Designing Expressive Interaction with Generative Artificial Intelligence

Xiaohan PENG [a,1]

[a] *LISN, Université Paris-Saclay, CNRS, Inria, France*
ORCiD ID: Xiaohan Peng https://orcid.org/0000-0002-5202-419X

Abstract. Current generative AI interfaces have posed challenges for creative professionals to interact with AI effectively. To address this issue from an interaction perspective, the primary focus of my research is designing expressive interaction with AI using mainly human-centered design approaches. More specifically, my research will delve into understanding how creative professionals use generative AI in professional settings, designing interactions with AI expressively and visually, and building engaging and controllable creative AI systems that keep a balanced agency between users and systems.

Keywords. Human-AI Interaction, Human-Centered Design, Creativity Support Tools, Expressivity

1. Context

Advancements in Generative Artificial Intelligence (GenAI) and Large Language Models (LLMs) have facilitated the production of high-quality synthetic images using descriptive text prompts [17,39], which has lowered the barriers to AI engagement for a wider range of audience including designers and artists. Current creative GenAI applications such as DALL·E 3 [35], stable diffusion [33], and midjourney [32] allow designers and artists to generate novel images, modify and iterate visual designs through text input. Research has shown that LLMs-based image generation can be especially useful in fast iteration and combining ideas in creative tasks [23]. Moreover, GenAI exhibits versatility across various creative fields by generating a range of outputs beyond images, such as music, videos, 3D models, and dance movements [47,37,1,9,29].

While current GenAI applications offer low-barrier text input access and great creative possibilities, non-AI professionals still often find it difficult to accurately express their intentions to AI and achieve their desired results [50,24]. More specifically, Subramonyam et al. have described the challenges with interacting with LLMs as intentionality gap (user only stating goals without other cognitive task processes), capability gap (user being unaware of AI's capabilities) and language gap (user unable to express themselves clearly and effectively to AI) [44].

---

[1] Corresponding Author: Xiaohan Peng, xiaohan.peng@universite-paris-saclay.fr

One of the most common technical solutions to these challenges is "prompt engineering", which is a means of structuring text in order to efficiently interact with AI models and tune the generation process towards desired outcomes [48,41], for instance in text-to-image generation, using the structure like "A MEDIUM of SUBJECT in the STYLE style" and specifying explicitly color, technique, relationships in space, and motifs [26]. There are a large number of existing studies on prompt engineering that have been focusing on effective strategies for prompt writing [27,25,36,10,26], interactive prompt feedback [40,12] as well as more customized suggested prompt refinement [6]. Prompt engineering has also been studied in design practice. A study conducted by Chiou et al. on co-ideation between participants and AI image generators has demonstrated that strategies such as preparing data, breaking design into steps, crafting specific styles and reviewing can help generate effective prompts in design[10].

Apart from formulating effective prompts, some research uses conversational and turn-taking approaches in the interactions with GenAI, in order to share semantics between humans and AI and develop repairing mechanisms [44]. For instance, ChatPainter, a GAN-based model, uses a Q&A interaction such as "Q: is the woman standing on the board? A: no she is beside it." to facilitate create image details that meet users' expectations [43]. ControlNet allows users to use additional text descriptions to have incremental spatial control on existing generated images with diffusion models [51].

While these studies have provided insights into utilizing different AI input strategies to generate more satisfying results, the interaction remains text-heavy and always ends up in "trials and errors". For visually oriented creative professionals, optimizing inputs in a rule-based way seems counter-intuitive, especially in divergent thinking processes [22]. Apart from prompt engineering and other techniques for detailing textual input, there seems to be a lack of studies on expressive and efficient approaches to empower designers and artists in interacting with GenAI.

In addition to inefficient and non-intuitive input interaction, I also see user engagement and control as another challenge in creative AI applications from a human-centered perspective. Despite GenAI has already been applied in a wide range of creative domains and practices such as quick exploration of visual ideas [46], UI design [34], fashion style clustering, forecasting and merging [18], speeding up design process [45], visual communication and fast prototyping [20,11], a large number of creative AI applications seem to be built in a fairly result-oriented way, where the results are presented but the creative process is missing. These tech-driven applications reflect an inclination within AI development to prioritize substitution over augmentation [46], and the absence of interaction makes them closer to "direct execution of goal" [44]. This potentially raises broader societal concerns such as deskilling [49]. On the other hand, the lack of engagement and control of the creative process also prompts users to view AI as an obedient tool. However, the relationship between human and AI can be rather versatile, and there can be different levels of agency in such interactions [21]. For instance, instead of functioning solely as a tool, AI can serve as a cooperative companion and assistant, guiding discussions and task outcomes [42]. Questions on how to maintain a balanced agency between humans and AI during interaction in context, and how to support such mixed-initiative and complex relationships remain to be answered.

