Exploring Generative Image AI for Design Evaluation in Human-Robot Interaction Research

YONG-GYUN GHIM, JOE CURTSINGER
UNIVERSITY OF CINCINNATI

This study explores generative image AI's capacity to rapidly create multiple realistic renderings for human-robot interaction (HRI) research. By utilizing freehand sketches and employing an iterative process of rendering generation and refinements in Vizcom and Photoshop, variations of service robot designs for four different use contexts were created and visualized as user evaluation media to understand the perception of robot morphology. The AI-assisted workflow, from sketching to rendering, is examined by investigating each step with outcomes, followed by reflections on the strengths and challenges of co-creating with generative image AI. This study contributes to design research and practice by proposing a novel application of generative image AI for creating variations of situated robot representations in HRI and by offering insights and suggestions for its effective use.

Keywords: Human-robot interaction, Generative image AI, Design evaluation, Product visualization, Service robot

1. INTRODUCTION

The recent advancement and widespread availability of generative image artificial intelligence (AI) tools have sparked tremendous interest among designers and researchers thanks to their capacity to facilitate design ideation and create realistic images quickly (Hoggenmueller et al., 2023; Rafner et al., 2023; Zhang et al., 2023). Industrial designers have started utilizing text-to-image tools and sketch-to-rendering tools in their design processes for early ideation and fast, realistic visualization of design concepts. As evidenced in the Spring 2023 issue of the Innovation magazine, dedicated to the topic of AI technologies and generative image AI software (IDSA, 2023), as well as IDSA's conferences and events in 2023 and 2024, there have been lively discussions in the industrial design community on AI, sharing and seeking methods of leveraging the potential of generative image AI tools for design processes and design education (Lin & Kim, 2023; Tellez, 2023). Design research has also recognized the potential of generative image AI as research methods and tools to envision novel concepts, reveal current assumptions, and evaluate designs early in the process (Hoggenmueller et al., 2023).

Within human-robot interaction (HRI) research, an early design evaluation through low-fidelity media has been discussed to be effective in assessing preliminary designs and identifying problems quickly before investing in the development of working robot prototypes (Clarkson & Arkin, 2007; Lindblom et al., 2020), which is costly and time-consuming. At the same time, the morphology of service robots requires in-depth investigation and development based on the assessment of how people perceive different designs. The increase of service robots at homes and in public spaces (Wirtz et al., 2018) challenges the widespread approach of anthropomorphism in HRI (Fong et al., 2003; Ishiguro, 2006; Fink, 2012; Phillips et al., 2018) and the questions around their proper appearances, of which industrial designers' contributions have been limited compared to the ones from roboticists and psychologists. Indeed, research shows contradictory effects of anthropomorphic design on people's perception and acceptance of robots. While anthropomorphic design can enhance a robot's understandability and interaction with people (Bartneck et al., 2020), its effects should be understood in relation to multiple aspects: for example, dependency on the use context (Goetz et al., 2003; Roesler et al., 2022), alignment with function (Fink, 2012), and the eerie sensation with the human-like appearance (Mori, 1970). In other words, anthropomorphic design may not always be the answer. This calls for heightened attention from industrial designers regarding the appropriate level of human-likeness of service robots, as human lives are gradually surrounded by a wide array of service robots in different functions, appearances, and use contexts. As mentioned above, an early design evaluation based on low-fidelity media, such as renderings, can be an efficient and effective way of measuring the level of human-likeness of a service robot and defining its appropriate appearance.

Accordingly, this paper taps into leveraging generative image AI's capacity to create variations of realistic renderings quickly for evaluation studies in HRI and examines the subsequent AI-assisted workflow. The second author participated in the study as a co-creative designer with AI to generate variations of service robot designs in four different use contexts. For each service context, an ideation sketch of a service robot in a three-quarter view was drawn manually by the second author and then imported into generative image AI software applications. Using a combination of text-to-image and sketch-to-image models and undergoing iterations of image generation and refinements, variations of service robot designs were developed and visualized as user evaluation media for a study of human perception of robot morphology. However, this paper does not cover the evaluation study but focuses on the AI-assisted workflow, examining its process and discussing the opportunities and challenges of co-creating with generative image AI. This study's contributions are twofold. Firstly, it proposes a novel application of generative image AI for design research in HRI: it leverages AI's capacity to generate design variations and visually situate them in context for the preparation of representation materials for evaluation studies. Secondly, it offers insights into the strengths and challenges of using generative AI for design practice and research.

