Fashioning the Future: Unlocking the Creative Potential of Deep Generative Models for Design Space Exploration

Richard Lee Davis
École Polytechnique Fédérale de Lausanne (EPFL)
Switzerland
richard.davis@epfl.ch

Thiemo Wambsganss
École Polytechnique Fédérale de Lausanne (EPFL)
Switzerland
theimo.wambsganss@epfl.ch

Wei Jiang
École Polytechnique Fédérale de Lausanne (EPFL)
Switzerland
weijiang9715@gmail.com

Kevin Gonyop Kim
ETH Zürich
Switzerland
kevkim@ethz.ch

Tanja Käser
École Polytechnique Fédérale de Lausanne (EPFL)
Switzerland
tanja.kaeser@epfl.ch

Pierre Dillenbourg
École Polytechnique Fédérale de Lausanne (EPFL)
Switzerland
pierre.dillenbourg@epfl.ch

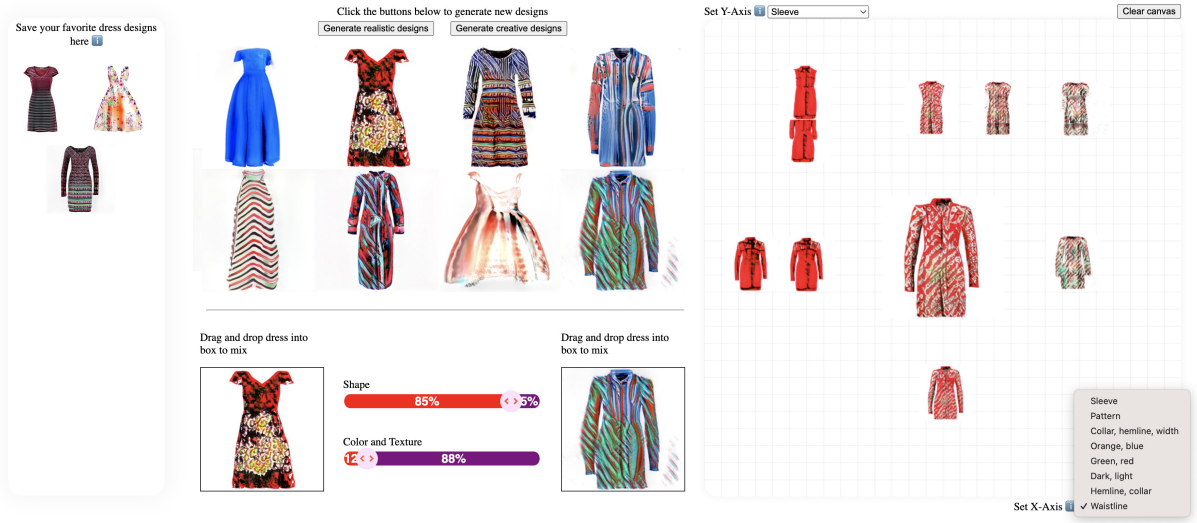

Figure 1: The current version of generative.fashion.

ABSTRACT

This paper investigates the potential impact of deep generative models on the work of creative professionals, specifically focusing on fashion design. We argue that current generative modeling tools lack critical features that would make them useful creativity support tools, and introduce our own tool, generative.fashion[1], which was designed with theoretical principles of design space exploration in mind. Through qualitative studies with fashion design apprentices, we demonstrate how generative.fashion supported both divergent and convergent thinking, and compare it with a state-of-the-art diffusion model, Stable Diffusion. In general, the apprentices preferred generative.fashion over Stable Diffusion, citing the features explicitly designed to support ideation. We conclude that the exploration and development of novel interfaces and interaction modalities that are theoretically aligned with principles of design space exploration is crucial for unlocking the creative potential of generative AI and advancing a new era of creativity.

[1]A live demo of the tool is available at https://generative.fashion.

CHI EA '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9422-2/23/04.
https://doi.org/10.1145/3544549.3585644

CCS CONCEPTS

• **Human-centered computing** → **Interactive systems and tools**; • **Computing methodologies** → **Artificial intelligence**.

KEYWORDS

Generative AI, Deep Generative Models, Creativity Support Tools (CSTs), Creativity, Design Space Exploration, Fashion Design, Ideation Process

ACM Reference Format:
Richard Lee Davis, Thiemo Wambsganss, Wei Jiang, Kevin Gonyop Kim, Tanja Käser, and Pierre Dillenbourg. 2023. Fashioning the Future: Unlocking the Creative Potential of Deep Generative Models for Design Space Exploration. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23), April 23–28, 2023, Hamburg, Germany.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3544549.3585644

1 INTRODUCTION

Deep generative models are neural networks that are capable of creating new things. Recent iterations of these models such as DALL·E [19], GPT-3 [1], and StyleGAN [11] have reached a point where the images, speech, and text that they generate is of such high quality that it is often indistinguishable from original work created by humans. These models typically require an enormous amount of data, compute, and technical know-how to train and run, which has meant that few people outside of academia or industry have been able to access or work with them. However, in the past year it has become trivial to generate content using large-scale generative models via web portals and APIs (e.g., ChatGPT, DALL·E), and highly-optimized open-source generative models can be downloaded, trained, and run on machines with consumer-grade GPUs (e.g., Stable Diffusion, StyleGAN2-ADA).

