"Imagine a Dress": Exploring the case of task-specific prompt assistants for text-to-image AI tools

Helen Weixu Chen*
w352chen@uwaterloo.ca
University of Waterloo
Waterloo, Ontario, Canada

Lesley Istead*
LesleyIstead@cunet.carleton.ca
Carleton University
Ottawa, Ontario, Canada

(a) swiss cheese (b) gold brocade with a black belt (c) surreal fire (d) shards of glass

Figure 1: Dresses generated from prompts created through a task-specific prompt assistant.

ABSTRACT

In this paper, we explore the impact of task-specific prompt assistants for text-to-image generative AI tools through a user study. Participants were asked to recreate a dress with SDXL using either a prompt assistant tailored to the dress design, or, no assistant at all. A detailed analysis of the results and feedback suggests that for this specific task, a tailored assistant improves result satisfaction and accuracy. This style of assistant helps users focus on the task by providing a detailed, visual and organized approach to describing the object—enabling faster production times and more accurate descriptions with less ambiguity.

CCS CONCEPTS

• **Computing methodologies** → **Visual content-based indexing and retrieval**; **Natural language generation**; • **Human-centered computing** → **User studies**; • **Applied computing** → **Media arts**.

KEYWORDS

generative AI, prompt engineering, image generation, fashion

*Both authors contributed equally to this research.

GI '24, June 03–06, 2024, Halifax, Nova Scotia

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM
https://doi.org/XXXXXXX.XXXXXXX

ACM Reference Format:
Helen Weixu Chen and Lesley Istead. 2024. "Imagine a Dress": Exploring the case of task-specific prompt assistants for text-to-image AI tools. In *Proceedings of Graphics Interface 2024 (GI '24).* ACM, New York, NY, USA, 8 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Text-to-image generative AI tools such as Stable Diffusion, DALL-E3, and DeepFloyd IF, have flooded the internet over the last two years. These tools provide anyone, including those without artistic talent, the ability to create images beyond their imagination effortlessly in a few seconds. Where previously designers and artists spent hours translating ideas into visuals, they can now materialize sophisticated works with a few keystrokes. In essence, generative AI tools have opened a door to uncharted territories of creativity and innovation, transforming average users into artists and significantly streamlining the creative process in professional settings.

While these tools seem straightforward to use, the challenge lies in crafting the perfect prompt to generate the desired image, a task commonly referred to as "prompt engineering." Prompt engineering, however, is not a one-size-fits-all endeavour. What works for one genre might not yield the same results for another. For instance, generating fine art paintings, such as impressionism, expressionism, fauvism, or cubism, demands a vastly different set of keywords than generating haute couture dress concepts.

To assist with prompt creation, many have turned to ChatGPT for assistance. However, using tools such as ChatGPT to produce prompts relies on the user's ability to describe their needs—replacing the need to produce one prompt, with another. In order to produce prompts for specific tasks, such as dress design, or car design, the user needs to know many details and specific language for the task. The average user may not know the required terminology, and even individuals experienced with the topic may struggle to formulate it as text. For many, using ChatGPT to assist with prompt engineering, may be unapproachable.

As previously mentioned, generative AI tools can be used for more than just the creation of general art—they can be used to generate very specific things, like dress designs, video game textures, or even cars. For specific tasks, it is common to produce generative AI tools for that task. Some of these AIs were trained on highly specific datasets. For example, *The Little Black Dress*, is a generative AI tool for generating new fashion designs [21]. While other generative AI tools combine standard models, like Stable Diffusion, with additional training data and guidance models or a LoRA to perform the task, for example, Toyota's vehicle generation models [1]. These approaches have limitations in that they may stifle creativity, and/or be challenging to produce—requiring both vast amounts of expertise, and, computational power to produce the AI model. We argue, that for many tasks, like fashion designs, video game texture generation, and human face/body generation, standard models, such as SDXL, are likely sufficient—but the user needs some form of assistance with prompt engineering to produce fast and high-quality results.

In this paper, we explore how a task-specific prompt engineering tool can assist users in producing visuals for a specific task. In our experiment, we selected dress design, as our specific task. Participants were asked to reproduce a dress from a photograph using either our prompt assistant, or, without any assistance and then asked to evaluate both the result and their experience.

