The AI generative text-to-image creative learning process: An art and design educational perspective

Tore Andre Ringvold, Oslo Metropolitan University
Ingri Strand, Oslo Metropolitan University
Peter Haakonsen, Oslo Metropolitan University
Kari Saasen Strand, Oslo Metropolitan University

Abstract

In today's constantly changing world technological developments in artificial intelligence (AI) can induce educational visions of both utopia and dystopia. New technologies and communication platforms can provide new forms and possibilities of learning. Creating an image has historically mostly been a human process of using knowledge and application of technique that demanded training. This image-making process changed with the invention, development and spread of the photographic camera, when creating a detailed visual representation of reality became a possibility without a complex process of craftsmanship and artistry. The nature of visual art changed but the visualisation of ideas and prefigurative thoughts could not necessarily be captured by a camera. With the development and spread of AI text-to-image generation, can this change the need for competency to visualise ideas in the way the camera changed the need for drawings and paintings as visual representations? This study explores how AI text-to-image generators can contribute to and change art and design education. We conducted exploratory experiments where we tested a variety of AI text-to-image generators and explored the outcome of using different generators, prompts and settings. Reflections were written down throughout the process. This was combined with an online ethnography on a text-to-image community. Different potentials of learning were identified, as well as issues of interaction and possible contexts of use. The results are discussed in a future learning context.

Keywords
Artificial intelligence, Text-to-image generation, Art and design education, Visualisation

Introduction
Education and the way we learn needs to adapt to society. Constant changes in technology have led to new cultures and places of learning (Thomas & Brown, 2011). Artificial intelligence (AI) represents a digital technology that has great power to change society. With the rapid development of AI technology, educators need to adapt to emerging technology that can potentially change how we work and produce. Artificial intelligence text-to-image technology has been developed since 2014 and increased in popularity after the introduction of DALL-E in 2021 (Cetinic & She, 2022). In 2022, this development accelerated rapidly with the launch of services such as Midjourney, OpenAI's DALL-E 2, Adobe Firefly, and Stable Diffusion. With the widespread use and rapid technological development, there is a need for research to critically look at the possible use and application of AI text-to-image generation. In this study, we ask: How can AI text-to-image generators contribute to and change art and design education?

1

Background

A short history of visualisation

Prefigurative thinking is seen as an important part of what constitutes being human (Fry, 2012). The prefigurative thought of an idea or concept can be materialised through medium and technique. Doing this, and also being good at it, are no easy tasks. It takes time to develop the knowledge and skills to produce traditional visual forms that constitute sensorimotor activities, such as drawing and painting. There were no quick-fix methods for creating quality visualisations before the introduction of photography. The process of image creation underwent a significant transformation with the advent, progress, and widespread use of the photographic camera (Gombrich, 1982). The camera enabled the creation of highly accurate visual depictions of reality without intricate craftsmanship and artistic techniques. However, as prefigurative thoughts cannot necessarily be captured by a camera, there is a need for skills in drawing and illustration (Nielsen, 2013). This raises the question if the development and spread of AI text-to-image generation can change the need for competencies to visualise ideas in the way the camera changed the need for drawings and paintings as mimetic visual representations.

Text-to-image generation technology

There are different text-to-image generative systems, but what they have in common is that textual inputs called *prompts* are interpreted by a system before images are created. The systems are trained on large datasets of text-and-image pairs from the web (Abdallah & Estevéz, 2023; Hutson et al., 2024). A prompt can lead to unexpected results, but at the same time the different models, such as Midjourney and OpenAI (DALL-E 2), provide tips on how to alter the style or format by adding specific terms. These tips can help one to affect the outcome and, by adjusting the prompt input, one can increasingly control the image-making process. We mostly used Midjourney version 4 in the beginning of our study. Our explorations coincided with significant developments of the generators, such as the release of Midjourney Version 5.1 and solutions to recurring issues such as the depiction of fingers (Verma, 2023). Later we have mostly used Midjourney version 5 and 6, as well as different non-payment generators such as Krea, Lexica, and Stable Diffusion Online. Due to the rapid development of the technology, we are more concerned with the overarching concept of AI text-to-image generators than specific technical aspects.

A changing art and design education providing competencies for the future

In Norwegian primary and lower secondary school, Art and crafts is a compulsory subject. The 2020 curriculum revision ensures a more future-directed education, in part by making technological use and programming more visible (Norwegian Directorate for Education and Training, 2020). The extensive development of AI in recent years renders it impossible to predict what the future will look like. Rather than transferring knowledge teachers must facilitate learning processes regarding future challenges. This necessitates that educators possess a distinct combination of knowledge, skills, and attitudes to use technologies effectively. This is often referred to as professional digital competence (Kelentrić et al., 2017). This is not just about having technical skills; it also involves understanding how to integrate technology into teaching and learning in ways that add value and enhance students' understanding. This also includes shifting the focus from teacher-led to learner-centred processes. The use of technology should therefore be determined by the needs and demands of the subject matter rather than forcing the subject to adapt to the technology. Technological

tools can aid in the illustration of complex concepts, facilitate broader discussions, and enable students to explore and learn at their own pace (European commission et al., 2017). The role of a digitally competent teacher also involves making responsible decisions about privacy and data security, understanding the ethical implications of technology use in the classroom, and fostering a respectful and inclusive digital learning environment. The 2023 GEM report, *Technology in education: a tool on whose terms?* (Global Education Monitoring Report Team, 2023) published by UNESCO argues that digital technologies must be used to support an education based on human interaction rather than aiming at substituting human interaction with digital technologies, ensuring learning processes where the student's learning is in focus.

Artificial intelligence will challenge schools' teaching and assessment practices, and a new government strategy (Kunnskapsdepartementet, 2023) requires schools to adopt AI in order to gain a basic understanding of how AI works, its solutions, and its limitations. Thus, teachers and students need sufficient digital competencies to use AI in an exploratory way, with curiosity, critical thinking, and ethical awareness.

Literature review

Widespread use of AI text-to-image generative technology is fairly recent. Naturally, this is an emerging research field where much of the literature is shared as pre-prints in non-peer reviewed archives. We have however chosen to limit our scope of this literature review to only include journal articles or conference proceedings. At the time of the study, there is little research on the use of text-to-image generative technology in K-12 arts education. We have therefore included research that explores the use of this technology in art and design processes in general.