The wide availability of these models has sparked an intense debate about the future of creative work. One one side are those who argue that these models will lead to mass unemployment of illustrators, writers, and photographers. On the other side are those who view these models as the latest iteration of technologies that support creative practices, with accompanying arguments that creatives who learn to use the tools will have no trouble remaining employed. We were motivated by this debate to investigate the ways that deep generative models might impact the work of creative professionals.

Though we are sympathetic with the latter perspective, we argue that most deep generative modeling tools lack critical features that would make them useful creativity support tools. To explore this hypothesis, we created a deep generative modeling tool called generative.fashion whose design was aligned with theoretical principles of design space exploration [5]. We evaluated this tool in two qualitative studies with seven fashion design apprentices who were each actively designing their own clothing collections. In addition to working with our tool the apprentices also worked with Stable Diffusion [20], a state-of-the-art diffusion model capable of producing high-quality images of practically anything when given a text description.

In these studies we observed how the apprentices found useful ways of integrating both generative.fashion and Stable Diffusion into their creative practices. These tools did not displace their existing methods, such as using Internet search engines like Pinterest to find specific designs, but rather provided the apprentices with new possibilities for creative exploration and inspiration. While the two generative modeling tools offered complementary forms of support for the apprentices' creative activity, the apprentices mostly favored designs created with generative.fashion, and stated that generative.fashion provided better support for creativity through most phases of the ideation process. This provided support for our hypothesis that the exploration and development of novel interfaces and interaction modalities is paramount for unlocking the creative potential of generative AI and ushering in a new era of creativity.

2 BACKGROUND

Creativity support tools (CSTs) are broadly defined as digital systems that encompass one or more creativity-focused features which are deployed to positively influence one or more phases of the creative process [3]. Though there are competing theories about creativity, most research on creativity support tools adheres to a view that the creative process is comprised of distinct phases of pre-ideation/background research, ideation, implementation, evaluation, and iteration [3, 4], and this work typically involves the development and evaluation of tools meant to support specific phases of this creative process.

Ideation is the part of the creative process characterized by both divergent and convergent thinking. Divergent thinking involves the generation of a wide variety of ideas, while convergent thinking involves the selection and refinement of a small set of ideas. While these are typically viewed as distinct cognitive activities, both can be characterized as ways of searching through the design space of a domain. The design space of a domain is a "representation of the ideas and concepts that designers develop over time to propose a design solution that materializes into a design artifact" [5, p. 1]. Divergent thinking can be framed as a broad exploration through design space where a large quantity of different ideas are collected, while convergent thinking can be framed as search within a small area of the space that involves the gradual refinement of a small set of ideas.

One of the benefits of framing divergent and convergent thinking in this way is that it helps inform the design of creativity support tools for ideation. Supporting divergent thinking means acting as a guide in the exploration of the design space. The tool should help users identify where they are in the space, allow them to intentionally navigate through the space, and bring them to parts of the space that they are unaware of. These functionalities can help the user break the design fixation which occurs when they become trapped in a small part of the design space [7] and can support them in developing new and surprising ideas [12]. To support convergent thinking, these tools should allow the user to zero in on a specific part of the design space and to explore subtle variations between ideas.

Deep generative models such as GANs [18], VAEs [13], and diffusion models [19] have a unique set of properties that makes them especially well-suited for supporting design space exploration. When trained on large, representative datasets, these models build enormously detailed and complex representations of the design spaces in their learned latent spaces. However, exploration of these latent spaces is difficult due to their high-dimensionality and lack of clear structure. These spaces contain hundreds to thousands of entangled dimensions, which means interpolating an image along a single axis is likely to change multiple properties of the image. Additionally, it can be difficult to intentionally locate a region of space that produces images with desired properties, and even within

a very small region of this space there can be an enormous degree of variation.

Our tool, generative.fashion, provides technical solutions to these problems and packages them in a web-based graphical user interface that is designed for those with no prior programming experience (Figure 2). The features of generative.fashion were deliberately designed to support styles of design space exploration associated with both divergent and convergent thinking. Multiple ways of locating starting points for exploration were provided, ranging from divergent (randomly sampling points in the latent space) to convergent (locating and returning images with high similarity to a text description provided by the user). Once promising starting points were found, users could continue their exploration of the latent space by using the style mixing interface, which allowed users to select multiple points in the space and follow the paths between them, or by using the design canvas, where users could explore the latent space in directions corresponding to sleeve length, pattern, and color by dragging and dropping images in a two-dimensional grid. Whether the user wanted to make a dramatic change to the shape of a design while not altering the pattern, or to change the color without altering the shape, or to make a small alteration to the hemline, or to jump to new points in the space and generate entirely new designs, these features allowed them to do so. We present an overview of the technical details of these features in the following section, and a full treatment can be found in [9] and [8].

3 THE GENERATIVE.FASHION DESIGN SPACE EXPLORATION TOOL

Our primary goal in developing generative.fashion was to create a tool that could support users in intentionally and meaningfully exploring the latent space of a deep generative model. Achieving this goal required us not only to create an interface to the model with novel interaction modalities, but also to develop new features of the model itself. We followed general principles of design as outlined by Shneiderman et al. to ensure that the tool would foster "easy exploration, rapid experimentation, and fortuitous combinations that lead to innovations" (p. 70).

