attention from focusing on the discussion around the actual design to the AI-generated images, and pushes the discussion of the actual design ideas in a ‘secondary discussion space’ what would be attended to much later than in the context of discussion without text-to-image AI. In other words, the temporary goal of the discussion is no longer to reach an agreement on design ideas, instead, the members of the team would focus on understanding each other. The convergence of semantic space is given time and space to work with before the team exercises on the design ideas. Furthermore, the aforementioned process of building a common verbal-visual vocabulary demands the team to spend time to bring each other on the same page semantically. The given time and space for semantic convergence and the urgency to reach semantic convergence bring changes to the social interaction dynamic.

In normal team dynamics as in Figure 11, the human agencies constantly check external references when discussing ideas. The generated ideas are feedback to the discussion, where the selection of ideas also happens. We refer to this collection of ideas as the ‘idea cloud’. The idea cloud is situated within the primary discussion space, meaning that the ideas within the ideas cloud are directly being discussed as well.

When text-to-image AI is integrated into the discussion, the dynamic changes, which is not unexpected but also not predicted as previous research has revealed that new digital tools will also influence the choices of existing tools for ideation [35]. The convergence of conceptual space, particularly the semantic space, happens during the construction of the text prompt instead of when the members are working on the ideas cloud. The human participants are exercising semantic convergence during prompt construction instead of during idea generation. We argue that this creates an extra