In the literature, we have seen several ways of incorporating AI images into creative processes. Chen et al. (2019) conducted a case study on the design of spoons. In the ideation phase, participants used a GAN-model that combined selected images of leaves and spoons, to generate a multitude of synthesized images. The participants drew their spoon-designs by hand later. Compared to the control group, who only used Google in their ideation phase, the AI participants produced a larger quantity, variety, and novelty of their designs (Chen et al., 2019). Liu et al. (2023) also investigated product design. They implemented the AI text-to-image generator 3DALL-E into a 3D-modelling program, which meant that their participants worked with text-to-3D. Participants were enthusiastic about the use of AI, especially for ideation, as it was a quick and easy source of inspiration and helped them avoid design fixation. However, some participants felt like the AI was driving the design process, as they adapted their creative process to the output (Liu et al., 2023). Mikkonen (2023) explored how Midjourney could be used to generate mood boards. He concluded that the AI quickly produced high quality images that were visually usable in design but was concerned with potential copyright issues (Mikkonen, 2023).

The use of AI text-to-image generators have also been explored in the context of visual art. Hutson and Lang (2023) incorporated AI into a digital media course, in which the students' generated images that were modified further in image editing software. The students found the AI to enhance their outcomes, as it helped them to structure and visualise their ideas. They were however unsure about future use, as they perceived that the AI did not surpass human creativity, but that the images had 'a similar, postcard-like quality that hindered their artistic potential' (Hutson & Lang, 2023, p. 11). Lyu et al. (2022) compared AI generated images resembling oil paintings made by artists and non-artists, along with a questionnaire answered

by the participants. While the non-artists were excited about making images that looked excellent and perceived that the AI helped them to visualise their imagination, the artists were dissatisfied due to a feeling of losing control of the process. Participants mentioned being surprised by the images that were generated, and two artist participants likened it to opening Pandora's box. Finally, in an expert ranking of the outcomes, the images made by non-artists obtained slightly more votes (Lyu et al., 2022). Ko et al. (2023) interviewed visual artists after they had learned about and tested DALL-E. The participants saw a great potential for using AI in early phases, such as to quickly visualise ideas, and generate reference images or material for visual communication. They did however also experience that the articulation of their ideas into prompt could be time-consuming and that they may lose the delicate mental imagery while finding the correct text-prompt (Ko et al., 2023). Vartiainen and Tedre (2023) conducted a similar study, where pre-service craft teachers and teacher educators participated in a workshop followed by a joint discussion. Participants saw opportunities for using AI to visualise impossible ideas and mentioned that it might support small children to articulate vague ideas. Coming from the field of craft, they were however concerned about the lack of embodiment and engagement with materials in the creative process (Vartiainen & Tedre, 2023).

Among the recurring findings is that the AI may stimulate ideation processes by quickly generating unforeseen suggestions that can be a source for inspiration and references. AI was also found to support the visualising of ideas regardless of craft skills. There are however several concerns, regarding the constraining of creativity, copyright issues, and the lack of embodied engagement during the process.

Methodology

We have chosen a qualitative approach, combining explorative experiments and online ethnography. This generated data in form of notes and image material. As this article builds on a conference paper presented at the PATT23 conference (Ringvold et al., 2023), the study consisted of two phases. The first phase was an intensive six-month period from November 2022 to May 2023, when we finished the first paper. The second phase was a continuous exploration from May 2023 to April 2024, where we have continued with explorative experiments individually, but also included other participants in workshops or teaching at our university.

Explorative experiments

Explorative experiments, based on Dyrssen (2010), has been the main method used in this study. Dyrssen states that, while explorative experiments cannot be validated, they allow the researchers to 'shake up ingrained patterns of thought; provide quick feedback, increased curiosity, and discoveries of hidden possibilities; reveal possible links and points that need to be mapped; and get the creative process moving forward' (Dyrssen, 2010, p. 229). Due to the rapid development of AI text-to-image generators, such explorations of the technology have given us valuable insight into how it can contribute to and change art and design education.

In the explorative experiments, each of the four authors tested AI text-to-image generators and explored the outcome of different generators, prompts, and settings in use. In the autumn of 2022, we had little to no experience with AI text-to-image generators. The prompts were therefore written based on our own imagination and curiosities, sometimes choosing to follow interesting idea strands. We acknowledge that our abilities to prompt, and what we prompt have naturally evolved throughout the period of this study. The combined variety of prompt

inputs, ranging from the abstract to the concrete, is reflected in the examples presented in the results section.

Throughout the process of explorative experiments, we wrote down our reflections and saved image material that comprise the empirical data. We shared experiences from the explorations of the AI text-to-image generators in frequent meetings. In these discussions, we also drew on our backgrounds in design and education. We have approximately 10 years of teaching experience each, and combined this covers teaching at all levels, ranging from Year 4 (nine-year-olds) to university level. Through reflective dialogue, we discovered central issues regarding our experiences of AI text-to-image generators, as well as staking out a path for further experiments. While these meetings were more frequent and structured during the first intensive phase, we have continuously shared our experiences and reflections throughout both phases.

During the second phase of explorative experiments, we also conducted three workshops. These had a diverse range of participants, as one was for Art and crafts teachers (approximately 80 participants), one for students in Year 6 (11-12 years old) (6 participants), and one for colleagues from all faculties of our university, including retirees (12 participants). From these workshops, we have made observation notes, taken photographs, and saved generated image material. We have also implemented the use of text-to-image generators in our teaching at the Art and design teacher education at BA and MA levels. Our experiences from working with these participants have become an increasingly important part of our shared reflections in the meetings during this second phase.

The images in this paper are created with Midjourney Versions 4, 5 and 6. Version 4 was the default model from November 2022 to May 2023. Version 5 was released on 4 May 2023, and Version 6 has been the current default version since February 2024 (Midjourney, n.d.-b).

Online ethnography
Midjourney is accessed through a server on the online community platform Discord (Discord, n.d.; Midjourney, n.d.-a). This facilitates interaction between users, as the generated images normally are visible also to other users. This led the first author to conduct an online ethnography, as described by Hart (2017) and Winter and Lavis (2020) in the text-to-image community of the Midjourney server. The online ethnography was conducted at three different timepoints during the first, intensive phase, capturing users' interactions through screenshots of images and text.

Thematic analysis
The empirical material from the explorative experiments and the online ethnography were analysed through a thematic analysis as derived from Braun and Clarke (2022). The reflective notes were organised and temporarily coded based on the similarity of the content. We identified these themes and similarity of content related to AI text-to-image generators' potential for learning, limitations, and hindrances, as well as issues of interaction and possible contexts of use. The notes were then read again through the lens of the emerging themes. The themes were further developed and refined through defining, renaming, and merging themes (Braun & Clarke, 2022). The results from this analysis are presented and discussed in the following section.

An art and design educational perspective on the AI text-to-image creative learning process

The results and discussion are presented in a bilateral manner of learning potentials on the one side and learning obstacles on the other. However, the identified qualities are not necessarily either a potential or a hindrance to learning; they can be both at the same time. Firstly, we present two models of how we identified the AI text-to-image creative process in a learning perspective.