Our starting point was the StyleGAN2-ada model [10] trained on the Feidegger dataset [15]. Although the source code for this model provided a basic method for GAN inversion (i.e., locating a point in the latent space that best matched an input image), it did not provide any other way of intentionally locating points in the latent space aside from random generation. We implemented an additional way of locating a point in the GAN latent space by providing a text box where users could write a short description of a design which would be used to locate closely-matching images embedded in the latent space. Since this functionality was not a part of the StyleGAN2-ada code, we explored two methods for adding this capability to the model. The first method was to randomly sample images from the latent space, then to pass these along with the text description through a CLIP [17] model to find a small number of images which most closely matched the text. The second method was to fine-tune a DALL·E model [17] on the Feidegger dataset, and then to pass the text descriptions to DALL·E and let it generate designs. Surprisingly, we found that the two methods produced similar results and chose to use the first method because it was more efficient and straighforward.

After users had harvested a crop of images from the GAN latent space using random generation, image search, or text search, they needed a way to begin converging on specific ideas. We provided two distinct functionalities to aid in this process. The first, style mixing, allowed users to blend two images by interpolating between them in latent space. While this functionality was a pre-existing property of the StyleGAN2-ada model [11], we developed a novel user interface to expose the full power of this functionality to the user. Two generated images could be dragged and dropped into the visual style mixing interface, and sliders allowed the user to mix and combine features from each of the designs into a single example. The coarse slider controlled the shape of the output, and the fine slider controlled the pattern and color. The output image was shown in the center of the latent-space exploration panel on the right in Figure 2.

The second was the latent-space exploration panel, which allowed users to take any generated image and move it along meaningful directions in the latent space. This provided users with a way to intentionally and meaningfully explore the latent space of the GAN. Each axis of this two-dimensional canvas corresponded to a semantically-meaningful direction in the latent space, and the direction corresponding to each axis could be changed using a drop-down menu. Dragging and dropping an image within the canvas was equivalent to interpolating a point in the latent space along the directions selected in the drop-down boxes, and each newly-generated point was shown on the canvas as a history point. For a simplified representation of this interface see Figure 3. Behind the scenes, we used a method described in [6] to identify semantically-meaningful directions using PCA on the latent w space corresponding to sleeve length, pattern, color, hemline, waistline, and more. For more information on this method and for more examples of the results of interpolating along the different principal components, see Figure 4.

4 EVALUATING GENERATIVE.FASHION WITH FASHION DESIGN APPRENTICES

To investigate whether and how these features might be used to support ideation in an authentic context, we conducted two qualitative studies with a group of seven fashion design apprentices (3F, 4M, ages 17-25) studying in a Swiss university of applied sciences. The goal of the first study was to investigate the apprentices' current use of technology in their creative practice, to introduce the generative.fashion tool and give them the opportunity to use it in a design problem, to observe how they used the different features of the tool during the design process, and to discuss with them how the tool might fit into their existing creative practice. In the second study, we introduced the apprentices to Stable Diffusion, a large-scale diffusion model with enormous expressive power. While generative.fashion lacked the expressive power of these large-scale diffusion models, we were curious to see if the features we had built into the tool to support intentional and meaningful design-space exploration might support to the apprentices' ideation process in ways that the diffusion models could not. The details of these two studies are provided in the following sections.