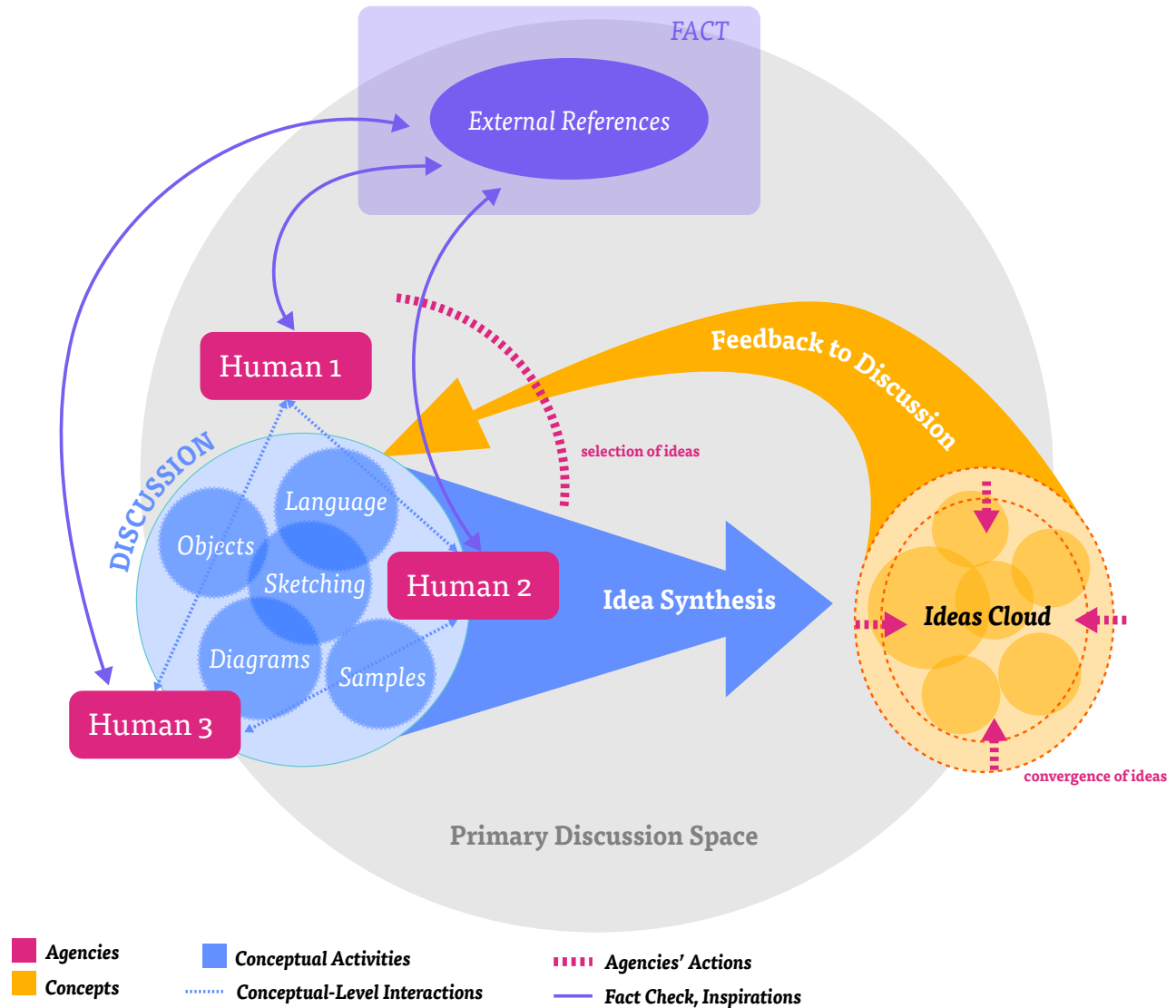


Figure 11: Dynamic diagram of the team discussion. The human agencies constantly check external references when discussing ideas. The generated ideas are feedback to the discussion, where the selection of ideas also happens. The idea cloud is situated within the primary discussion space, meaning that the ideas within the ideas cloud are directly being discussed as well.

step before the team reaches the final solution, which allows all parties to synchronise with each other semantically before reaching the stage of evaluating ideas (Figure 12) for physicalisation. In Figure 12, we observed three changes: the convergence of ideas happened around text prompts instead of the ideas cloud, the fact-check happened with the generated images by text-to-image AI instead of external references, and the ideas cloud is being pushed to the secondary discussion space. The text-to-image AI serves three purposes within the collaboration. First, text prompts as a medium that provokes a natural occurrence of semantic convergence through the discussion of the generated images amongst team members.

Second, the social interpretation of text-image relations through big data which in other words reflects social norms and biases that are innate to text-to-image AI as previous research has revealed [53]. Third, text-to-image AI engages human agencies to spend more effort pondering the synthesis of intermediate results before reaching an idea. It is because text-to-image AI not only demands that human agencies translate their ideas into text-prompt which text-to-image AI responds to through images, but it also visualises the thoughts of team members with its interpretation through the data it has been trained on. The visualisation is often not as exact as expected

according to the results we have found, but it triggers more discussion amongst the human members. Very often further inquiries were made through further generation of images. In hindsight, the extended discussions on the texts and images both demanded and generated by the text-to-image AI prolong the discussion before moving on to a concrete idea. The focus of discussion, which we refer to as the primary discussion space, does not contain the direct discussion of the ideas. Instead, discussions of ideas are pushed to the secondary discussion space, meaning that they will be discussed later than in the context without text-to-image AI.

In addition to looking at how the social dynamics change among the human members in the team collaboration, it is also crucial to mention that text-to-image AI in this context is not merely a tool but an entity with its agency. The generated images as AI's interpretation of the text prompts are being taken with weight which induces the change in the social dynamics. The coalition of humans and text-to-image AI forms a constituency [58] that influences the outcome of the design, where a 'wrong' interpretation is being taken seriously and its potential connotations and reasons are being discussed by the human members. In the post-questionnaire, multiple participants expressed that text-to-image AI felt like a conversation partner. Which, we can see an assemblage forming in the process. Although in this experiment text-to-image AI is not fully a collaborator, it has displayed a certain agency. Whether the agency is being dedicated by human members or innate to the text-to-image AI is out of the current scope of research. However, in our results, the images generated by text-to-image AI nudged the direction of the discussion, which is typically a role played only by humans. As the human members allowed text-to-image AI to influence the design ideas, text-to-image AI has taken a presence within the collaboration.