2 RELEVANT WORK

Generative AI technologies are capable of producing a wide range of content, including text, audio, images, videos, and three-dimensional models, among others [14]. ChatGPT from OpenAI generates text in response to user requests [7, 20]. DALL-E3, Stable Diffusion, SDXL, and DeepFloyd-IF are all text-to-image generative models that produce images based on textual user input [2, 5, 6, 8, 10, 12]. These tools continue to grow and improve as their linguistic understanding and vocabulary improve. However, all of these tools require users to input a text-based **prompt** which can be very challenging to produce.

2.1 Task-Specific Generative AI

Text-to-image generative AI tools can be used to produce images of nearly anything. However, some entities use these tools to assist in the design process for a very specific task. For example, video game artists might use these tools to rapidly produce textures for models. Or, a car company, might use a generative AI tool to imagine new car designs as inspiration.

Some organizations and groups construct their own generative AI models, or, provide additional training and guidance models, such as LoRAs, to existing AI models to create customized utilities for generating specific results such as textures, and dress or car designs.

"The Little Black Dress", for example, is a generative AI model trained on vintage and modern patterns along with patterns found in fashion magazines. This model produces imaginative new fashion designs that can be used to inspire dressmakers [21]. Similar models exist for face and body creation, along with body animations. These human models were trained using only relevant data—that of faces and human bodies and they provide a web-based interface to specify any of the desired traits, such as gender, eye colour, and age in the resultant image [18].

However, one of the downsides of a model that is trained on highly specific data, is that it has limited creativity. While "The Little Black Dress" can produce a seemingly infinite number of new, never-seen-before fashions, it cannot produce fashions made of novel materials, themes, or shapes that don't exist within the models training parameters (i.e., a dress made out of hamburgers).

Another common approach is to add customization to existing models and/or provide specific guidance models. For example. Toyota is in the process of developing a "drag-aware" generative car design tool that uses Stable Diffusion and a custom guidance model [1]. Producing these guidance models requires significant experience and can be costly.

Instead, we propose, that for some tasks, like generating textures, or, inspiring dress designs, that standard generative AI models like SDXL are not only sufficient but permit oddly creative and inventive combinations. The barrier to using these tools for a specific purpose is the creation of prompts.

2.2 Prompt Engineering

Creating a prompt that results in a satisfactory image can be challenging [4, 11, 16].

Prompts contain two parts, a description of the subject or image contents, and, a number of prompt modifiers that describe the desired style, visual properties, and quality. These prompt modifiers are essential in producing high-quality, pleasing output. However, inexperienced users of text-to-image tools may be unaware these modifiers are needed, or unsure of which ones will assist in producing a pleasing image. Hence, prompt engineering can be exceptionally overwhelming [11].

To assist users with prompt engineering, a large number of guides exist across the internet, released shortly after the introduction of tools like Midjourney and DALL-E2 [9, 15]. Oppenleander further investigated prompt modifiers, breaking them into various classes such as "style modifiers" that describe the visual style (e.g., "impressionist", "by Stephen Cosgrove"), and "quality boosters" that affect the output quality (e.g., "high quality", "4K") [16]. While Liu and Chilton conducted a detailed study on the impact of various prompt modifiers and model hyperparameters (i.e., number of generations) on the coherency and quality of the output image [11]. Both Oppenleander and Liu present prompt generation guides based on their work. From their results, both papers presented design guidelines for users to use when writing prompts. These guides can help users with subject descriptions, prompt modifiers, and keyword ordering. However, this "suggested reading" takes time to understand and these guides do not offer automated or assisted prompt creation. These documents and websites do not offer automatic or assisted prompt creation.

Aside from guidelines, there exist tools to assist with the automatic generation of prompts. Cao et al presented "BeautifulPrompt", which used the ChatGPT formula to generate training prompts for their fine-tuned prompt-generation model. Reinforcement learning was used to reward prompts that produced the highest-quality, most-pleasing images [4]. This paper is built upon previous work by Pavlichenko called "BestPrompt", a human-assisted genetic algorithm that identifies combinations of subjects and keywords to produce quality images [17].