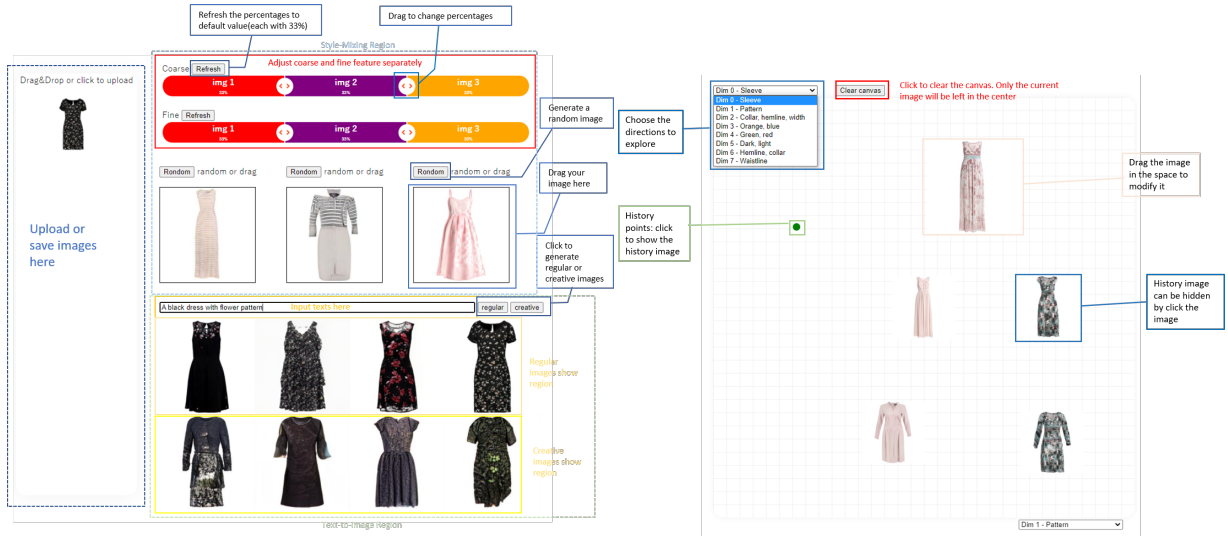

Figure 2: The initial version of the generative.fashion tool. Images could be initially generated via text descriptions using the text box. They could then be dragged to the style-mixing region or saved in the sidebar. Users could selectively combine elements from three designs using the visual style-mixing panel. The output image would be shown in the center of the canvas on the right. The 2D-dimensional canvas represents the design space with two meaningful attributes assigned to the horizontal and vertical axes. These attributes could be changed by using a drop-down menu for each axis. Dragging the image within the canvas was equivalent to moving through the latent space of the GAN in semantically meaningful directions.

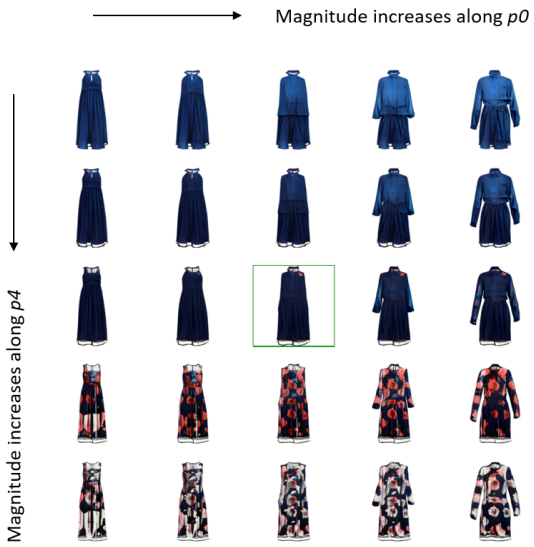

Figure 3: Examples of interpolating images simultaneously along two meaningful directions in the latent space (sleeve and pattern) found using PCA. The image in the green box shows the original image with 0 magnitude.

4.1 Study 1: User Testing generative.fashion with Fashion Design Apprentices

The apprentices were broken into three groups, with one researcher embedded in each of the groups. The groups were asked to use a collaborative whiteboarding tool to put together a research book containing one or more dress designs for a design persona. Each of the group members took turns using the generative.fashion tool for 10-15 minutes to create dress designs. While one group member was using generative.fashion, the others used their laptops to find and create dress designs using the tools and methods that they would normally use. Finally, after creating the research book, each of the apprentices created a final sketch of a dress for the client in accordance with a set of explicit guidelines[2]. While the apprentices worked on their projects, the researchers took detailed observational field notes on which features of the generative.fashion tool students used and how their use of these features evolved as the project progressed.

During the study each of the researchers led two semi-structured focus groups with the small group of 2-3 apprentices that they were embedded with. The focus group method [14] was used to investigate the apprentices' shared experiences related to technology use in their work as fashion designers, to surface their collective understanding of how the features of the generative.fashion tool might be different from the tools they were already using, and to allow the apprentices to collectively identify salient aspects of the tool. The first focus group took place at the midpoint of the study, after the apprentices had worked with generative.fashion and completed their research books. In this focus group apprentices were asked about the usefulness and usability of the tool, the creativity support provided by the tool, and how the tool compared to other tools and methods. The second focus group took place at the conclusion of the study after they had created their final sketches.

[2]The activity and materials were developed in consultation with the fashion design instructor to ensure their authenticity.

Figure 4: Examples of meaningful principle components p found in the latent space by PCA. (Dim k, Layers a-b) represents the k'th principle component applied in layers a to b in the 14-layer synthesis network. The images in green boxes are the original images with 0 magnitude. For each p, we show the result of two images, one in-sample and another out-of-sample.

In this focus group, the apprentices were asked to explain how their final sketches were influenced by the different tools (including generative.fashion) and asked to describe the activities and settings in which the tool might be most useful. During the focus groups, the researchers took detailed notes and acted as moderators with the goal of ensuring that all apprentices had the opportunity to share their thoughts and perspectives. This moderation approach was intended to compensate for some of the limitations of the focus group method related to individuals dominating the discussion [22]. Shortly after the conclusion of the study, the researchers conducted a debriefing session [16] to share and reflect on emergent findings.

The notes from the observation, the focus groups, and the debriefing session were analyzed using a hybrid process of deductive and inductive thematic analysis [2] to identify overarching themes. The notes were first analyzed according to a set of deductive codes derived from our research objectives, and during this deductive coding process emergent themes that surfaced were assigned inductive codes. Finally, these codes were grouped into a small number of themes that captured important aspects of the participants' experiences and ways of using the generative.fashion tool during the study. We elaborate on these themes in the following sections.

4.1.1 Better Support for Convergent than Divergent Exploration. Activities associated with the careful refinement of a single idea were coded as convergent, while activities that resulted in the generation of a number of new ideas were coded as divergent. Based on our observations, the apprentices mainly used the tool to support convergent ideation. When using the tool in a convergent manner, apprentices typically started with a single text prompt like "classic gray dress", viewed the resulting image, and then made a series of minor changes to the text in an attempt to tweak the output (e.g, adding the substring "with black buttons"). Then, the apprentices would either make small changes in the style-mixing area or skip the style-mixing area entirely, before moving on to the design canvas to generate a number of designs with minor variations. Each step of the process resulted in further refinements to a dress design, and once the apprentice reached the end of this process they would copy the final dress into the research book.