5 CONCLUSION - A NEW MODEL OF DESIGN IDEA GENERATION

In this paper, we have described an investigation of how text-to-image AI can potentially be used as a brainstorming aid for design students. Using the semester-long design assignment where students were asked to work on self-defined projects that related to wearable, tangible, and physical material and contexts, we experimented with giving the students the opportunities to use text-to-image AI to facilitate their ideation process. Within the context, we have identified three traits: *Enhancing verbal articulation*, where constructing a prompt engages the designer to articulate with words and the connotation between these words. Particularly because text-to-image AI models are trained on databases of human-created text and image pairs, that reflect genre, bias, and abstract connections between text and image. Evaluating the words of a prompt by evaluating the images generated, engages the designer in an exercise of verbal articulation and creates a deeper understanding of the connection between the word and the imagery. *Semantic convergence between collaborators* can be seen to emerge from the designer's encounter with the ambiguity of materiality, shapes, and symbols in the generated images. The ambiguity of the imagery invites collaborators to share interpretations, from which they gain the opportunity to understand what the others see in their semantic space. *Secondary discussion space* is being created by introducing

text-to-image AI as a non-human agency within the collaboration. The co-creation collaboration between human and non-human (AI) collaborators changes the traditional human-only team dynamic and creates a secondary discussion space as a buffer for social tension, thereby potentially creating a constructive atmosphere.

Although this paper only investigated the context of using text-to-image AI for ideation in wearable, tangible, and physical material-related projects for inexperienced students, we believe the identified traits could potentially be applied in other contexts with different material focuses.

Whether the findings are context-specific or not, it suggests a holistic view on evaluating machine learning models for creative collaborations should not only be looking at the *quality* or how a *final product* can be generated to fulfil a creative task but also the possibility of creating new connections for epistemological exercises. As such, AI can become a medium for linking different cognitive spaces for *thought experiments* or *sketching*, thereby extending the opportunities for semantic convergence in design processes.

REFERENCES

- [1] Tessa Aarts, Linas K. Gabriellaitis, Lianne C. de Jong, Renee Noortman, Emma M. van Zoelen, Sophia Kotea, Silvia Cazacu, Lesley L. Lock, and Panos Markopoulos. 2020. Design Card Sets: Systematic Literature Survey and Card Sorting Study. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference (Eindhoven, Netherlands) (DIS '20)*. Association for Computing Machinery, New York, NY, USA, 419–428. <https://doi.org/10.1145/3357236.3395516>
- [2] AEC Magazine AEC Magazine. 2022. Midjourney Architecture: Hassan Ragab. Retrieved Jan 10, 2023 from <https://aecmag.com/ai/midjourney-architecture-hassan-ragab-conceptual-ai-artist/>
- [3] Hosam Al-Samarraie and Shuhaila Hurmuzan. 2018. A review of brainstorming techniques in higher education. *Thinking Skills and Creativity* 27 (2018), 78–91. <https://doi.org/10.1016/j.tsc.2017.12.002>
- [4] Kristina Andersen, Ron Wakkary, Laura Devendorf, and Alex McLean. 2019. Digital crafts-machine-ship: creative collaborations with machines. *Interactions* 27, 1 (Dec. 2019), 30–35. <https://doi.org/10.1145/3373644>
- [5] Cynthia J Atman, Justin R Chimka, Karen M Bursic, and Heather L Nachtmann. 1999. A comparison of freshman and senior engineering design processes. *Design studies* 20, 2 (1999), 131–152.
- [6] Linden J Ball and Bo T Christensen. 2019. Advancing an understanding of design cognition and design metacognition: Progress and prospects. *Design Studies* 65 (2019), 35–59.
- [7] Jesse Josua Benjamin, Arne Berger, Nick Merrill, and James Pierce. 2021. Machine Learning Uncertainty as a Design Material: A Post-Phenomenological Inquiry. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (Yokohama, Japan) (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 171, 14 pages. <https://doi.org/10.1145/3411764.3445481>
- [8] Mark Blythe, Kristina Andersen, Rachel Clarke, and Peter Wright. 2016. Anti-Solutionist Strategies: Seriously Silly Design Fiction. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 4968–4978. <https://doi.org/10.1145/2858036.2858482>
- [9] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (Jan. 2006), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- [10] Jorge D Camba, Mark Kimbrough, and Eunsook Kwon. 2018. Conceptual product design in digital and traditional sketching environments: a comparative exploratory study. *Journal of Design Research* 16, 2 (2018), 131–154.
- [11] Peter C.-H. Cheng, Ric K. Lowe, and Mike Scaife. 2001. Cognitive Science Approaches To Understanding Diagrammatic Representations. In *Thinking with Diagrams*, Alan F. Blackwell (Ed.). Springer Netherlands, Dordrecht, 79–94. https://doi.org/10.1007/978-94-017-3524-7_5
- [12] 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 (Pittsburgh, PA, USA) (DIS '23)*. Association for Computing Machinery, New York, NY, USA, 1941–1954. <https://doi.org/10.1145/3563657.3596001>
- [13] Mao-Lin Chiu. 2002. An organizational view of design communication in design collaboration. *Design studies* 23, 2 (2002), 187–210.
- [14] HHCM Christiaans and Kees H Dorst. 1992. Cognitive models in industrial design engineering: a protocol study. *Design theory and methodology* 42, 1 (1992),