2. BACKGROUND

2.1 SERVICE ROBOT MORPHOLOGY AND DESIGN EVALUATION

The appearance of a service robot is dependent on the context. Roesler et al.(2022) argued that the preferred level of human-likeness for a robot varies across different use contexts. Service robots can be classified according to personal or professional use (International Federation of Robotics, n.d.), which can also be interpreted as the distinction between domestic and public spaces in terms of their operational environment. Within this classification, the use contexts further diversify according to specific applications: for example, at-home care and household robots for domestic environments; restaurant serving, airport cleaning, supermarket, and delivery robots for public spaces. Each context requires robots equipped with functions specific to their environment, exhibiting their unique task nature. Particularly, the nature of the tasks determines the level of social functions of the robot, depending on how much the tasks involve human interaction or physical demands, which subsequently affects the desired level of human-likeness (Roesler et al., 2022). People expect greater human-likeness in robots that interact more closely and frequently with humans during their tasks.

Researchers in HRI have aimed to understand the preferred level of human-likeness for robot appearance and enhance their acceptability (Goetz et al., 2003; Phillips et al., 2018; Kunold et al., 2023). The most common method in such studies involves presenting representations of robots to study participants through various types of media (e.g., still images, videos, physical robots) and evaluating their perception of and interaction with the robots. Multiple robot designs are prepared and compared by one of the following approaches: creating a single design and comparing it with existing robots (DiSalvo et al., 2002), generating design variations for a specific body part (Goetz et al., 2003), collecting or comparing images of existing robots (Phillips et al., 2018; Randall & Šabanović, 2023), or gathering images of different designs for the same body part from existing robots (Kalegina et al., 2018). In conventional product development, industrial designers create multiple design options for decision-making at gate reviews or focus-group studies along the design process. Similarly, for a more accurate evaluation of human perception of robot appearance, it is desirable to compare and evaluate multiple design variations within the same context and development requirement, not partially but presenting their designs as a whole in the context. However, the high cost of robot development has made it prohibitive to invest in the development of design variations of a robot, especially just to evaluate desired human-likeness. Even creating CAD-based renderings for multiple robot designs requires a significant investment in time and resources.

2.2 GENERATIVE IMAGE AI FOR DIVERGENT DESIGNS AND SCENARIO CREATION

In their examination of the co-creative workflow between human designers and AI in interactive product design, Chiou et al. (2023) highlighted generative image AI's capacity for divergent ideation and its applicability in generating images for concept, scenario, and form. They observed the divergent results from abstract prompts facilitated the divergent nature in the design process. Among the three image themes, AI-generated scenario images were found to be effective in conveying "the interaction between the artifact, user, and environment" (Chiou et al., 2023). In the field of architectural design, Zhang et al. (2023) examined how sketches used as inputs in generative image AI can assist designers in producing design variations and quickly creating realistic images from sketches. This approach avoids time-intensive tasks while maintaining fidelity to the original design intent. Vizcom (https://vizcom.ai), a sketch-to-image generation tool, has gained popularity among industrial designers thanks to its capacity to quickly and faithfully translate early design ideas represented in sketches into realistic renderings. Subsequently, this paper finds the aspects of divergent ideation and realistic scenario image creation in generative image AI useful for the study of robot morphology. By presenting variations of robot designs in context, it addresses the desirable evaluation approach mentioned above. Moreover, it reduces the heavy time investment through AI's ability to swiftly convert sketches into renderings.

3. AI-ASSISTED WORKFLOW AND RESULTS

A first-person method (Chien & Hassenzahl, 2017; Lucero et al., 2019) was employed in this paper for the exploration and reflection of the AI-assisted workflow. As a co-creative designer collaborating with AI, the second author developed and visualized service robot designs in four different use contexts. Due to the lack of prior experience with generative image AI software during the study, the second author had to explore and develop a workflow of using generative image AI tools, specifically for creating evaluation materials related to HRI research. The study spanned seven weeks in the fall of 2023, from late October to early December. During this period, the second author dedicated 10 hours per week to exploring the development of the AI-assisted workflow and creating and visualizing service robot designs. The authors of this paper selected four types of service robots that were either already deployed in real-world environments or soon-to-be-launched products: robots for restaurants, supermarkets, delivery, and household use. The sequence of the study is as follows.

STEP 1: CONTEXT MAPPING

As a first step of the study, the four robot types were arranged on a 2x3 matrix based on the division between domestic and public spaces, as well as the degree of human-robot interaction (Figure 1).