While all seven of the apprentices primarily used the tool in this manner, two of the apprentices shifted to using the design canvas feature in a divergent way as the project progressed. Instead of using the design canvas to refine a design, they shifted to dragging the dress across large distances. This resulted in the generation of dresses with strange forms and vivid colors which had little in common with their original designs.

Despite the fact that we did not observe much activity related to divergent ideation, the apprentices viewed the tool as one that could help them find inspiration and produce new ideas. One said, "If you want to mess around and get new ideas and inspiration, ...mess around for a few minutes you have something pop up out of nowhere". Another stated that "What I liked about the canvas is depending on where you drag the dress you get things that you haven't really seen in daily life, or in pictures".

4.1.2 Support for Intentionality and Sense of Ownership. When asked to compare the generative.fashion tool to their existing practices, which were using Internet search engines like Google and Pinterest to look for ideas, some apprentices said that the tool offered support that these existing tools did not. The primary benefit mentioned by apprentices was that generative.fashion made it possible to realize one's own ideas, which provided more of a sense of ownership over the dress designs. One apprentice stated that first-year designers should use this tool instead of Google, since the tool would allow them to "to get their creativity and images in their head more refined". This would allow novices to "build a foundation for their own creativity", allowing them to develop a stronger identity before being influenced by others' work online.

4.1.3 Proposed changes to the system. The primary criticism voiced by all of the apprentices was that the output image quality was too low. They compared the tool's output to that of Internet search engines, saying "on Google or Pinterest it's possible to find extremely accurate images if you know what to search for". Some apprentices

said the images were too small and blurry, while others were unsatisfied with the system's inability to produce fine details such as specific types of buttons or patterns.

4.2 Study 2: User Testing generative.fashion and Diffusion Models with Fashion Design Apprentices

Between the first and second study a number of changes were made to the generative.fashion tool to better support divergent ideation and to improve usability. The primary change was that the text-prompt was replaced with two buttons that would randomly generate "traditional" or "creative" designs by sampling from smaller or larger volumes of the latent space. This feature was meant to bootstrap divergent thinking by presenting users with a variety of designs sampled from random points in the GAN latent space at the very start of the design process. Additionally, the style-mixing panel was simplified by reducing the number of designs to be mixed from three to two and the position of different features was rearranged to better indicate the intended workflow. The second version of the generative.fashion tool can be seen in Figure 1 and a live demo of the tool can be found at https://generative.fashion.

In addition to evaluating the impact of these changes on the apprentices creative practices, we were also interested in performing a more direct comparison between our tool and Stable Diffusion, an extremely powerful, expressive, and general generative model. Our choice to introduce this tool was motivated by criticisms of generative.fashion related to low image quality and inability to produce highly-accurate images using text prompts. While the Stable Diffusion model provided solutions to these problems, it lacked features provided by generative.fashion that were specifically designed to support intentional design-space exploration. By comparing the two tools, we hoped to learn whether the unique features of generative.fashion could offer advantages during the ideation process over the more general, text-based interface to the Stable Diffusion model.

The same seven fashion design apprentices who took part in the first study also took part in the second study. The apprentices were split into two groups, with one researcher embedded within each group. The apprentices within a group did not collaborate with one another, but worked individually on their design collections.

The study took place over three hours. The initial activity that the apprentices took part in was spending 15 minutes sketching an initial design for their collection. After completing this sketch they moved on to working with the tools to generate ideas. One group of apprentices worked with the generative.fashion tool and the other group worked with the diffusion modeling tool. After 45 minutes had passed, the apprentices were asked to stop using the tool and to sketch a new design inspired by their work with the tool.

In the next phase the apprentices swapped tools. If they had been working with the generative.fashion tool, they switched to working with the diffusion model (and vice versa). Again, they spent 45 minutes generating ideas, after which they spent 15 minutes sketching a new design. Finally, all of the apprentices displayed their three drawings on a central table for a gallery walk facilitated by the fashion design instructor.

The data collection and analysis methods were the same as those used in Study 1. Researchers took observational field notes during the activities, and led three focus group discussions during the study. The first and second focus groups were conducted at the group level and took place after the apprentices worked with one of the tools. These semi-structured discussions were focused on the usefulness and usability of the tool, the creativity support provided by the tool, and how the tool compared to other tools and methods. The final focus group discussion took place with the entire class after the gallery walk. In this discussion apprentices were asked to explain how the different sketches were influenced by the tools, to talk about how the tool might have helped them come up with ideas that they wouldn't have otherwise discovered, and to describe the activity and context in which they would use these tools again. After the study concluded, the researchers debriefed to compare notes and surface insights from their observations and focus group interviews, and the full set of notes were analyzed using the a hybrid process of deductive and inductive thematic analysis.