However, these general guides and assistants may not produce the best prompts for a particular task because they would lack the specificity and terminology of that task. Furthermore, the cost of building guidance models may be too high for many groups.

Another approach used by several to generate large volumes of prompts for training is to create a formula for ChatGPT or other LLM that will produce prompts [4]. These formulas instruct the LLM on how to interpret user input and produce a properly formatted prompt output complete with modifiers.

To test this, we created a formula in ChatGPT to generate dress design prompts. This formula included rules for each design feature along with conditional logic, prompt modifiers, and an output grammar for the final prompt. While the formula worked in general, we observed a few flaws.

- It assumes the user knows all of the appropriate terminology and needs for dress designs.
- The output prompt grammar was not strictly adhered to, ChatGPT occasionally added items such as colour or length when these were not provided.
- We found that ChatGPT did not understand some design elements, such as odd or imaginative detailing, and hence these were not entered into the grammar correctly. They were often ignored, or appended in odd places.

Due to the strict and detailed nature of this task, *we did not see any particular advantage to using ChatGPT over a web-based form as a prompt assistant.* Furthermore, we note that any LLM-based assistant relies on the users' knowledge of the subject to form the prompt. While these LLM-based assistants can provide lists of options for different design details, along with examples, to the user, tools like ChatGPT do not provide the option of giving embedded image or video assistance. Furthermore, from a development perspective, while both an LLM formula and a web-based form assistant would require careful thought to the generated prompt syntax, using an LLM formula has an additional challenge of producing text that the LLM will be able to follow precisely. Hence, we propose that for task-specific prompts, a web form-based assistant is both simpler for developers to produce and easier for users to articulate the task.

3 METHOD

To investigate the impact of a task-specific prompt assistant, we first chose a task—designing dresses. We then produced two web applications, the first application is a simple, but standard, text-to-image AI interface with text boxes for positive and negative prompts (as seen in Figure 2).

The second application is a web form, which is our prompt engineering tool. We noted that specific tasks, such as dress, clothing, vehicle, etc., design can be roughly broken down into a decision tree of individual construction or design elements. For example, a car has wheels of a particular size and a dress is made of one or more types of material. Our prompt engineering assistant enumerates the different dress design options such as silhouette, material, and sleeve length. This web form provides visual reference radio buttons for each of the design options, but also provides text boxes for manual input should the user want a more customized dress (e.g., a dress made out of Swiss cheese). At the bottom of the web form is a button used to "Submit" the prompt, which results in the creation of both positive and negative prompts (including various prompt modifiers, silently added) that are then submitted to text-to-image AI for image generation. Figure 3 shows our web form-based assistant.

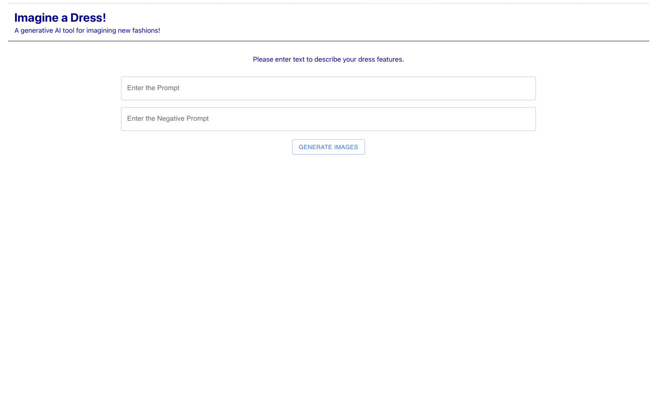

Figure 2: A visualization of the user interface for application B

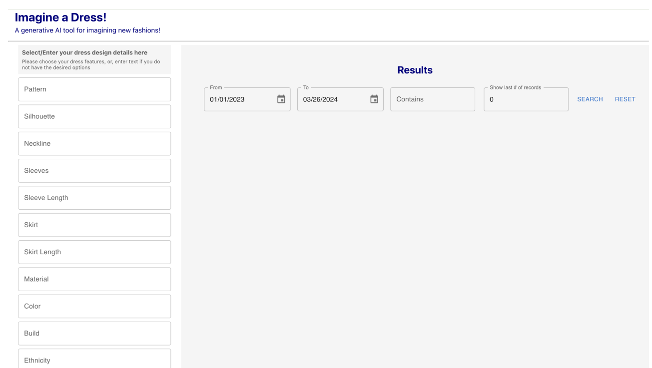

Figure 3: A visualization of the user interface for application A

For both applications, we use the SDXL model for image generation, and four results are generated and displayed to the user for each prompt.