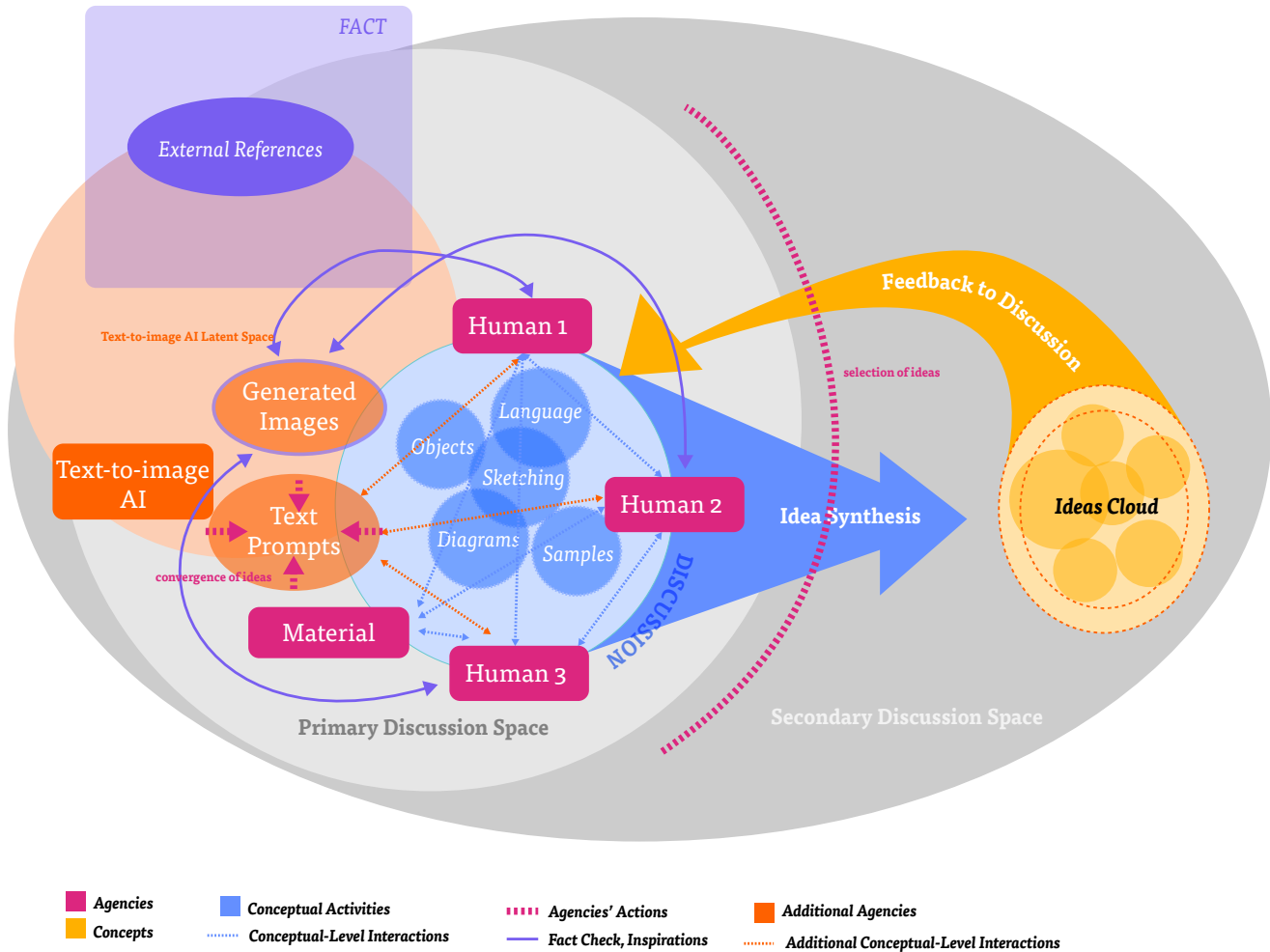


Figure 12: Dynamic diagram of the team discussion with text-to-image AI integrated. In this diagram, we observed three changes: the convergence of ideas happened around text prompts instead of the ideas cloud, the fact-check happens with the generated images by text-to-image AI instead of external references, and the ideas cloud is being pushed to the secondary discussion space.

- 131–140.
- [15] Herbert H. Clark and Susan E. Brennan. 1991. Grounding in communication. In *Perspectives on socially shared cognition*. American Psychological Association, Washington, DC, USA, 127–149. <https://doi.org/10.1037/10096-006>
- [16] Brock Craft and Paul Cairns. 2009. Sketching Sketching: Outlines of a Collaborative Design Method. In *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology* (Cambridge, United Kingdom) (BCS-HCI '09). BCS Learning & Development Ltd., Swindon, GBR, 65–72.
- [17] Nigel Cross. 2004. Expertise in design: an overview. *Design studies* 25, 5 (2004), 427–441.
- [18] Gilles Deleuze, 1925–1995., Félix Guattari, 1930–1992., and Brian Massumi. 1988. *A thousand plateaus : capitalism and schizophrenia*. Athlone Press, London. Section: XIX, 610 p. : ill. ; 24 cm.
- [19] Ellen Yi-Luen Do and Mark D. Gross. 2001. Thinking with Diagrams in Architectural Design. *Artificial Intelligence Review* 15, 1/2 (2001), 135–149. <https://doi.org/10.1023/a:1006661524497>