4.2.1 Supporting Both Convergent and Divergent Exploration. In contrary to the first study where the apprentices mainly used the generative.fashion tool for convergent exploration, in this study we observed a better balance between convergent and divergent exploration approaches. The apprentices used the random-generation functionality to produce wider varieties of dress designs, and then used the design canvas to explore large volumes of the design space (Figure 5). Many of the apprentices stated that the tool's support for divergent exploration was useful in the context of their projects. One apprentice said using the tool "made my brain go places it hadn't gone before", and another said, "it was useful for my project…it helps for getting outside of a thought box". The apprentices agreed that the designs produced by generative.fashion were surprising, saying "it produced mega-creative patterns" and "the tool produced infinitely many possibilities".

Opinions about the tool's support for convergent thinking were more varied. Some apprentices found that the tool offered enough control to hone in on a specific idea, while others were frustrated with their inability to "get where [they] wanted to go".

4.2.2 Comparing generative.fashion to Stable Diffusion. During the gallery walk, each apprentice took turns presenting their three sketches and explaining which elements were inspired by the use of the different tools. When asked which sketches they were happiest with, six out of the seven apprentices indicated that the drawing created with the generative.fashion tool was their favorite, and stated that they preferred working with generative.fashion over Stable Diffusion. When these apprentices were asked to explain why they preferred using generative.fashion, they provided a number of reasons. Some apprentices valued the ability to explore different patterns and colors without modifying the form of the dress, while others found that the tool made it easier for them to explore different silhouettes and forms. In contrast to generative.fashion, the version of Stable Diffusion used by the apprentices did not provide ways to make these kinds of fine-grained changes to the garments since desired adjustments required the apprentices to input a modified text prompt, which would generate a completely new output image.

A number of apprentices found inspiration in the surprising details generated by generative.fashion, such as pointy shoulders,

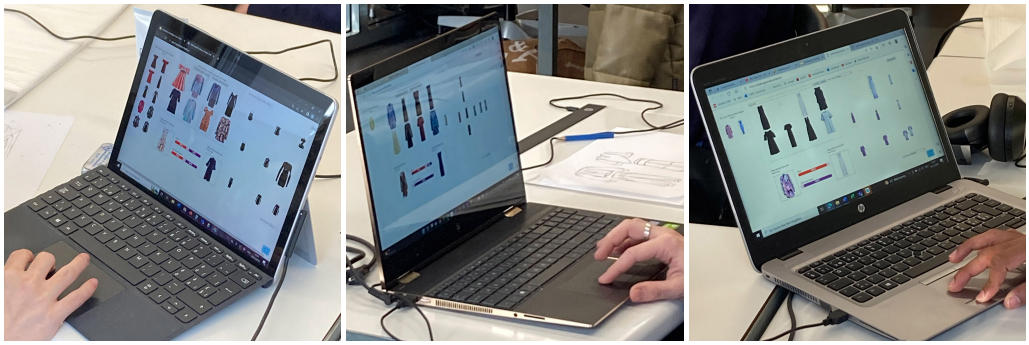

Figure 5: Three different apprentices using the generative.fashion tool to conduct divergent exploration of the design space. Note that the images generated in the design canvas on the right of the screen cover a large area, which corresponds to a large volume of the latent space.

irregular folds and cuts, and spiky sleeves. Many of these details were included in the apprentices' sketches and were the features that the apprentices liked the most. One apprentice said, "I really like this dress because of the shape and the line down the middle…which was inspired by the tool". In contrast, the apprentices found the outputs produced by Stable Diffusion less inspiring. One apprentice said, "I felt like it was less creative because it was so specific, there wasn't much to change. No room for imagination because it was so accurate." Another said, "when I looked up something I got what I expected, nothing unexpected".

However, one aspect of Stable Diffusion was felt to provide advantages over generative.fashion: the images Stable Diffusion produced were of higher quality and the output was more accurate than generative.fashion. One apprentice stated, "I can be more detailed with specific things like buttons, pockets, colors…It's accurate, you can even look up brands and not recognize any of the pieces but it fit the aesthetic". Another apprentice explained that Stable Diffusion might be useful to further refine specific aspects of ideas created using the generative.fashion tool, saying "If I specifically needed a pocket or sleeve, I would maybe use Stable Diffusion because there are more specific images there".

5 DISCUSSION

We distill our findings from these two studies into three insights related to the use of deep generative models for creativity support. All of these insights are concerned with different ways of controlling the stochasticity or unexpectedness of the generative model's outputs to support specific types of ideation activity.

5.1 Exposure to unexpected regions of the design space supports divergent ideation

At the start of the divergent thinking phase, we found that the apprentices appreciated when the model produced unpredictable or surprising outputs, and that these sorts of outputs were rarely produced via text prompting. With both generative.fashion and Stable Diffusion, using text prompts appeared to short circuit the divergent ideation process. With Stable Diffusion, the apprentices explicitly said that the outputs produced via text prompting were too accurate, "leaving no room for imagination", and that the model produced "nothing unexpected". And while generative.fashion could not match the accuracy of Stable Diffusion, we observed that apprentices who started by inputting text prompts mostly skipped over the divergent thinking phase entirely. However, after replacing the text-prompt in the generative.fashion tool with buttons for randomly sampling and generating images from the GAN latent space, apprentices engaged in more activities associated with divergent ideation and explicitly stated that the unexpected outputs were inspiring.