### Modes of creation processes from a learning perspective

Creating visualisations using text-to-image generators is a cyclic, iterative process. We identified this image creation process as consisting of several stages, as visually presented in Figure 1. The first of these is the idea stage. This can be a prefigurative thought at the beginning of a process, the refinement of previous ideas, an adaptation of a previous idea, or a surprising new idea that originated from a previous process. Taking the step from idea to a written prompt, one needs to shape the visual idea or concept through articulation (stage 2). After feeding the prompt in writing (stage 3), a black box process (Bunge, 1963) gives one the results (stage 4). Results can be implemented (used) or aborted (left unused). They can also be refined, adapted, or used to start new ideas in a cyclic process. These post-result actions are triggered by how the prompter evaluate the prompts, identifying what needs modifying, deleted, or added in the prompt. An example of this is when using terms with multiple meanings or specific cultural references in the prompts. While this might not be clear to the user initially, what needs to be modified in the prompt is usually easily detected in an evaluation of why the generated image differs from the users' idea.
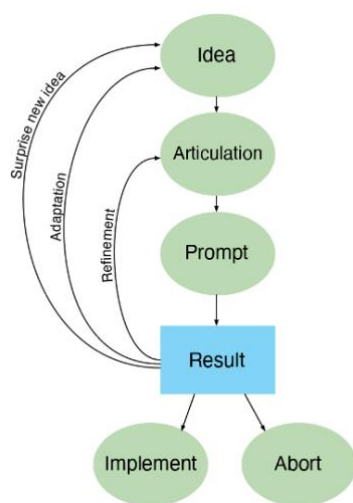
Figure 1. The identified image creation process using AI text-to-image generation.

In addition, we propose an understanding of the learning process that takes place while using AI text-to-image technology (Figure 2). In a cyclic, interactive learning process between the human user and text-to-image AI technology, a variety of actions can take place in an interchange between A: Seeing opportunities and B: Finding limitations. These actions can contribute to the cyclic development of learning processes of varying length. Variations also apply to the actions involved in the cycle. As signified by the multiple double-pointed arrows in

the model, the actions do not take place in a linear fashion, the user rather moves between some or all of the actions during an interactive learning process.
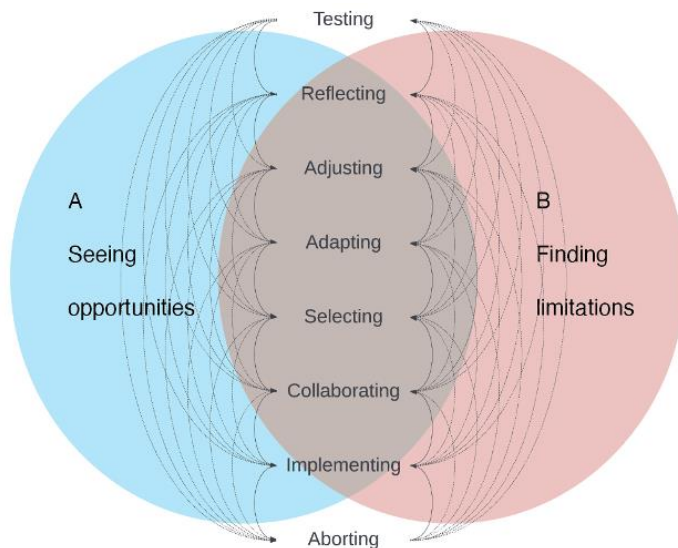
Figure 2. A proposed representation of the cyclic human user and AI text-to-image interactive learning process.

Potential learning opportunities

In this section, the identified potential learning opportunities are presented and discussed.

Enabling visualising abilities, a democratisation of visualising?

Artificial intelligence text-to-image technology provides the user with a powerful tool to generate visualisations without much effort. This ease of access to visualising allows for playful image-making. The image creation could, especially at the beginning of using the technology, be about testing limits, and seeing what is possible. Using this technology, students can possibly push their own boundaries of imagination. The ease of access can trigger a willingness to try and experiment with creative image making. Early adolescents may experience a more critical view of their produced drawings, as described by Lowenfeld (1947). For the non-professional image creator, text-to-image generation can provide a beneficial training ground and a tool for visualising. We see this as especially relevant for pupils who are hesitant to enter the creative processes of image making. By having access to new tools of visualising prefigurative thinking, more visualising processes can be materialised than if not. This harmonises with the results of Lyu et al. (2022), where the non-artist participants were excited by the possibility to create images that looked professional. AI text-to-image generators can also be used by those who are physically not able to use and master traditional forms and techniques of visualising, which mean that AI can allow for greater inclusion in visualising learning activities. Artificial intelligence text-to-image visualising can also allow for imagery which might not be possible at all or on the borders of or outside the parameters of skill. This imagery can include impossible situations or complex emotional expressions which can be hard to portray. An example of this is shown in Figure 3.1, where the intention was to show a person in a state of despair and rage bordering on crying. Producing this image with conventional photography would probably need a model or actor and a photographer at a very high skill level. This can open new possibilities of

visualising outside the real. In other attempts we explored impossible combinations, such as in Figure 3.2, an iteration of a 1910 archaeological dig of an ancient robot.
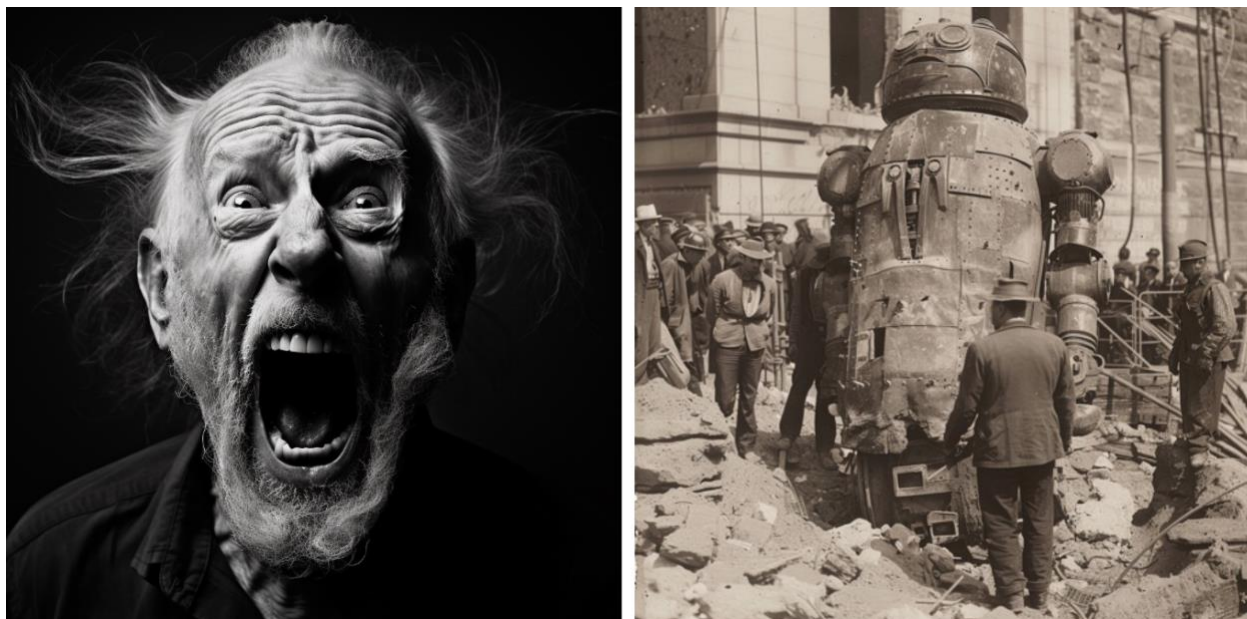
Figure 3.1 (left) and 3.2 (right). Generating emotionally complex or impossible imagery.