2. Research Questions

In my research, the key research questions I am interested in are:

1. What are the interactions in users' current creative practice, and how this can inform us to build a more controllable and meaningful system for professional design?
2. How to use expressive but controllable modalities and interaction techniques in designing interaction to support creative professionals to engage with GenAI?
3. How to integrate GenAI system interpretation and results meaningfully in design process to help users produce results more aligned with their expectations?

3. Methodology

In this section I will briefly introduce the theories and methodology I am likely to apply in my future research. I aim to use a mixed methodology in my research, where my primary methodology will be rooted in human-centered design, while also remaining receptive to theories and methodologies originating from other communities.

3.1. Generative Theories of Interaction as Methodological Building Block

In current human-AI collaboration research, a large number of methodologies applied seem to be tech-centered or highly reliant on existing technology. To ground my research from a human-centered perspective and envision interaction without being constrained by existing technology and existing usages of technology, one of the theoretical constructs that I will apply in my research is generative theories of interaction. A generative theory of interaction is grounded in empirically-based human activity and behavior theories, built based on concepts and generative principles related to the theories, and enables analytical, critical and constructive lenses to understand existing artifacts and inform future design [3]. In the context of HCI I see it can fit in my research both theoretically (as a theoretical foundation) and methodologically (using other theories as part of building blocks to inform design) since it constructs a generative bridge between HCI concepts and creating new designs. Two human-centered design theories that are closely relevant to my research and well-suited for the construct of generative theories are instrumental interaction [2] and human-computer partnership [30]. Since my research focuses on instant and immediate interaction between creative professionals and AI also new software adapting and appropriating patterns, these two theories will be great fits since they are suited for interactions in short-term and mid-term [3]. However, it might also be interesting to design for collaboration (although not being my main focus at the moment) and observe long-term community practices of using expressive AI interaction in the creative community, which might make community and common objects a relevant theory in a long-term project.

Considering using generative theory of interaction as a methodological building block, I would use "expressivity" as the main concept in my future research. Since in recent literature researchers have investigated the term "expressivity" in HCI context and found its meanings in sensories, dynamic forms, social activity and values [8], the level of concept abstraction should be appropriate, although the concept's detailed definition

might still require more in-depth elaboration. In terms of principles, initial thoughts are "input richness", "expressive awareness" and "output variability". "Input richness" and "output variability" are inspired by the principle of expressivity for human-computer partnership [30]. However, here I use two lower-level interpretations of the principle and they refer to the different levels of abstraction or resolution supported by input, and the capability of producing various and customized output based on different inputs. "Expressive awareness" is inspired by previous works on expressive communication devices [15,16]. From my perspective "expressive awareness" refers to the implicit sense of presence and connection during the interaction. The analytical, critical and constructive lenses that can be applied to the concept and principles remain to be researched from other existing literature.

3.2. Mixed Perspectives in Design

HCI methodologies that I apply in research will be primarily from a human-centered perspective. Depending on the stage of the research, different methodologies, mostly empirical will be applied such as semi-structured interviews, walkthroughs [28], thematic analysis and structured observation [14,5,19,22]. In the process of design, I will mainly design from third-person and second-person perspectives, meaning designing for users based on user studies and co-designing with users in context [13]. The decision on the perspectives is largely motivated by the discipline and the initiative to build more human-centered technology. Thus, design decisions are primarily informed by user perspectives or co-evolved in an engaging process. However, I am also open to using first-person methods such as research-through-design, which will be further deliberated depending on the subdomain of future research.

3.3. Learning and Getting Inspired from Users in Creative Practice

As an increasing number of GenAI applications built with different approaches are positioning at creativity support, one concern that comes into place is who the potential users will be for these new systems and how they will interact with them. From a human-centered perspective, I prefer to gather insights about users and draw design implications at the initial stage of design. In one completed project, a preliminary study has been conducted to understand how visually oriented users can interact with GenAI in their moodboard process, especially in terms of prompt formulation. The focuses are on the interaction and workflow with existing moodboard tools, how they prompt textually with GenAI, and their responses and actions towards the AI-generated output. The study included two think-aloud sessions, one of which was creating themed images with GenAI. An appropriated version of "Wizard of Oz" method [7] has been applied, which is instead of directly interacting with the GenAI, the experimenter played a mediator between the participant and the AI encouraging them to express their intents with technical impact minimized.