- [20] Martin J. Eppler and Roland Pfister. 2011. Sketching as a Tool for Knowledge Management: An Interdisciplinary Literature Review on Its Benefits. In *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies* (Graz, Austria) (i-KNOW '11). Association for Computing Machinery, New York, NY, USA, Article 11, 6 pages. <https://doi.org/10.1145/2024288.2024301>
- [21] Jonathan Fish and Stephen Scrivener. 1990. Amplifying the mind's eye: Sketching and visual cognition. *Leonardo* 23, 1 (1990), 117.
- [22] Patricia I Fusch and Lawrence R Ness. 2015. Are We There Yet? Data Saturation in Qualitative Research. *The Qualitative Report* 20, 9 (2015), 1408–1416. <http://nsuworks.nova.edu/tqr/vol20/iss9/3>
- [23] William Gaver. 2012. What Should We Expect from Research through Design?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) (CHI '12). Association for Computing Machinery, New York, NY, USA, 937–946. <https://doi.org/10.1145/2207676.2208538>
- [24] William Gaver, Peter Gall Krogh, Andy Boucher, and David Chatting. 2022. Emergence as a Feature of Practice-Based Design Research. In *Proceedings of the 2022 ACM Designing Interactive Systems Conference* (Virtual Event, Australia) (DIS '22). Association for Computing Machinery, New York, NY, USA, 517–526. <https://doi.org/10.1145/3532106.3533524>