3.1 Study Design

To study the impact of the prompt assistant, we conduct an experiment using the two aforementioned web applications. Participants were divided randomly into two groups. Both groups were asked standard demographic information, but also, of their experience with generative AI tools. Participants were then provided with five images of dresses and asked to choose one (as seen in Figure 4). They were asked to recreate the selected dress to the best of their ability using text-to-image generation.

Figure 4: The five image options, labelled I to V, were provided to participants as references for a generation.

Participants randomly assigned to Group A were directed to use our web form prompt assistant to generate the selected dress, while participants in Group B were directed to the control application, which provides only text boxes for positive and negative prompts as described previously.

Participants were then asked about their satisfaction with the result, how long it took them to achieve that result, and how closely they think it matches the selected source dress using scales from 1-5.

Following the dress recreation, participants were asked to create any dress of their choice. Participants in Group A were asked how useful they felt the prompt assistant was on a scale from 1-5. Participants in Group B were asked how difficult, on a scale from 1-5, they felt producing a text prompt for the task was.

Both groups were asked to upload both their prompts and their preferred image for each task. They were then asked to provide any feedback if they had any.

4 PARTICIPANT DEMOGRAPHICS

Participants were recruited through social media, in particular, Reddit. Only participants between the ages of 18 and 64 were eligible. A total of 41 participants were initially recruited; however, one individual did not complete the survey. This results in data from 40 participants being analyzed. These participants were randomly assigned by a link switcher into two groups (Group A and Group B) of 20 each for equitable survey distribution. The gender distribution within Group A included 11 female participants, eight male participants, and one individual who identified differently, while Group B comprised 11 females and nine males. Most participants for both groups (Group A: 70%, Group B: 66.7%) were aged between 18 and 24 years, followed by those in the 25 to 34 age group.Most participants also indicated a relatively high level of experience with generative AI tools—with Group A (4.65 out of 6) and Group B (4.95 out of 6). However, only 65% of participants in Group A felt somewhat comfortable with generative AI tools, in contrast to 42.9% of participants in Group B feeling **very** comfortable.

5 RESULTS

Overall, those in Group A, who used the prompt assistant, were more satisfied with their results than those who did not (Group B). Group A reported an average result satisfaction of 4.1, while Group B reported an average of 3.4 (as shown in Figure 5).

We asked participants how long it took them, approximately, to achieve a result they were satisfied with. The distribution of times is shown in Figure 6, with Group A averaging 5.725 minutes, and Group B averaging 5.6 minutes. We consider that while the prompt

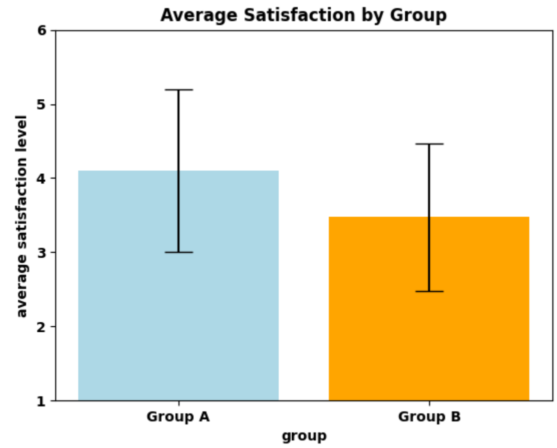

Figure 5: Average satisfaction with resulting images from both participant groups.