These findings indicate that it is important to provide users with ways of rapidly generating multiple outputs from a large volume of the design space, since this will aid them in finding interesting and unexpected regions of the design space in which to continue their exploration. Text prompts may be ill-suited for this task since it is challenging for a user to write a description of a design that they aren't expecting to find. Put differently, it is not reasonable to expect that a user can describe a region of design space that they don't know exists.

For generative.fashion, it was trivial to implement features that could expose users to new regions of the design space. This was because the latent space of the underlying generative model closely corresponded to the dress design space, which meant that randomly sampling points from the latent space would reliably produce recognizable dress designs. Implementing such a feature remains an open challenge for high-capacity, general-purpose diffusion models such as Stable Diffusion. While small regions of the Stable Diffusion latent space may correspond to the design space of different domains, it is not clear how to define their location such that randomly sampling from these subspaces would consistently produce images corresponding to a given design space. Paradoxically, this suggests that smaller, less powerful models trained on content from a specific domain may provide better support for divergent ideation than larger, more expressive models.

5.2 Control over model stochasticity supports the transition from divergent to convergent ideation

After identifying promising regions in the design space for further exploration, users should be able to set constraints on model stochasticity that support intentional and meaningful exploration of those regions. For the generative.fashion tool, these constraints took two forms. First, users were able to choose meaningful directions in the design space to explore, and second, they were able to control the size of the steps that they took in these directions. These constraints made it possible for users to intentionally explore individual regions of the design space in the design canvas, and to explore the design space between regions of interest by using the style-mixing panel.

In the design canvas, users were able to select meaningful directions in the design space to explore by assigning properties such as sleeve length, pattern, color, hemline, and neckline to the x- and y-axes. While moving an image along one of these axes, model stochasticity was tightly constrained to only affect the property of the dress that the user wished to change. Additionally, the user could control the amount of stochasticity applied to this property of the dress by moving the image over larger or smaller distances. In practice, we found that the apprentices used these features in two distinct ways. First, apprentices used these features to map out regions of the design space by dragging and dropping images across large areas of the design canvas (see Figure 5 for three examples of how apprentices in Study 2 used the design canvas in this way). Second, they used these features to hone in on a design by exploring small areas in the design canvas, which resulted in increasingly similar designs with small variations.

The version of Stable Diffusion the apprentices used did not provide the ability to intentionally move in meaningful directions of the design space. To change aspects of a design, an apprentice had to tweak the text description and submit this to the model, which would generate an entirely new batch of images. With no way to pin down specific aspects of a generated image such as the form, color, or pattern, intentional exploration of the design space was unpredictable. Apprentices explicitly mentioned this as a downside of the Stable Diffusion model, and asked for the ability to mix multiple outputs or pin down specific aspects of generated images.

Based on these findings, we argue that it is important to provide ways of exploring the latent space of a generative model by changing specific elements of the model's output without modifying other aspects of the generated image. These features support the transition from divergent to convergent ideation as users move from mapping out a region of the design space to honing in on a specific design.

5.3 High-quality, predictable outputs support convergent ideation

Finally, high-quality, precise model outputs are useful during the final phase of the convergent thinking process. In contrast to generative.fashion, Stable Diffusion produced clearer, higher-resolution images, and when provided with detailed prompts it produced designs with more accurate details and styles. Some apprentices stated that this made Stable Diffusion more useful when focusing on smaller details, such as a pocket or a sleeve.

5.4 Limitations and Next Steps

These findings are preliminary, as there are several limitations that limit their generalizability and validity. First, we did not triangulate our findings based on observational field notes with interaction data collected by the system. Using interaction data would provide us with a second source of data related to which features the apprentices used more frequently, the order in which they used different features, and the ways in which they used these features as they moved through the ideation process. Second, the sample size (N=7) places limits on the generalizability of our findings. Finally, while observational data indicated that certain features of the tool supported specific forms of ideation, a controlled experiment would be needed to firmly establish causal links. These limitations suggest future work that explores the specific impact of different features on creative practices with larger groups of users in different contexts.

6 CONCLUSION

Deep generative models can play an important role in supporting the work of creative professionals, but current tools lack critical features that could unlock their potential. These models are uniquely capable of learning vast and complex representations of design spaces, but users lack intuitive ways of exploring these spaces in intentional and meaningful ways. When augmented with features that provide users with ways of controlling the stochasticity of the model's outputs, these models are better able to support both convergent and divergent ideation. In the generative.fashion tool, we implemented features which constrained model outputs in ways that were aligned with theories of design space exploration and found that these features were successful in supporting fashion design apprentices throughout the ideation process. Additionally, the apprentices stated that they preferred generative.fashion over Stable Diffusion for most aspects of the ideation process, despite the fact that the underlying model for generative.fashion was far less powerful. These findings provide support for our hypothesis that unlocking the potential of deep generative models for creative support depends on the development of interfaces and functionalities that are specifically designed to support design space exploration, and provides some assurance that the features we built into generative.fashion were successful in providing this support. We hope to bring attention to our theorized connection between the learned latent space of deep generative models and the design space of a domain, and view the development of tools grounded in this theory as a promising area for future research on creativity support tools and design space exploration.

ACKNOWLEDGMENTS

Thanks to Carmen, Peter, and the apprentices who took part in our studies.