Apart from understanding users from a third-person perspective, more engaging activities such as participatory design workshops can be conducted to involve users in the design process [31]. In the scope of my research, participatory design is particularly suitable in the initial stage of exploring how artists and designers create professional work and how they are currently using GenAI in their practice. Activities such as co-

brainstorming, co-design with AI and finding creative usages of AI can be explored with various themes, such as the creative needs for precision or abstraction, their expectations of the roles of AI, their preferred ways of expressing intentions to AI.

3.4. Interaction in Human-AI Collaboration

To design interaction that is expressive but also controllable in professional creative settings, in my research design decisions are mainly drawn from relevant HCI theories, prototyping activities as well as empirical studies. Theories such as instrumental interaction [2], technical reasoning [38] and activity theory [4] used in HCI context work as an aid in creating a connection from thinking on objects of interests and actions to designing interaction. Prototyping as a process, particularly paper prototyping, prompts reasoning interactively, also enables a representation in physical form and helps form a rough model for interaction. Studies can also inform interaction design. In one of my completed projects, the results show that some designers consider "ease of steering" important thus they prefer using a system with simple interfaces and low cognitive workload; some designers think "system assumptions" affect controllability, thus they sometimes prefer manual control rather than let AI or the computer "make interpretations" for them; some designers think being able to control the input to AI precisely is the key to control, while some others interpret control as controlling over the whole AI generation process. These findings suggest that designers' definitions of "control" over the AI affect their choice of interaction and inform design possibilities from each or combined implications.

4. Preliminary results

In the first project completed during my PhD research, I developed an interactive ideation system that allows designers to explore and express themselves when working together with GenAI tools using multimodal interaction. The system is built based on design implications obtained from the preliminary study, and it is digital web-based moodboard tool that incorporates GenAI functionalities with multimodal prompts including images, colors and semantics rather than pure text prompts. In comparative structured observation study with 12 professional designers, I compared the multimodal prompt moodboard tool with a baseline system that uses text prompts. The results have highlighted that prompt decomposition (from text to multimodal) allows users to explore underlying image connections and express intentions to GenAI. The work suggests design implications for designing expressive and controllable human-GenAI interaction in a design context and provides insights into the potential of using multimodal input to help designers better express themselves.

5. Discussion and Future work

Findings from my finished research project indicate that utilizing multimodal input rather than relying solely on textual input enhances designers' ability to express ideas and explore diverse creative AI system usages. Given that my completed work incorporated images, colors, and semantics as AI input, future projects can delve into varied aspects which can include investigating the suitability of different modalities for creative AI in-

teraction in different use cases, assessing the stability and reliability of these modalities in AI interaction, exploring the expressive potential inherent in these modalities, and determining meaningful ways to combine them in future AI-incorporated designs. Apart from modalities, I am also interested in the potential of expressive and artistic mediums, e.g. drawing and sketching. In my future research, I would like to explore using drawing and sketching as an interaction technique or a part of design activity in collaborating with AI.

Implicated by different "control" interpretations, my future work will also dive deeper into specific controllability needs and navigate the balance between these different facets of control. In order to achieve this goal, part of my future work would focus on rethinking the interaction paradigm in human-AI from a user action point of view. I aim to use more expressive interaction techniques such as sketching to build a controllable system. Additionally, I am also interested in designing more fluid relationships and control between human and AI with different levels of agency, rather than using AI solely as a tech-driven obedient tool. In the long term, I envision my research potentially extending beyond GenAI interaction with creative professionals to include a broader demographic in their daily routines.

6. Acknowledgements

References

[1] Andrea Agostinelli, Timo I. Denk, Zalán Borsos, Jesse Engel, Mauro Verzetti, Antoine Caillon, Qingqing Huang, Aren Jansen, Adam Roberts, Marco Tagliasacchi, Matt Sharifi, Neil Zeghidour, and Christian Frank. Musiclm: Generating music from text, 2023.

[2] Michel Beaudouin-Lafon. Instrumental interaction: an interaction model for designing post-wimp user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '00, page 446–453, New York, NY, USA, 2000. Association for Computing Machinery.