An online arena for sharing and collaborative creativity

Midjourney and Discord provide an arena for knowledge-sharing and collaborative generative art and design processes. These processes take place via a variety of chatrooms, show-and-tell rooms and other channels of communication. In these multi-human and machine collaborations, several users and the AI generative technology are part of discussions. Through discussions and testing, they collaborate on developing prompts, aesthetic qualities, and designs. This online space for potential collaboration and co-learning provides a learning environment independent of place. Figures 4.1, 4.2, 4.3 and 4.4 provide examples of Midjourney users collaborating and helping each other to achieve a desired image. Collaboration and seeking guidance from others through sharing is an important part of creative professional work and are also highlighted in new curricula in primary and secondary education (Norwegian Directorate for Education and Training, 2020).
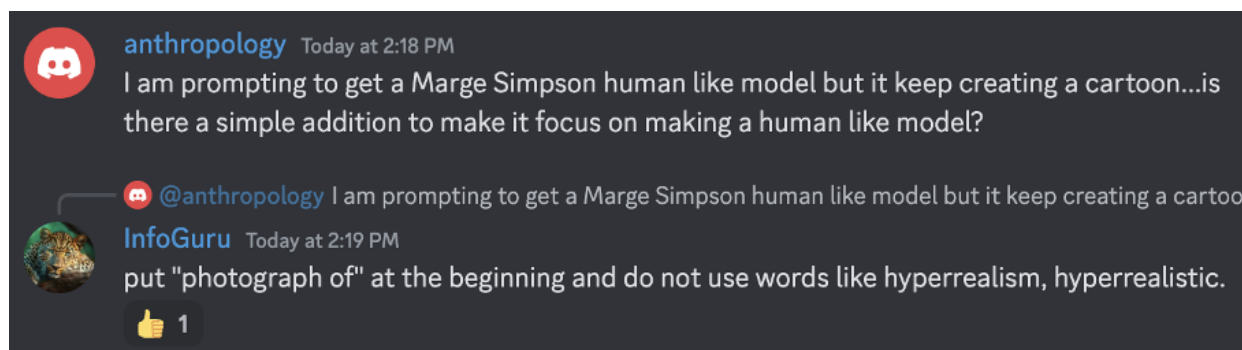
Figure 4.1. Requesting prompt help. Discord, Midjourney server, prompt-chat screenshot.

Figure 4.2. Help suggested. Discord, Midjourney server, prompt-chat screenshot.
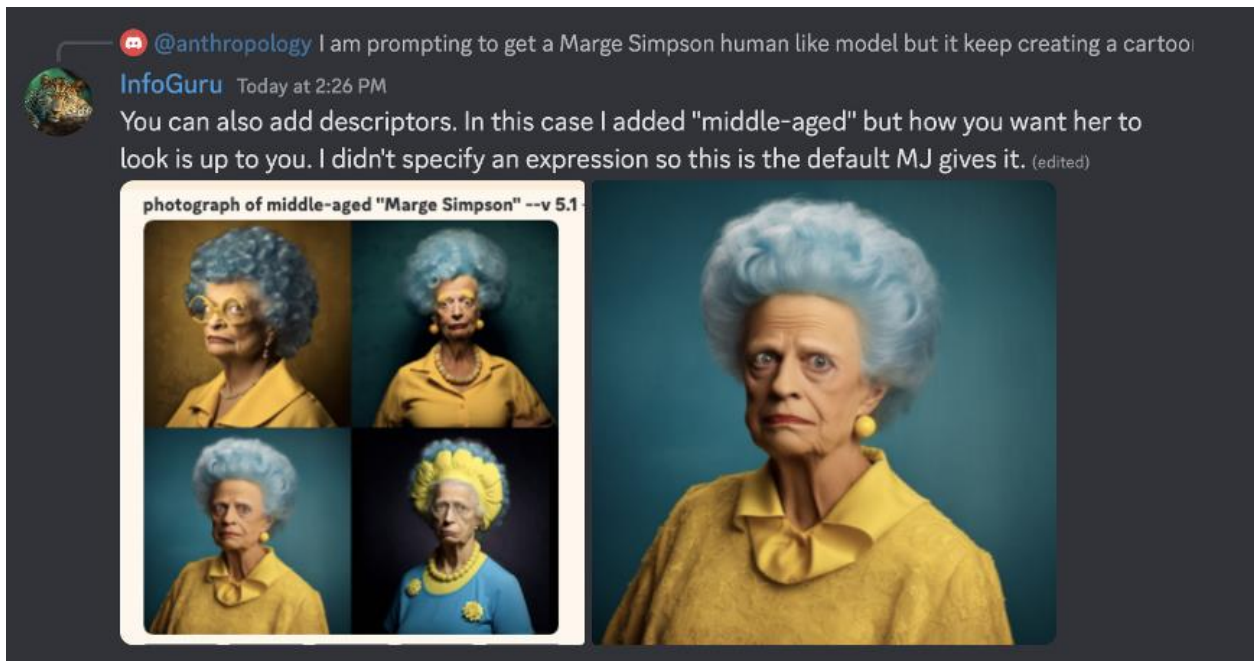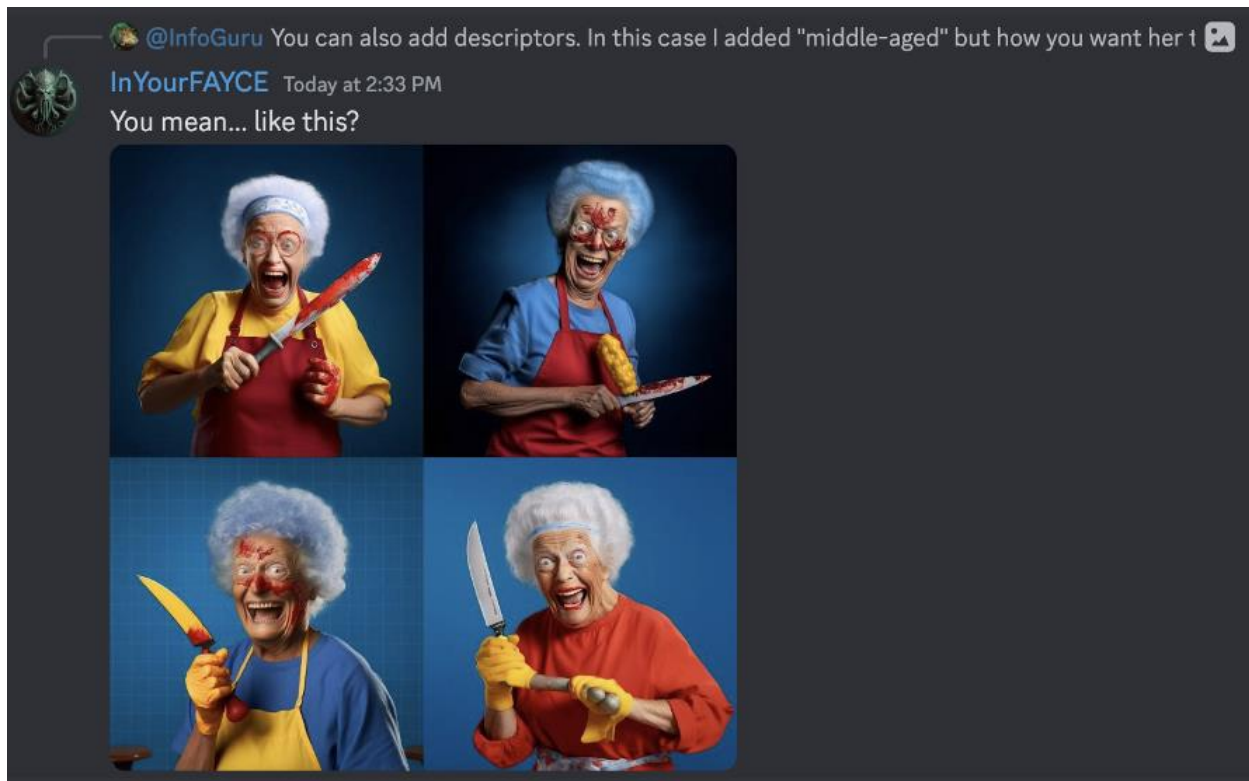