- [25] William W. Gaver, Jacob Beaver, and Steve Benford. 2003. Ambiguity as a Resource for Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) (CHI '03). Association for Computing Machinery, New York, NY, USA, 233–240. <https://doi.org/10.1145/642611.642653>
- [26] John Gero and Julie Milovanovic. 2022. Creation and characterization of design spaces. In *DRS Biennial Conference Series, DRS2022*, D. Lockton, P. S. Lenzi, Hekkert, A. Oak, J. Sádaba, and P. Lloyd. (Eds.). Design Research Society, Bilbao, Spain. <https://dl.designresearchsociety.org/drs-conference-papers/drs2022/researchpapers/44>
- [27] John S Gero and Udo Kannengiesser. 2004. The situated function–behaviour–structure framework. *Design studies* 25, 4 (2004), 373–391.
- [28] Milene Gonçalves, Carlos Cardoso, and Petra Badke-Schaub. 2014. What inspires designers? Preferences on inspirational approaches during idea generation. *Design studies* 35, 1 (2014), 29–53.
- [29] Bruna Goveia da Rocha and Kristina Andersen. 2020. Becoming Travelers: Enabling the Material Drift. In *Companion Publication of the 2020 ACM Designing Interactive Systems Conference*. ACM, Eindhoven Netherlands, 215–219. <https://doi.org/10.1145/3393914.3395881>
- [30] Bruna Goveia Da Rocha, Johannes M. L. Van Der Kolk, and Kristina Andersen. 2021. Exquisite Fabrication: Exploring Turn-taking between Designers and Digital Fabrication Machines. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–9. <https://doi.org/10.1145/3411764.3445236>
- [31] Hamid Hassanzadeh. 2022. How midjourney and dall-e 2 help designers to create unique concepts? Retrieved Jan 10, 2023 from <https://parametric-architecture.com/how-midjourney-and-dalle-2-help-designers-to-create-unique-concepts/>
- [32] Serra Utkum Ikiz. 2022. 10 designers and their impressive Midjourney Works you should know. Retrieved Jan 10, 2023 from <https://parametric-architecture.com/10-designers-and-their-impressive-midjourney-works-you-should-know/>
- [33] Adobe Inc. 2023. *Adobe Firefly*. Adobe Inc. <https://firefly.adobe.com/>
- [34] Discord Inc. 2023. *Discord*. Discord Inc. <https://discord.com/>
- [35] Ben Jonson. 2005. Design ideation: the conceptual sketch in the digital age. *Design Studies* 26, 6 (Nov. 2005), 613–624. <https://doi.org/10.1016/j.destud.2005.03.001>
- [36] Manolya Kavakli and John S Gero. 2002. The structure of concurrent cognitive actions: a case study on novice and expert designers. *Design Studies* 23, 1 (2002), 25–40. [https://doi.org/10.1016/S0142-694X\(01\)00021-7](https://doi.org/10.1016/S0142-694X(01)00021-7)
- [37] Janin Koch, Andrés Lucero, Lena Hegemann, and Antti Oulasvirta. 2019. May AI? Design Ideation with Cooperative Contextual Bandits. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK) (CHI '19). Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3290605.3300863>
- [38] Ilpo Koskinen, John Zimmerman, Thomas Binder, Johan Redstrom, and Stephan Wensveen. 2013. Design Research Through Practice: From the Lab, Field, and Showroom. *IEEE Transactions on Professional Communication* 56, 3 (2013), 262–263. <https://doi.org/10.1109/TPC.2013.2274109>
- [39] Yuyu Lin, Jiahao Guo, Yang Chen, Cheng Yao, and Fangtian Ying. 2020. It Is Your Turn: Collaborative Ideation With a Co-Creative Robot through Sketch. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3313831.3376258>
- [40] Vivian Liu and Lydia B Chilton. 2022. Design Guidelines for Prompt Engineering Text-to-Image Generative Models. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 384, 23 pages. <https://doi.org/10.1145/3491102.3501825>
- [41] Midjourney. 2022. *Midjourney*. Midjourney Inc. Retrieved Jan 10, 2023 from <https://midjourney.com>