[3] Michel Beaudouin-Lafon, Susanne Bødker, and Wendy E. Mackay. Generative Theories of Interaction. *ACM Transactions on Computer-Human Interaction*, 28(6):45:1–45:54, November 2021.

[4] Olav Bertelsen and Susanne Bødker. *Activity Theory*, pages 291–324. 04 2003.

[5] Adrien Bousseau, Theophanis Tsandilas, Lora Oehlberg, and Wendy E. Mackay. How novices sketch and prototype hand-fabricated objects. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, page 397–408, New York, NY, USA, 2016. Association for Computing Machinery.

[6] Stephen Brade, Bryan Wang, Mauricio Sousa, Sageev Oore, and Tovi Grossman. Promptify: Text-to-Image Generation through Interactive Prompt Exploration with Large Language Models. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, UIST '23, pages 1–14, New York, NY, USA, October 2023. Association for Computing Machinery.

[7] Jacob T. Browne. Wizard of oz prototyping for machine learning experiences. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI EA '19, page 1–6, New York, NY, USA, 2019. Association for Computing Machinery.

[8] Miguel Bruns, Stijn Ossevoort, and Marianne Graves Petersen. Expressivity in interaction: a framework for design. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA, 2021. Association for Computing Machinery.

[9] Yang Chen, Jingwen Chen, Yingwei Pan, Xinmei Tian, and Tao Mei. 3d creation at your fingertips: From text or image to 3d assets. In *Proceedings of the 31st ACM International Conference on Multimedia*, MM '23, page 9408–9410, New York, NY, USA, 2023. Association for Computing Machinery.

[10] Li-Yuan Chiou, Peng-Kai Hung, Rung-Huei Liang, and Chun-Teng Wang. Designing with AI: An Exploration of Co-Ideation with Image Generators. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*, DIS '23, pages 1941–1954, New York, NY, USA, July 2023. Association for Computing Machinery.

[11] John Joon Young Chung. Artistic user expressions in ai-powered creativity support tools. In *Adjunct Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*, UIST '22 Adjunct, New York, NY, USA, 2022. Association for Computing Machinery.

[12] Niklas Deckers, Maik Fröbe, Johannes Kiesel, Gianluca Pandolfo, Christopher Schröder, Benno Stein, and Martin Potthast. The infinite index: Information retrieval on generative text-to-image models. In *Proceedings of the 2023 Conference on Human Information Interaction and Retrieval*, CHIIR '23, page 172–186, New York, NY, USA, 2023. Association for Computing Machinery.

[13] Audrey Desjardins, Oscar Tomico, Andrés Lucero, Marta E. Cecchinato, and Carman Neustaedter. Introduction to the special issue on first-person methods in hci. *ACM Trans. Comput.-Hum. Interact.*, 28(6), dec 2021.

[14] Jérémie Garcia, Theophanis Tsandilas, Carlos Agon, and Wendy E. Mackay. Structured observation with polyphony: a multifaceted tool for studying music composition. In *Proceedings of the 2014 Conference on Designing Interactive Systems*, DIS '14, page 199–208, New York, NY, USA, 2014. Association for Computing Machinery.

[15] William Gaver, Andy Boucher, Dean Brown, David Chatting, Naho Matsuda, Liliana Ovalle, Andy Sheen, and Michail Vanis. Yo–yo machines: Self-build devices that support social connections during the pandemic. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, New York, NY, USA, 2022. Association for Computing Machinery.

[16] William Gaver and Frances Gaver. Living with Light Touch: An Autoethnography of a Simple Communication Device in Long-Term Use. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, pages 1–14, New York, NY, USA, April 2023. Association for Computing Machinery.

[17] Jonathan Ho, Chitwan Saharia, William Chan, David J. Fleet, Mohammad Norouzi, and Tim Salimans. Cascaded diffusion models for high fidelity image generation. *Journal of Machine Learning Research*, 23(47):1–33, 2022.

[18] Youngseung Jeon, Seungwan Jin, Patrick C. Shih, and Kyungsik Han. FashionQ: An AI-Driven Creativity Support Tool for Facilitating Ideation in Fashion Design. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, pages 1–18, New York, NY, USA, May 2021. Association for Computing Machinery.

[19] Taslim Arefin Khan, Dongwook Yoon, and Joanna McGrenere. *Designing an Eyes-Reduced Document Skimming App for Situational Impairments*, page 1–14. Association for Computing Machinery, New York, NY, USA, 2020.