Figure 4.3. Help suggested, follow-up. Discord, Midjourney server, prompt-chat screenshot.

Figure 4.4. A third user joins. Discord, Midjourney server, prompt-chat screenshot.

A visual aid for developing perspectives and composition

Artificial intelligence images can be useful visual aids for creative work. In our exploration, one of the authors tried this approach to draw compositions for a comic book, including an image of a large dragon sleeping in a parking house. Using AI generated images, he could explore how to draw this composition. Images created with Midjourney often do not meet the creator's expectations, but the images can be used as sublayers for reference in constructing illustrations. By prompting several iterations of a dragon and the environment and combining multiple images to get a reliable perspective (Figure 5), the drawing could be completed. In another case AI images were used as visual aid to draw a greeting card to a boy's confirmation, a traditionally Lutheran but now also a secular coming of age-ceremony. When prompting, the term 'coming of age' didn't produce anything useful, neither did 'confirmation'. Being more concrete, describing the desired motif, such as 'young man' or '15-year-old boy in a suit, standing next to' or 'leaning against a tree', did produce enough imagery to be used when making a hand drawn and water colour card (Figure 6). Other visual aids in this matter included architectural designs, different perspectival problems, light and shadows, body postures and anatomy studies. The final sketches and drawings were based on parts of AI generated images built upon various iterations in an interaction between articulating and implementing or aborting the results before articulating new prompts.

Figure 5. Dragon and environment image generations combined with author's own sketch work (far right), used for reference in creation process.

Figure 6. Imagery from the process of making a water colour greeting card.

A multitude of idea generations

The study identified that, in creating object designs, the AI generated a multitude of varying idea visualisations. A series of relevant designs were easily created when the object shape was conventional, such as a ring (Figure 7.1), or a conventional object design such as a car or a toaster (Figure 7.2). The AI does not rethink a functional need or type of use but generates imagery based on convention if the name of the object is used in the prompt. In a search for new ways of thinking about object design, we have explored the use of more descriptive prompts, such as 'a new and ground-breaking technological device that can warm up slices of bread in a distant future created with an unknown material' (Figure 7.3). Another way of creating interesting results or suggestions useful for further explorations, are mashups of existing design visualisations. The variations are more decorative rather than shape variations.

In a creative learning process, students' abilities to generate ideas differ. A common obstacle in the creative process can be a lack of variation of ideas in the early stages, leading to unsatisfactory results that do not fulfil their potential. The ease of generating idea visualisations can help prevent concept or design fixation (Schut et al., 2020), which is also demonstrated in previous research on AI in product design (Chen et al., 2019; Liu et al., 2019). Artificial intelligence text-to-image technology can provide new avenues for solutions to a problem or new ways of seeing and understanding.

Figure 7.1. Engineering department alumni ring design propositions.

Figure 7.2. Different toaster design propositions.

Figure 7.3. Toaster design propositions attempting to avoid the object design conventions.

Potluck visualisation as a tool for creativity

Artificial intelligence text-to-image generation can be a game of chance, providing surprising image results. This potential potluck quality of AI text-to-image generations led us to look at our initial ideas in new and unforeseen ways. Opening new avenues of thought can be rewarding in the creative process. While providing new possibilities, it also emphasises the need for a more critical creator role, as further described below. Using chance or coincidence in idea generation

can contribute to new perspectives that enrich the creative process (Fazel & Almousa, 2021). Surprising image results, unrelated to the original idea, can be exciting to develop further. In these cases, the process, or 'dialogue' with the text-to-image generator, produces new and unexpected outcomes. This resonates with Oppenlaender (2022, p. 198), who has referred to unexpected results and serendipity, or two of the participants of Lyu et al. (2022) who likened the use of AI as opening Pandora's box.

The element of surprise has also a comical or amusing side to it which can be a social factor when it comes to learning. When prompting together in the classroom, students of both Year 6 and teacher education could be motivated by attempting combinations and aesthetic qualities that brought laughter to the group. This light-hearted side to prompting and text-to-image generation can benefit group dynamics which are valued in an educational setting. Humour can contribute to the learning environment in a positive way, helping the students in their willingness to focus on tasks (Banas et al., 2011).