- [42] Newfacet. 2022. NEWFACET: FUSING AI AND FASHION. Retrieved Jan 10, 2023 from <https://www.designboom.com/design/newfacet-imagines-abstract-fashion-campaign-ai-program-stable-diffusion-12-29-2022/>
- [43] Jeffrey V. Nickerson, James E. Corter, Barbara Tversky, Yun-Jin Rho, Doris Zahner, and Lixiu Yu. 2013. Cognitive tools shape thought: Diagrams in design. *Cognitive Processing* 14, 3 (2013), 255–272. <https://doi.org/10.1007/s10339-013-0547-3>
- [44] Dimas Rifqi Novica, Elizabeth Wianto, and Susana Andrade Campos. 2023. Drawing and ideation process in design education: A systematic literature review. *Cogent Arts & Humanities* 10, 1 (2023), 2219487.
- [45] Alex Faickney Osborn. 1953. *Applied Imagination: Principles and Procedures of Creative Thinking*. Scribner, New York, NY, USA. <https://books.google.com.tw/books?id=0u49AAAAYAAJ>
- [46] J. Poon and M.L. Maher. 1997. Co-evolution and emergence in design. *Artificial Intelligence in Engineering* 11, 3 (1997), 319–327. [https://doi.org/10.1016/S0954-1810\(96\)00047-7](https://doi.org/10.1016/S0954-1810(96)00047-7)
- [47] Aditya Ramesh, Prafulla Dhariwal, Alex Nichol, Casey Chu, and Mark Chen. 2022. Hierarchical text-conditional image generation with clip latents. *arXiv preprint arXiv:2204.06125* 1, 2 (2022), 3.
- [48] Jeba Rezwana and Mary Lou Maher. 2021. COFI: A Framework for Modeling Interaction in Human-AI Co-Creative Systems. In *Proceedings of the Twelfth International Conference on Computational Creativity, México City, México (Virtual), September 14-18, 2021*, Andrés Gómez de Silva Garza, Tony Veale, Wendy Aguilar, and Rafael Pérez y Pérez (Eds.). Association for Computational Creativity (ACC), 444–448. <https://computationalcreativity.net/iccc21/wp-content/uploads/2021/09/Wkshp2.pdf>
- [49] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. 2021. High-Resolution Image Synthesis with Latent Diffusion Models. arXiv:2112.10752 [cs.CV]
- [50] Kevin Roose. 2022. A.I.-generated art is already transforming creative work. Retrieved Jan 10, 2023 from <https://www.nytimes.com/2022/10/21/technology/ai-generated-art-jobs-dall-e-2.html>
- [51] Jeremy Roschelle. 1992. Learning by Collaborating: Convergent Conceptual Change. *The Journal of the Learning Sciences* 2, 3 (1992), 235–276. <http://www.jstor.org/stable/1466609>
- [52] Simeon J Simoff and Mary Lou Maher. 2000. Analysing participation in collaborative design environments. *Design studies* 21, 2 (2000), 119–144.
- [53] Ramya Srinivasan and Kanji Uchino. 2021. Biases in Generative Art: A Causal Look from the Lens of Art History. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (Virtual Event, Canada) (FAccT '21). Association for Computing Machinery, New York, NY, USA, 41–51. <https://doi.org/10.1145/3442188.3445869>
- [54] Jakob Tholander and Martin Jonsson. 2023. Design Ideation with AI - Sketching, Thinking and Talking with Generative Machine Learning Models. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (Pittsburgh, PA, USA) (DIS '23). Association for Computing Machinery, New York, NY, USA, 1930–1940. <https://doi.org/10.1145/3563657.3596014>
- [55] Anne Tomes, Caroline Oates, and Peter Armstrong. 1998. Talking design: negotiating the verbal–visual translation. *Design Studies* 19, 2 (April 1998), 127–142. [https://doi.org/10.1016/S0142-694X\(97\)00027-6](https://doi.org/10.1016/S0142-694X(97)00027-6)
- [56] Elvia Vasconcelos. 2022. Weaving activism into the fabric of everyday life - Technical University of Technology. <https://ddw.nl/en/programme/8078/weaving-activism-into-the-fabric-of-everyday-life>
- [57] Elvia Vasconcelos. 2023. The Sketching Table. https://www.instagram.com/the_sketching_table/
- [58] Ron Wakkary. 2021. *Things we could design: For more than human-centered worlds*. MIT press, Cambridge, Massachusetts.