[20] Hyung-Kwon Ko, Gwanmo Park, Hyeon Jeon, Jaemin Jo, Juho Kim, and Jinwook Seo. Large-scale Text-to-Image Generation Models for Visual Artists' Creative Works. In *Proceedings of the 28th International Conference on Intelligent User Interfaces*, IUI '23, pages 919–933, New York, NY, USA, March 2023. Association for Computing Machinery.

[21] Janin Koch, Nicolas Taffin, Michel Beaudouin-Lafon, Markku Laine, Andrés Lucero, and Wendy E. Mackay. ImageSense: An Intelligent Collaborative Ideation Tool to Support Diverse Human-Computer Partnerships. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW1):45:1–45:27, May 2020.

[22] Janin Koch, Nicolas Taffin, Andrés Lucero, and Wendy E. Mackay. SemanticCollage: Enriching Digital Mood Board Design with Semantic Labels. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*, DIS '20, pages 407–418, New York, NY, USA, July 2020. Association for Computing Machinery.

[23] Chinmay Kulkarni, Stefania Druga, Minsuk Chang, Alex Fiannaca, Carrie Cai, and Michael Terry. A word is worth a thousand pictures: Prompts as ai design material, 2023.

[24] Q. Vera Liao, Hariharan Subramonyam, Jennifer Wang, and Jennifer Wortman Vaughan. Designerly Understanding: Information Needs for Model Transparency to Support Design Ideation for AI-Powered User Experience, February 2023. arXiv:2302.10395 [cs].

[25] Jiachang Liu, Dinghan Shen, Yizhe Zhang, Bill Dolan, Lawrence Carin, and Weizhu Chen. What makes good in-context examples for gpt-3?, 2021.

[26] Vivian Liu and Lydia B Chilton. Design Guidelines for Prompt Engineering Text-to-Image Generative Models. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, pages 1–23, New York, NY, USA, April 2022. Association for Computing Machinery.

[27] Vivian Liu and Lydia B Chilton. Design Guidelines for Prompt Engineering Text-to-Image Generative Models. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, pages 1–23, New York, NY, USA, April 2022. Association for Computing Machinery.

[28] Danielle Lottridge and Wendy E. Mackay. Generative walkthroughs: to support creative redesign. In *Proceedings of the Seventh ACM Conference on Creativity and Cognition*, C&C '09, page 175–184, New York, NY, USA, 2009. Association for Computing Machinery.

[29] Ryan Louie, Andy Coenen, Cheng Zhi Huang, Michael Terry, and Carrie J. Cai. Novice-ai music co-creation via ai-steering tools for deep generative models. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, page 1–13, New York, NY, USA, 2020. Association for Computing Machinery.

[30] Wendy E. Mackay. Creating human-computer partnerships. In Hugo Plácido da Silva and Pietro Cipresso, editors, *Computer-Human Interaction Research and Applications*, pages 3–17, Cham, 2023. Springer Nature Switzerland.

[31] Wendy E. Mackay and Michel Beaudouin-Lafon. Participatory Design and Prototyping. In *Handbook of Human Computer Interaction*, pages 1–33. Springer International Publishing, February 2023.

[32] Midjourney. Midjourney. https://www.midjourney.com/, 2024. [Accessed 26-01-2024].

[33] Emad Mostaque. Stable Diffusion Public Release — Stability AI — stability.ai. https://stability.ai/news/stable-diffusion-public-release, 2023. [Accessed 26-01-2024].

[34] Mohammad Amin Mozaffari, Xinyuan Zhang, Jinghui Cheng, and Jin L.C. Guo. GANSpiration: Balancing Targeted and Serendipitous Inspiration in User Interface Design with Style-Based Generative Adversarial Network. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, pages 1–15, New York, NY, USA, April 2022. Association for Computing Machinery.

[35] Openai. DALL·E 3 — openai.com. https://openai.com/dall-e-3, 2024. [Accessed 26-01-2024].

[36] Jonas Oppenlaender. Prompt engineering for text-based generative art. *arXiv preprint arXiv:2204.13988*, 2022.

[37] Ben Poole, Ajay Jain, Jonathan T. Barron, and Ben Mildenhall. Dreamfusion: Text-to-3d using 2d diffusion, 2022.

[38] Miguel A. Renom, Baptiste Caramiaux, and Michel Beaudouin-Lafon. Exploring technical reasoning in digital tool use. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, New York, NY, USA, 2022. Association for Computing Machinery.