Developing and strengthening a visual reference language
We identified articulation as an important stage between idea and prompting (Figure 1). The training of articulation from prefigurative thought to writing can contribute to the development and strengthening of a visual reference language. On the one hand, the ability to use the subjects' terminology is often highlighted in curricula (Norwegian Directorate for Education and Training, 2020). On the other hand, when articulating and writing prompts, one might use words or phrases not necessarily viewed as correct in the Art and design discipline. However, they work well when prompting due to their widespread use online, for example in gaming communities.

Novel users of text-to-image technology seems to differ in their ability to articulate and prompt accordingly to increasingly control the outcome of the generative process, based on our experiences from the workshops we have conducted. In one task we asked participants to prompt about their dream houses. Their backgrounds were diverse, ranging from advisors to teachers and students affiliated with various institutes and departments. There was also a variation in age and experience with digital technology. The outcomes from the prompts greatly varied. Some quickly understood that they had to adapt their prompt to express a different idea. However, others seemingly didn't comprehend that they had to describe what they wanted themselves. It appeared as if they thought the AI had intelligence, that it could understand what they meant. Much of this could be related to age - those with less experience and who are not as confident with digital technology, didn't quite grasp how it works. The Year 6 students, on the other hand, had a more intuitive understanding of the prompting, which they likened to chatting with a person. Those with more experience intuitively understood the task and adjusted their prompt accordingly. This concurs with Pennefather (2023) who highlights the importance of prompting multiple times: 'With little effort comes little result' (Pennefather, 2023, p. 206).

The 'concept-articulating catalyst wizard', a changing role of the image creator?
The study's exploration identified that the role of creator in creative processes using AI text-to-image technology differs from the creator's role in traditional image making such as drawing or painting. We see the creator's role shifting towards that of an art director, composer, editor, or selector. The process role of editing and catalysing black box processes consists of articulating,

testing, developing, adapting, refining, selecting, and editing by starting new cyclic processes, as described in Figures 1 and 2. This shift in the creator's role questions what kind of knowledge and skills will be needed in future creative processes. With a shift towards editing and selecting rather than producing, a critical mindset should be an important part of future creator and design competencies. Such a mindset should be critical of results and open to different solutions of visualising prefigurative thoughts.

In this AI-assisted creative learning process, the idea and articulation stages can be seen as more important, compared to a traditional image making process, due to the ease of generating or producing. This added importance of imagining and articulating what you imagine highlights the need for a focus on prefigurative thinking and communication skills in future education.

Potential learning obstacles

In the following section, we will present and discuss the potential learning obstacles identified through our study.

Chasing the centaur, not getting what one wants

An example from our testing was to create a centaur by using different prompt-writing approaches. Whether writing a short prompt, such as 'centaur', or describing what a centaur is, the results were mainly images of horses. Other results depicted a man standing in front of a horse-like body, or a human torso attached to a horse's back (Figure 8.1). Although other Discord users in the Midjourney community had managed to create centaurs, a successful prompt copied and pasted from the community also elicited poor results. Other absurd combinations were also difficult to accomplish. The first attempts at creating AI images may be fascinating, but the wow factor will not necessarily last for long. Trying to create something based on ideas and imagination may lead to disappointing results that do not match how one visualised the ideas in the first place. The natural limitations of a given technology or tool will limit the possibilities.

Figure 8.1. A selection of the failed centaurs.

In an ongoing process with several iterations of different centaurs ending up as horses, figure 8.2 is the closest we got so far. This was partly based on tips found in relevant discussions from the community, where phrases like 'wearing four graceful horse legs' was helpful in generating centaurs instead of horses. Other useful phrases could be that the centaur was holding

something, to indicate that it had arms. Writing 'wearing a helmet' could indicate a human head as opposed to a horse's head. This way of articulating gave us gradually better control over the results. Several results were abandoned, but parts of the prompts that gave somewhat satisfying results were reused in new combinations. In this way we redefined how we articulated, and results could be implemented in new iterations, or abandoned.

Figure 8.2. One of the most successful centaurs so far.

Bias and stereotypes

As the images generated by text prompts are based on image-and-text pairs from the internet, biases and stereotypes may be reproduced, which is also mentioned by Vartiainen and Tedre (2023). If ethnicity or gender was not specified in a prompt, a white male was often featured in the results in our early attempts unless the prompts contained words typically associated with woman. In 2024 Google's Gemini created images perhaps being too woke (Kleinman, 2024). Even if the biases are reduced, the datasets still need to be trained on classifications that put human beings into categories (Crawford, 2021).

In some attempts to see whether Midjourney made female or male superheroes, most results produced different versions of Superman in a style resembling acrylic paintings. With shorter prompts, Midjourney may create more surprising results, although one can recognise its default visual style (Pennefather, 2023). Often, digital illustrations of young women with long dark hair appear in the results based on short prompts, even if the prompt suggests content without any people in it, such as mood boards of a specific mood or hue of colour (Mikkonen, 2023). By prompting 'superhero' Midjourney may recreate image data connected to well-known

superheroes in our popular culture. One way to avoid the typical *Midjourneysce* style can be to ask for photographs, or other artistic styles. Adding a year or an era in the prompt will also produce results with more variations, both in style and in motif. Asking for black and white photos of female superheroes in the 1890s (Figure 9, left) gave more gritty results than just prompting for female superheroes. Adding older, elderly or geriatric to the prompt can result in more grown up or middle-aged people (Figure 9, right). To avoid both biases and results that look too much like existing imagery, a more detailed prompt is therefore recommended.

Figure 9. Black and white photo of two female superheroes. Left: 1890s, right: geriatric.

Ethics, privacy, copyright, and censorship

Who owns images created with AI? Being trained on large datasets from the internet, there is no guarantee that images will not violate any copyrights (Abdallah & Estévez, 2023; Hutson et al., 2024, p. 32). In a classroom setting one can explore and experiment with images without violating copyrights (Bergman, 2021). However, as Midjourney shares the generated images with the community and the creator alike, one can question which copyrights are potentially violated with each image generation. Making images in an AI-based process provides a natural ground for discussing ethical dilemmas of copyright infringements and obstacles of censorship. Due to different types of censorship, some of the AI models have certain constraints. With Midjourney, one cannot create imagery based on prompts that suggest sexual content, while materials potentially violating copyright or personal data issues seem to be accepted. Compared to Midjourney (Figure 10), Adobe Firefly is more restrictive regarding copyright issues as also Hutson pointed out (2024, p. 111). These limitations in technology due to censorship, copyright or privacy issues limit the user's freedom to express their prefigurative thoughts in visualisations.