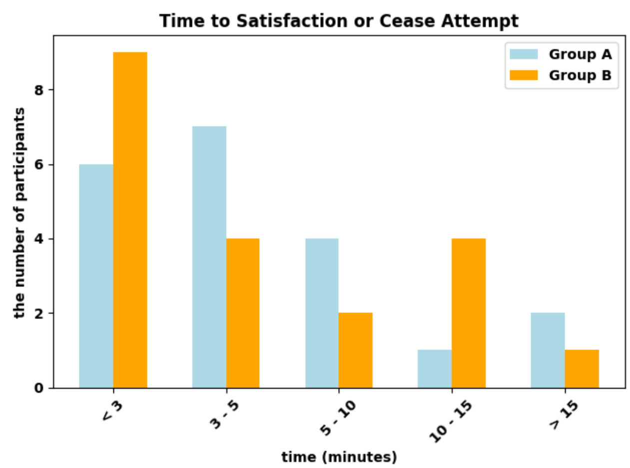

Figure 6: The self-evaluated time it took for participants in Group A and Group B to achieve a result with which they were satisfied in a given task.

assistant may reduce the frustration of creating a text prompt that produces a satisfactory image, it enumerates a large number of dress features and options. This enumeration may cause users to slow down, and be more organized and careful when describing the desired dress. Furthermore, 40% of participants completed all available fields even when it was not necessary, this was likely due to participants' inexperience with the tool along with a lack of provided instruction. Filling out these unneeded fields extends the time required to generate a satisfactory result. This small difference in time could also be attributed to the fact that users in Group B self-reported as being more experienced with generative AI than those in Group A.

When participants were asked to compare the similarity of the selected dress to the generated dress, those in Group A reported an average similarity rating of 3.6, while those in Group B reported an average of 3.7.

Our analysis indicated a linear relationship between satisfaction, similarity, and time. In particular, a higher similarity rating correlated with greater satisfaction, and participants were also more satisfied with results when they spent less time generating the prompt. The relationship was stronger in Group B, possibly due to having lower expectations of the quality of generated images based on input prompts. We noted that 70% of participants in Group B revised their prompts, while only 35% of participants in Group A made revisions.

We found it peculiar that participants in Group A rated the similarity between the reference and generated images lower, on average than those in Group B, so we conducted the further analysis. We found that participants had varied standards and definitions of similarity. Based on comments left by some, they found it hard to generate images that replicated the model, background, and dress of the reference sample exactly. Therefore, they assigned a lower similarity rating, even when the dress closely aligned with the reference sample. Better instructions for assessing similarity may have assisted participants.

In order to more accurately assess participants' generative efforts, we identified 11 common criteria against which to re-evaluate their work. These criteria include pattern, skirt, neckline, sleeve presence, sleeve length, skirt length, overall shape, model ethnicity, embellishments, colour, and background setting. A dress which meets a specific criterion is awarded a score of one; if it partially meets the criterion, it receives a score of 0.5; if it fails to meet the criterion, it is scored zero (as shown in Table 2). By using the scoring system we designate as "accuracy", we found that participants in Group A had an average accuracy of 7.69 out of 11 and those in Group B had an average accuracy of 6.55 out of 11. This higher average suggests that, in general, Group A did a better job than Group B in the image generation task. However, it should be noted that the average self-reported similarity rating from Group A was slightly lower than that from Group B. Considering the accuracy score, participants from Group A may have undervalued the similarity of their generated images. The spectrum of accuracy scores is shown in Figure 8. The range of accuracy scores shows that some individuals from Group A excelled, achieving scores above 9. Such high-performance levels, however, were less seen among participants from Group B. Additionally, when the Cumulative Relative Frequency reaches approximately 0.4 and above, the Empirical C.D.F for Group A consistently leads that of Group B. This indicates that Group A achieved better scores at the same levels of cumulative relative frequency compared to Group B.

By reviewing the average satisfaction, similarity, and accuracy scores, Female participants in general had a better performance compared to male participants, as shown in Table 1. To be specific, females self-reported higher scores in both satisfaction and similarity. Moreover, the average scores among female individuals exceeded that of the collective group, while the self-reported scores from male participants fell below the group's average. The pattern is consistent with the accuracy scores calculated by our evaluation system. That female participants scored, on average, better than male participants in all categories (satisfaction, similarity, and accuracy), may reflect the nature of the task as female participants may be more familiar with terminologies associated with dress designs [19].

Table 1: Average Scores for Satisfaction, Similarity, and Accuracy across gender and group divisions.