REFERENCES

[1] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel Ziegler, Jeffrey Wu, Clemens Winter, Chris

Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language Models Are Few-Shot Learners. In *Advances in Neural Information Processing Systems*, Vol. 33. Curran Associates, Inc., 1877–1901.

[2] Jennifer Fereday and Eimear Muir-Cochrane. 2006. Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. *International Journal of Qualitative Methods* 5, 1 (March 2006), 80–92. https://doi.org/10.1177/160940690600500107

[3] Jonas Frich, Lindsay MacDonald Vermeulen, Christian Remy, Michael Mose Biskjaer, and Peter Dalsgaard. 2019. Mapping the Landscape of Creativity Support Tools in HCI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–18. https://doi.org/10.1145/3290605.3300619

[4] Jonas Frich, Michael Mose Biskjaer, and Peter Dalsgaard. 2018. Twenty Years of Creativity Research in Human-Computer Interaction: Current State and Future Directions. In *Proceedings of the 2018 Designing Interactive Systems Conference*. 1235–1257.

[5] John Gero and Julie Milovanovic. 2022. Creation and Characterization of Design Spaces. In *DRS2022: Bilbao*. https://doi.org/10.21606/drs.2022.265

[6] Erik Härkönen, Aaron Hertzmann, Jaakko Lehtinen, and Sylvain Paris. 2020. Ganspace: Discovering Interpretable Gan Controls. *Advances in Neural Information Processing Systems* 33 (2020), 9841–9850.

[7] David G. Jansson and Steven M. Smith. 1991. Design Fixation. *Design Studies* 12, 1 (Jan. 1991), 3–11. https://doi.org/10.1016/0142-694X(91)90003-F

[8] Wei Jiang, Richard Lee Davis, Kevin Gonyop Kim, and Pierre Dillenbourg. 2021. Neural Design Space Exploration. In *35th Conference on Neural Information Processing Systems (NeurIPS 2021)*.

[9] Wei Jiang, Richard Lee Davis, Kevin Gonyop Kim, and Pierre Dillenbourg. 2022. GANs for All: Supporting Fun and Intuitive Exploration of GAN Latent Spaces. In *NeurIPS 2021 Competitions and Demonstrations Track*. PMLR, 292–296.

[10] Tero Karras, Miika Aittala, Janne Hellsten, Samuli Laine, Jaakko Lehtinen, and Timo Aila. 2020. Training Generative Adversarial Networks with Limited Data. *Advances in Neural Information Processing Systems* 33 (2020), 12104–12114.

[11] Tero Karras, Samuli Laine, and Timo Aila. 2019. A Style-Based Generator Architecture for Generative Adversarial Networks. *arXiv:1812.04948 [cs, stat]* (March 2019). arXiv:1812.04948 [cs, stat]

[12] Kevin Gonyop Kim, Richard Lee Davis, Alessia Eletta Coppi, Alberto Cattaneo, and Pierre Dillenbourg. 2022. Mixplorer: Scaffolding Design Space Exploration through Genetic Recombination of Multiple Peoples' Designs to Support Novices' Creativity. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3491102.3501854

[13] Diederik P. Kingma and Max Welling. 2013. Auto-Encoding Variational Bayes. *arXiv preprint arXiv:1312.6114* (2013). arXiv:1312.6114

[14] Richard A. Krueger. 2014. *Focus Groups: A Practical Guide for Applied Research*. Sage publications.

[15] Leonidas Lefakis, Alan Akbik, and Roland Vollgraf. 2018. Feidegger: A Multi-Modal Corpus of Fashion Images and Descriptions in German. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*.

[16] Shannon A. McMahon and Peter J. Winch. 2018. Systematic Debriefing after Qualitative Encounters: An Essential Analysis Step in Applied Qualitative Research. *BMJ Global Health* 3, 5 (Sept. 2018), e000837. https://doi.org/10.1136/bmjgh-2018-000837

[17] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, and Jack Clark. 2021. Learning Transferable Visual Models from Natural Language Supervision. In *International Conference on Machine Learning*. PMLR, 8748–8763.

[18] Alec Radford, Luke Metz, and Soumith Chintala. 2015. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *arXiv preprint arXiv:1511.06434* (2015). arXiv:1511.06434

[19] Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever. 2021. Zero-Shot Text-to-Image Generation. In *International Conference on Machine Learning*. PMLR, 8821–8831.

[20] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. 2022. High-Resolution Image Synthesis with Latent Diffusion Models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 10684–10695.

[21] Ben Shneiderman, Gerhard Fischer, Mary Czerwinski, Mitch Resnick, Brad Myers, Linda Candy, Ernest Edmonds, Mike Eisenberg, Elisa Giaccardi, Tom Hewett, Pamela Jennings, Bill Kules, Kumiyo Nakakoji, Jay Nunamaker, Randy Pausch, Ted Selker, Elisabeth Sylvan, and Michael Terry. 2006. Creativity Support Tools: Report From a U.S. National Science Foundation Sponsored Workshop. *International Journal of Human–Computer Interaction* 20, 2 (May 2006), 61–77. https://doi.org/10.1207/s15327590ijhc2002_1

[22] Janet Smithson. 2000. Using and Analysing Focus Groups: Limitations and Possibilities. *International Journal of Social Research Methodology* 3, 2 (Jan. 2000), 103–119. https://doi.org/10.1080/136455700405172