Environmental issues need considering when using AI, like other internet-based technology such as cloud storage and search engines. The impact spans from mining minerals to coal driven energy consuming data centres and water consumption (Crawford, 2021). These issues are

complex but in using AI generators one should be aware of the impact this has on the environment. AI generators are also dependent on human repetitive underpaid work such as training the datasets through labelling (Crawford, 2021, p. 63).

Figure 10. Prompting a realistic Mickey Mouse in Midjourney (left) and Adobe Firefly (right)

If AI text-to-image generation is to be used in schools, there are several considerations when it comes to navigating the complex terrain of copyright and privacy regulations. Elementary schools seldom allocate funds for platforms shielded by paywalls. While numerous non-payment AI tools exist, our experience indicates that these often lack the comprehensive filters found in their paid counterparts. For instance, unintentional generation of images with sexual references. The non-payment generator Krea generated close to pornographic images as a result of the innocent prompt 'sunset on a beach'. Privacy concerns also arise when uploading images that might jeopardize personal privacy, such as faces or other identifiable data. This not only opens the door to potential digital bullying but also highlights the critical importance of privacy awareness. There have been incidents where students use AI tools to make fake, realistic, pornographic images of other students (Jargon, 2023). This deep-fake technology has rapidly advanced in the last years, making it possible to create highly realistic media content that convincingly mimics the appearances of individuals. This technology carries the potential for abuse, including the spread of false information and digital depiction of users (Karnouskos, 2020; Pennefather, 2023).

Non-payment AI tools tend to feature a considerable amount of advertising. Educational institutions are prohibited from exposing students to advertisements in class. To bypass these ads, users often need to register or subscribe, which requires personal information such as an email address. Should a student be required to disclose such information, school administrators must adhere to strict privacy regulations. Without proper agreements between the school and the data processor, the risk associated with privacy becomes considerably high. If not properly safeguarded, personal information pertaining to students and teachers could be exploited, leading to potential exposure to various risks. Our exploration has revealed that AI tools behind paywalls generally offer superior image resolution, more effective copyright and privacy filters,

absence of advertisements, and faster image generation. It presents a valuable opportunity to reflect on the importance and nuances of these matters in the digital age educational setting.

A cop-out?

The low-effort ease of image making with AI makes us question whether its extensive use can result in a non-critical view on the benefits of making. If the use of AI text-to-image technology replaces sensorimotor making activities such as drawing, what is possibly lost in a creation process consisting less of producing and more of articulating and selecting? Training is essential in traditional image-making processes consisting of applying sensorimotor techniques. Drawing skills are developed over multiple years of practice. If such sensorimotor skills for producing imagery based on prefigurative thinking become superfluous, what is the incentive in education to develop sensorimotor skills for mimetic drawing, as described by Nielsen (2013)? Is this ease of use a hindrance to learn or can it possibly free up more time to focus on other knowledges and skills necessary in a future learning environment? The use of AI technology in learning processes demands a critical teaching mindset, ensuring the necessary training and development of the skills needed.

Concluding remarks

This has been an initial study aiming to explore the use of AI text-to-image generators in a broad fashion. The study has identified several possible avenues for further research, such as how the potentials of the technology can be used in K-12 arts education. This include how AI text-to-image generators can stimulate the pupils' creative processes or how generated images can be combined with traditional crafts. However, with the widespread use and development of this technology, we emphasize the need for a critical perspective in future research.

Artificial intelligence text-to-image technology can contribute to and change future Art and design education in various ways. It can contribute to increased opportunities for training prefigurative thinking, providing new ways of visualising and co-creating. This can represent a democratisation of visualising prefigurative thinking, as creators are not being restricted by their limitations in skills or techniques. The use of AI can simplify and enrich image making, design and creative processes. Artificial intelligence text-to-image technology can represent a useful tool for creative processes and developing articulation for visualising the imaginary. Its use can also represent limitations to creativity and contribute to ethical questions and issues of bias being raised. With the application and use of AI text-to-image technology in art and design education spreading, we need to question what kinds of competencies are needed in future learning processes.

AI text-to-image generated image copyright statement by authors

All AI text-to-image generated images in this article were created using Midjourney, except one image using Adobe Firefly (Figure 10). To create images with Midjourney you will need a subscription plan. According to Midjourneys content rights, the creator of the images owns all created assets (Midjourney, n.d.-c.).

References

Abdallah, Y. K., & Estévez, A. T. (2023). Biomaterials Research-Driven Design Visualized by AI Text-Prompt-Generated Images. *Designs*, 7(2), 48. http://dx.doi.org/10.3390/designs7020048

Banas, J. A., Dunbar, N., Rodriguez, D., & Liu, S. J. (2011). A Review of Humor in Educational Settings: Four Decades of Research. *Communication Education*, *60*(1), 115–144. https://doi.org/10.1080/03634523.2010.496867

Bergman, S. (2021). Copying and cultural references as tools in image creation: Copyright challenges in teaching. *Nordic Journal of Art & Research*, 10(1). https://doi.org/10.7577/information.3692

Braun, V., & Clarke, V. (2022). *Thematic analysis: a practical guide.* SAGE.

Bunge, M. (1963). A General Black Box Theory. *Philosophy of Science*, 30(4), 346–358. http://www.jstor.org/stable/186066

Cetinic, E. & She, J. (2022). Understanding and Creating Art with AI: Review and Outlook. *ACM Transactions on Multimedia Computing, Communications and Applications*, 18(2), Article 66. https://doi.org/10.1145/3475799

Chen, L., Wang, P., Dong, H., Shi, F., Han, J., Guo, Y., Childs, P. R. N., Xiao, J., & Wu, C. (2019). An artificial intelligence based data-driven approach for design ideation. *Journal of Visual Communication and Image Representation*, *61*, 10-22. https://doi.org/10.1016/j.jvcir.2019.02.009

Crawford, K. (2021). *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press. https://doi.org/10.2307/j.ctv1ghv45t

Discord. (n.d.). *IMAGINE A PLACE...* Discord. https://discord.com/

Dyrssen, C. (2010). Navigating in Heterogeneity: Architectural Thinking and Art-Based Research. In M. Biggs, H. Karlsson (Eds.), The Routledge Companion to Research in the Arts (pp. 223–239). Routledge.

European Commission, Joint Research Centre, Redecker, C., Punie, Y. (2017). European framework for the digital competence of educators: DigCompEdu, (Y. Punie, ed.) Publications Office. https://data.europa.eu/doi/10.2760/159770

Fazel, M., & Almousa, S. (2021). Chance, Ambiguity, and Indeterminacy as Idea-Generating Mediums Applied in Creative Design: Encountering Uncertainty in Mediums of Drawing in the Design Process. In F. Bianconi & M. Filippucci (Eds.), *Digital Draw Connections: Representing Complexity and Contradiction in Landscape* (pp. 267–293). Springer International Publishing. https://doi.org/10.1007/978-3-030-59743-6_12

Fry, T. (2012). *Becoming Human by Design*. Bloomsbury Publishing Plc.