	Group A		**Group B**	
Gender	**Female**	**Male**	**Female**	**Male**
Satisfaction	4.27	3.8	3.7	3.222
Similarity	3.9	3.2	3.727	3.666
Accuracy	8.15	7.125	6.818	5.333

We took a closer look at the prompts given by participants in both groups. Prompts from participants in Group A were very similar due to the assistant following strict grammar, and while the prompt assistant provided the ability, via text boxes, to give custom design elements we found that participants did not take this opportunity. Prompts from participants in Group B, by comparison, were much shorter and lacked any kind of consistent structure or format. Generally speaking, Group B's prompts lacked both descriptive terminology for the dress but also prompt modifiers. Additionally, some of the descriptive language that was used in these prompts was very emotional or ambiguous. This is reflected in Figure 7. The prompts for both dresses produced by Group A participants are very similar, differing only in the specific dress details. One example prompt is given below:

> **Prompt**: A above calf-length blue polka dot dress; circle skirt; sweetheart neckline; short flutter sleeves; standing up; alone; full body shot; full body portrait; long shot; isolated; isolated on white background; plain background; well-lit;
> **Negative prompt**: Blurriness; pixelation; distortion; visual artifacts; compression artifacts; noise; graininess; discoloration; overly loose or ill-fitting; lighting issues; overexposure; underexposure; harsh shadows; distractions; clutter; crowds;

While the prompts used by the Group B participants were vastly different. For the blue dress:

> blue polka dot dress for sale

And for the purple dress:

> a purple but uncomfortable dress

Given the lack of specificity in these Group B prompts, it is easy to see why the results differed from the reference (Figure 7).

5.1 Task Selection

The reference images provided to participants represented fairly simple dresses with common silhouettes, materials, and either simple or regular patterns (if any), as shown in Figure 4. However, despite the relative simplicity of each dress, we noted that nearly 50% of participants selected option I to recreate. This dress is the simplest, with an easy-to-describe silhouette, solid colour, and lacking patterns or complex materials. This dress may have been more commonly selected due to it being the first in the list, or, because people tend to prefer simple tasks when presented with options of varying difficulty [3]. Within Group A, those who chose Option I self-reported higher satisfaction (4.5 vs. 3.83) and similarities (3.75 vs. 3.5) scores than those who chose different options. Furthermore, the average scores of individuals who chose Option I were better

Group A Sample Group B

(a)

Group A Sample Group B

(b)

Figure 7: Image participants from Group A (left) and Group B (right), with the reference image positioned centrally.

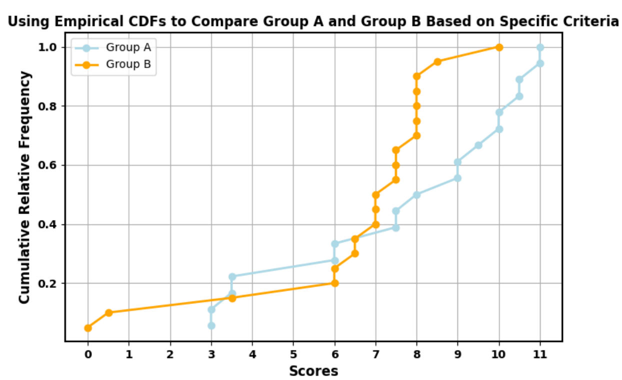

Figure 8: The distribution of user-generated images evaluated based on 11 specific criteria.

than the overall group performance. In Group B, there was a slight difference. The average satisfaction score for participants selecting Option I was marginally lower (3.429) compared to those who opted for other options (3.5). Several examples of output from Group A and B participants can be found in Figure 9, and the corresponding scores in Table 2.

5.2 Overall Experience

Participants in Group A found the prompt assistant beneficial while rating its utility, reflecting this in their feedback with an average score of 4.2 out of 5 on a scale from 1 to 5. Participants in Group B, who had to manually create prompts, noted that they found the task to be of average difficulty (2.5 out of 5 on a scale from 1 to 5)

An analysis of participant feedback indicated that 95% of participants from Group A found our prompt assistance to be useful, making the task easy and enjoyable. While 40% of participants in Group B emphasized the importance of being specific and detailed in their prompts. They expressed a desire for a tool that could supply professional vocabularies, which is what the task-specific prompt assistant provided to Group A. An analysis of output images indicates that individuals who applied specific words in their prompts, such as sparkly, chic silver, halter, etc., reported higher satisfaction rates in Group B.