[39] Chitwan Saharia, William Chan, Saurabh Saxena, Lala Li, Jay Whang, Emily Denton, Seyed Kamyar Seyed Ghasemipour, Burcu Karagol Ayan, S. Sara Mahdavi, Rapha Gontijo Lopes, Tim Salimans, Jonathan Ho, David J Fleet, and Mohammad Norouzi. Photorealistic text-to-image diffusion models with deep language understanding, 2022.

[40] Téo Sanchez. Examining the Text-to-Image Community of Practice: Why and How do People Prompt Generative AIs? In *Proceedings of the 15th Conference on Creativity and Cognition*, C&C '23, pages 43–61, New York, NY, USA, June 2023. Association for Computing Machinery.

[41] Victor Sanh, Albert Webson, Colin Raffel, Stephen H. Bach, Lintang Sutawika, Zaid Alyafeai, Antoine Chaffin, Arnaud Stiegler, Teven Le Scao, Arun Raja, Manan Dey, M. Saiful Bari, Canwen Xu, Urmish Thakker, Shanya Sharma, Eliza Szczechla, Taewoon Kim, Gunjan Chhablani, Nihal V. Nayak, Debajyoti Datta, Jonathan Chang, Mike Tian-Jian Jiang, Han Wang, Matteo Manica, Sheng Shen, Zheng Xin Yong, Harshit Pandey, Rachel Bawden, Thomas Wang, Trishala Neeraj, Jos Rozen, Abheesht Sharma, Andrea Santilli, Thibault Févry, Jason Alan Fries, Ryan Teehan, Stella Biderman, Leo Gao, Tali Bers, Thomas Wolf, and Alexander M. Rush. Multitask prompted training enables zero-shot task generalization. *CoRR*, abs/2110.08207, 2021.

[42] Albrecht Schmidt, Passant Elagroudy, Fiona Draxler, Frauke Kreuter, and Robin Welsch. Simulating the Human in HCD with ChatGPT: Redesigning Interaction Design with AI. *Interactions*, 31(1):24–31, January 2024.

[43] Shikhar Sharma, Dendi Suhubdy, Vincent Michalski, Samira Ebrahimi Kahou, and Yoshua Bengio. Chatpainter: Improving text to image generation using dialogue, 2018.

[44] Hariharan Subramonyam, Christopher Lawrence Pondoc, Colleen Seifert, Maneesh Agrawala, and Roy

Pea. Bridging the gulf of envisioning: Cognitive design challenges in llm interfaces, 2023.

[45] Jakob Tholander and Martin Jonsson. Design Ideation with AI - Sketching, Thinking and Talking with Generative Machine Learning Models. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*, DIS '23, pages 1930–1940, New York, NY, USA, July 2023. Association for Computing Machinery.

[46] Mathias Peter Verheijden and Mathias Funk. Collaborative diffusion: Boosting designerly co-creation with generative ai. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI EA '23, New York, NY, USA, 2023. Association for Computing Machinery.

[47] Benedikte Wallace, Clarice Hilton, Kristian Nymoen, Jim Torresen, Charles Patrick Martin, and Rebecca Fiebrink. Embodying an Interactive AI for Dance Through Movement Ideation. In *Proceedings of the 15th Conference on Creativity and Cognition*, C&C '23, pages 454–464, New York, NY, USA, June 2023. Association for Computing Machinery.

[48] Jules White, Quchen Fu, Sam Hays, Michael Sandborn, Carlos Olea, Henry Gilbert, Ashraf Elnashar, Jesse Spencer-Smith, and Douglas C. Schmidt. A prompt pattern catalog to enhance prompt engineering with chatgpt, 2023.

[49] Allison Woodruff, Renee Shelby, Patrick Gage Kelley, Steven Rousso-Schindler, Jamila Smith-Loud, and Lauren Wilcox. How Knowledge Workers Think Generative AI Will (Not) Transform Their Industries, March 2024.

[50] J.D. Zamfirescu-Pereira, Richmond Y. Wong, Bjoern Hartmann, and Qian Yang. Why Johnny Can't Prompt: How Non-AI Experts Try (and Fail) to Design LLM Prompts. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, pages 1–21, New York, NY, USA, April 2023. Association for Computing Machinery.

[51] Lvmin Zhang, Anyi Rao, and Maneesh Agrawala. Adding conditional control to text-to-image diffusion models, 2023.