Gombrich, E. H. (1982). *The Image and the Eye*. Phaidon Press Limited.

Hart, T. (2017). Online Ethnography. In J. Matthes, C. S. Davis & R. F. Potter (Eds.) *The International Encyclopedia of Communication Research Methods*, 1-8. John Wiley & Sons. https://doi.org/10.1002/9781118901731.iecrm0172

Hutson, J., & Lang, M. (2023). Content creation or interpolation: AI generative digital art in the classroom. *Metaverse*, *4*(1), 13.

Hutson, J., Lively, J., Robertson, B., Cotroneo, P., Lang, M. (2024). *Creative Convergence: The AI Renaissance in Art and Design.* Springer Series on Cultural Computing. Springer Cham. https://doi.org/10.1007/978-3-031-45127-0

Jargon, J. (2023, 02. November) Fake Nudes of Real Students Cause an Uproar at a New Jersey High School. *The Wall Street Journal.* https://www.wsj.com/tech/fake-nudes-of-real-students-cause-an-uproar-at-a-new-jersey-high-school-df10f1bb?mod=tech_lead_story

Karnouskos, S. (2020) Artificial Intelligence in Digital Media: The Era of Deepfakes, in *IEEE Transactions on Technology and Society*, *1*(3), pp. 138-147, https://doi.org/10.1109/TTS.2020.3001312

Kelentrić, M., Helland, K., & Arstorp, A.-T. (2017). *Professional Digital Competence Framework for Teachers*. Senter for IKT i utdanningen.

Kleinman, Z. (2024, 28. February). Why Google's 'woke' AI problem won't be an easy fix. *BBC.* https://www.bbc.com/news/technology-68412620?trk=public_post_comment-text

Ko, H.-K., Park, G., Jeon, H., Jo, J., Kim, J., & Seo, J. (2023). *Large-scale Text-to-Image Generation Models for Visual Artists' Creative Works* Proceedings of the 28th International Conference on Intelligent User Interfaces, Sydney, Australia. https://doi.org/10.1145/3581641.3584078

Kunnskapsdepartementet [Ministry of Education and Research]. (2023). *Strategi for digital kompetanse og infrastruktur i barnehage og skole [Strategy for digital competence and infrastructure in kindergardens and school].* https://www.regjeringen.no/no/dokumenter/strategi-for-digital-kompetanse-og-infrastruktur-i-barnehage-og-skole/id2972254/

Liu, V., Vermeulen, J., Fitzmaurice, G., & Matejka, J. (2023). *3DALL-E: Integrating Text-to-Image AI in 3D Design Workflows* Proceedings of the 2023 ACM Designing Interactive Systems Conference, Pittsburgh, PA, USA. https://doi.org/10.1145/3563657.3596098

Lowenfeld, V. (1947). *Creative and Mental Growth*. Macmillan.

Lyu, Y., Wang, X., Lin, R., & Wu, J. (2022). Communication in Human–AI Co-Creation: Perceptual Analysis of Paintings Generated by Text-to-Image System. *Applied Sciences*, *12*(22). https://doi.org/10.3390/app122211312

Midjourney. (n.d.-a). *Discord Interface*. Midjourney. https://docs.midjourney.com/docs/midjourney-discord

Midjourney, (n.d.-b). *Model Versions*. Midjourney. https://docs.midjourney.com/docs/model-versions

Midjourney, (n.d.-c). *Terms of Service*. Midjourney. https://docs.midjourney.com/docs/terms-of-service

Mikkonen, J. (2023) Advent of GAN: How does a generative AI create a moodboard? In Holmlid, S., Rodrigues, V., Westin, C., Krogh, P. G., Mäkelä, M., Svanaes, D., Wikberg-Nilsson, Å (eds.), *Nordes 2023: This Space Intentionally Left Blank*, 12-14 June, Linköping University, Norrköping, Sweden. https://doi.org/10.21606/nordes.2023.114

Nielsen, L. M. (2013) Visualising ideas: a camera is not enough, in Reitan, J. B., Lloyd, P., Bohemia, E., Nielsen, L. M., Digranes, I., & Lutnæs, E. (eds.), *DRS // Cumulus: Design Learning for Tomorrow*, 14–17 May, Oslo, Norway. https://doi.org/10.21606/learnxdesign.2013.103

Norwegian Directorate for Education and Training. (2020). Curriculum for Art and Crafts (KHV01-02). The National Curriculum *Kunnskapsløftet* 2020. https://www.udir.no/in-english/curricula-in-english/

Oppenlaender, J. (2022). The Creativity of Text-to-Image Generation. In *25th International Academic Mindtrek Conference (Academic Mindtrek 2022)*, November 16–18, 2022, Tampere, Finland. ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3569219.3569352

Pennefather, P.P. (2023). *Creative Prototyping with Generative AI: Augmenting Creative Workflows with Generative AI.* Apress. https://doi.org/10.1007/978-1-4842-9579-3

Ringvold, T. A., Strand, I., Haakonsen, P., & Saasen Strand, K. (2023). AI Text-to-Image Generation in Art and Design Teacher Education: A Creative Tool or a Hindrance to Future Creativity? *The Proceedings for the 40th International Pupils' Attitudes Towards*

Technology Conference, Liverpool John Moores University, Liverpool, 31 October – 3 November. Retrieved from https://openjournals.ljmu.ac.uk/PATT40/article/view/1350

Schut, A., Klapwijk, R., Gielen, M., van Doorn, F., & de Vries, M. (2020). Uncovering early indicators of fixation during the concept development stage of children's design processes. *International Journal of Technology and Design Education*, 30(5), 951–972. https://doi.org/10.1007/s10798-019-09528-2

Thomas, D., & Brown, J. S. (2011). *A New Culture of Learning: Cultivating the Imagination for a World of Constant Change.* CreateSpace Independent Publishing Platform.

Vartiainen, H., & Tedre, M. (2023). Using artificial intelligence in craft education: crafting with text-to-image generative models. *Digital Creativity*, 34(1), 1-21. https://doi.org/10.1080/14626268.2023.2174557

Verma, P. (2023, 26 March). AI can draw hands now. That's bad news for deep-fakes. *Washington Post*. https://www.washingtonpost.com/technology/2023/03/26/ai-generated-hands-midjourney/

Winter, R., & Lavis, A. (2020). Looking, but not listening? Theorizing the practice and ethics of online ethnography. *Journal of Empirical Research on Human Research Ethics*, 15(1–2), 55–62.