Furthermore, we noted that most participants in Group B did not provide negative prompts or misused them. Only 7 out of 20 participants provided negative prompts, with two providing prompts that were counterproductive. Negative prompts help achieve better results by specifying what you do not want ("ugly", "bad lighting") [13].

5.3 Creative Outputs

Participants were also asked to create a dress of their own imagining using the respective tool. Figure 10 illustrates some of the creations produced by participants. We note that Group A exhibited less creativity in the "free choice" generation compared to Group B. Although the prompt assistant provided users with a textbox for manual/custom input, participants generally did not use them—choosing only from the options listed instead. Greater visibility of the option to add one's own text, from either a tutorial, examples, or helpful mouse-overs, could encourage users to use this ability and improve their creativity. Additionally, these images matched the prompts much more closely than those produced by Group B. Finally, while Group B's works exhibited more creativity, the lack of prompt modifiers resulted in poorly framed dresses presented in busy settings.

6 CONCLUSION

In this paper, we examined the impact of a task-specific prompt assistant for generating images. We asked participants to recreate dresses using text-to-image generative AI using either a web form prompt assistant, or, no assistant at all.

Our analysis of the responses suggests that when using generative AI for a specific task, such as designing a dress, a web form-based prompt assistant is helpful. The prompt assistant helped users organize their ideas by providing a detailed breakdown of the task, complete with reference images. This resulted in participants being more satisfied with their output, and, the output matching the prompt more accurately.

While on average, participants using the prompt assistant took slightly more time to produce a satisfactory result than those that did not use the assistant, we believe this was due to inexperience using the tool. Due to the number of design features and options presented by prompt assistants, first-time users might have been overwhelmed. We believe that if the task had been repeated, allowing users to gain familiarity with the tools, participants in Group A, who used the prompt assistant, would be able to produce a satisfactory result much faster than those in Group B. This was a limitation of the study and something we wish to investigate further moving forward.

Table 2: A demonstration of our accuracy assessment system for dress Option I. Note that participants from Group A are shown in blue, and those of Group B are shown in orange.

Participant	Pattern	Skirt	Neckline	Sleeve	Sleeve Length	Skirt Length	Shape	Ethnicity	Embellishment	Colour	Background	Score
GA1	1	0.5	0	1	1	1	1	1	1	1	0.5	9
GA2	0	0	0	1	1	0.5	1	1	1	1	1	7.5
GB1	0	0	0	0	0	0	0	0	0	0.5	0	0.5
GB2	0	1	0	1	1	0	1	0	1	1	0	6

(a) reference (b) GA1 (c) GA2 (d) GB1 (e) GB2

Figure 9: Reference and select participant results from Group A and B.

(a) Group A: (b) Group A (c) Group B (d) Group B

Figure 10: A selection of the participants' imaginative creations from both groups.

ACKNOWLEDGMENTS

We sincerely thank all the participants for spending time on the survey and for sharing their insightful thoughts. We also greatly appreciate the anonymous reviewers for their valuable feedback, which significantly helped improve our paper.

REFERENCES

[1] Nikos Arechiga, Frank Permenter, Binyang Song, and Chenyang Yuan. 2023. Drag-guided diffusion models for vehicle image generation. arXiv:2306.09935 [cs.LG]

[2] Ali Borji. 2023. Generated Faces in the Wild: Quantitative Comparison of Stable Diffusion, Midjourney and DALL-E 2. arXiv:2210.00586 [cs.CV]

[3] Donald J. Campbell. 1988. Task Complexity: A Review and Analysis. *The Academy of Management Review* 13, 1 (1988), 40–52. http://www.jstor.org/stable/258353

[4] Tingfeng Cao, Chengyu Wang, Bingyan Liu, Ziheng Wu, Jinhui Zhu, and Jun Huang. 2023. BeautifulPrompt: Towards Automatic Prompt Engineering for Text-to-Image Synthesis. arXiv:2311.06752 [cs.CL]

[5] Giannis Daras and Alexandros G. Dimakis. 2022. Discovering the Hidden Vocabulary of DALLE-2. arXiv. https://arxiv.org/abs/2206.00169

[6] Chengbin Du, Yanxi Li, Zhongwei Qiu, and Chang Xu. 2023. Stable Diffusion is Unstable. In *Advances in Neural Information Processing Systems*, A. Oh, T. Neumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine (Eds.), Vol. 36. Curran Associates, Inc., 58648–58669. https://proceedings.neurips.cc/paper_files/paper/2023/file/b733cdd80ed2ae7e3156d8c33108c5d5-Paper-Conference.pdf

[7] Wenxiang Jiao, Wenxuan Wang, Jen tse Huang, Xing Wang, Shuming Shi, and Zhaopeng Tu. 2023. Is ChatGPT A Good Translator? Yes With GPT-4 As The Engine. arXiv:2301.08745 [cs.CL]

[8] Enis Karaarslan and Ömer Aydın. 2024. Generate Impressive Videos with Text Instructions: A Review of OpenAI Sora, Stable Diffusion, Lumiere and Comparable Models. *TechRxiv* (Feb. 2024). https://doi.org/10.36227/techrxiv.170862194.43871446/v1

[9] LearnPrompting. [n. d.]. Learn prompting: Your guide to communicating with AI. https://learnprompting.org/docs/image_prompting/intro

[10] Bolong Liu, Hao Zhang, Jie Liu, and Qiang Wang. 2023. ACIGS: An automated large-scale crops image generation system based on large visual language multi-modal models. In *2023 20th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*. 7–13. https://doi.org/10.1109/SECON58729.2023.10287530

[11] Vivian Liu and Lydia B Chilton. 2022. Design Guidelines for Prompt Engineering Text-to-Image Generative Models. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>New Orleans</city>, <state>LA</state>, <country>USA</country>, </conf-loc>) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 384, 23 pages. https://doi.org/10.1145/3491102.3501825

[12] Yiyang Mei. 2024. Prompting the E-Brushes: Users as Authors in Generative AI. *SSRN Electronic Journal* (March 2024). https://doi.org/10.2139/ssrn.4751625

[13] Daiki Miyake, Akihiro Iohara, Yu Saito, and Toshiyuki Tanaka. 2023. Negative-prompt Inversion: Fast Image Inversion for Editing with Text-guided Diffusion Models. arXiv:2305.16807 [cs.CV]

[14] Fiona Fui-Hoon Nah, Ruilin Zheng, Jingyuan Cai, Keng Siau, and Langtao Chen. 2023. Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research* 25, 3 (July 2023), 277–304. https://doi.org/10.1080/15228053.2023.2233814
[15] OpenAI. [n. d.]. Prompt Engineering. https://help.openai.com/en/articles/6654000-best-practices-for-prompt-engineering-with-openai-api
[16] Jonas Oppenlaender. 2023. A taxonomy of prompt modifiers for text-to-image generation. *Behaviour amp; Information Technology* (Nov. 2023), 1–14. https://doi.org/10.1080/0144929x.2023.2286532
[17] Nikita Pavlichenko and Dmitry Ustalov. 2023. Best Prompts for Text-to-Image Models and How to Find Them. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '23)*. ACM. https://doi.org/10.1145/3539618.3592000
[18] Generated Photos. [n. d.]. AI Human Generator. https://generated.photos/human-generator
[19] Amelia B. Williams. 2022. *Sustainable Fashion Leadership: The Female Phenomenon*. Ph. D. Dissertation. https://proxy.lib.uwaterloo.ca/login?url=https://www.proquest.com/dissertations-theses/sustainable-fashion-leadership-female-phenomenon/docview/2715691622/se-2 Copyright - Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works; Last updated - 2024-03-19.
[20] Tianyu Wu, Shizhu He, Jingping Liu, Siqi Sun, Kang Liu, Qing-Long Han, and Yang Tang. 2023. A Brief Overview of ChatGPT: The History, Status Quo and Potential Future Development. *IEEE/CAA Journal of Automatica Sinica* 10, 5 (May 2023), 1122–1136. https://doi.org/10.1109/JAS.2023.123618
[21] Pinar Yanardag and Emily Salvador. 2018. The Little Black Dress. https://lbd-ai.com/

Received 6 April 2024; revised 19 May 2024