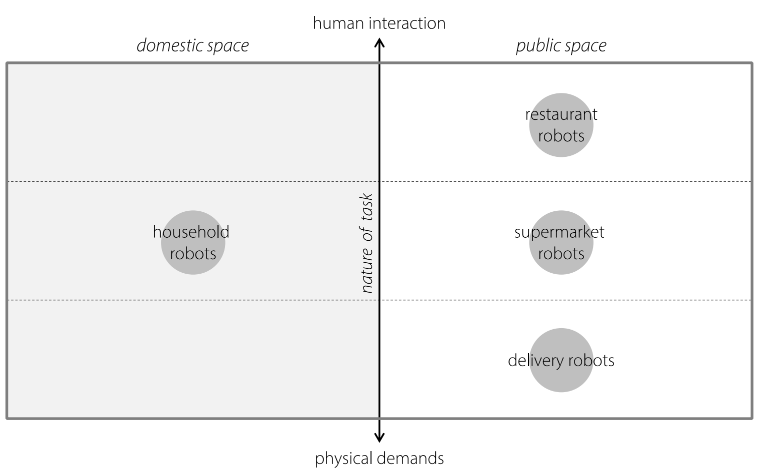

Figure 1. Service robot mapping according to context

The vertical axis represents the nature of tasks, along which the robot types were arranged based on the authors' speculation regarding how much and how frequent service robots in each context would socially interact with human users compared to their physical work. For instance, restaurant robots might closely and frequently interact with customers to greet, receive orders, serve dishes, and process payments. In contrast, delivery robots might have brief interactions with customers solely for package delivery, while supermarket or household robots could fall in between, requiring more physical tasks than restaurant robots but still interacting with human users more frequently than delivery robots.

STEP 2: IDEATION SKETCHING

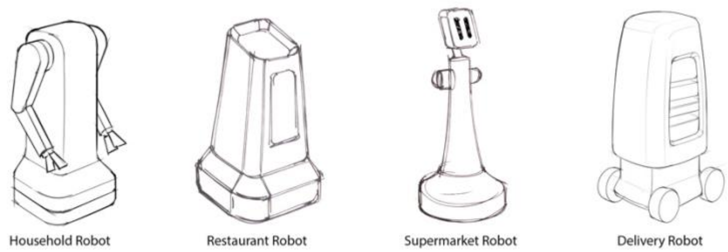

Figure 2. Sketch drawings of service robot designs

In the subsequent phase, a series of ideation sketches was rapidly created for each designated robot type (Figure 2). These sketches, each completed in one to three minutes through analog sketching techniques, aimed to conceptualize both the form and function of each robot type corresponding to its position on the 2x3 matrix. From this set of sketches, a key sketch was chosen from each category for further development.

STEP 3: RENDERING GENERATION

Figure 3. Initial AI-generated outputs of service robot designs

The ideation sketches were imported into AI software to be rendered with applied materials and within accurate lighting composition. Each sketch was uploaded into the "Render" environment of Vizcom, accompanied by a text prompt outlining the viewpoint, subject matter, and desired design aesthetics. For instance, one prompt specified "Front 3/4 View, Robot, Modern, Minimal Surfacing." The initial AI-generated renderings provided foundational representations of materiality and lighting but lacked contextual backgrounds and discernable design details (Figure 3). While these renderings mostly aligned with the original design intentions, the AI occasionally misinterpreted some sketch elements, resulting in unintended design features. For example, some contour lines were transformed into parting lines, functionally obscure components were added, and a screen display with facial expressions turned into a seemingly control panel. Although these unexpected outcomes may inspire new ideas, the authors deemed them undesirable for the study's purpose, necessitating additional measures to fix them.

STEP 4: CONTEXT INTEGRATION

Figure 4. An initial AI-generated rendering of a service robot in the supermarket context

The rendered outputs were imported into Photoshop to be integrated into contextual backgrounds aligned with each robot's intended environment (Figure 4). Leveraging the Photoshop generative AI tool (i.e., Generative Fill), a background image matching the robot's perspective and lighting was created. Robot concepts were extracted from their original Vizcom outputs and pasted into the foreground of the contextual images generated by Photoshop's AI software. Subsequently, manual painting tools were also utilized to blend the robot concepts seamlessly into the AI-generated background.

STEP 5: REFINEMENT

Figure 5. Final AI-generated outputs of service robot designs

After integrating the rendered robot image into its contextual background, manual design refinements were applied using brush tools in Photoshop to enhance design intent, material properties, and lighting effects on the robot itself. The modified image was then imported into the "Refine" environment of Vizcom. Once again, a text prompt accompanied the image, outlining the rendered viewpoint, subject matter, and desired aesthetic, now with the inclusion of contextual environment. At this stage, one text prompt read, "Front 3/4, Robot Rover, Modern, Minimal Surfacing, In a Restaurant." The AI software generated a photorealistic rendering of the refined robot concept, seamlessly integrated within the specified context. Figure 5 displays the four different types of service robots rendered in their specific contexts.

STEP 6: VARIATION AND SELECTION
The final AI-generated output was uploaded back into the Vizcom "Refine" environment. By making subtle adjustments to the text prompts and the influence of the original image, a spectrum of design variations was explored for each robot type, ranging from *objectified* to *human-like* robot morphology. Using this method, over twenty distinct design concepts were rapidly generated within Vizcom for each robot type. From these iterations, three concepts were chosen in varying levels of human-likeness as the definitive representations for each of the four robot types. The selection process yielded a total of twelve final outputs, collectively showcasing a diverse range of human-likeness. For example, Figure 6 illustrates the design variations for delivery robots arranged along the human-likeness level scale. This arrangement is based on the authors' assumption, aiming to demonstrate how these variations can serve as evaluation material for measuring human-likeness levels and examining the influence of robot morphology on human perception of robot designs.

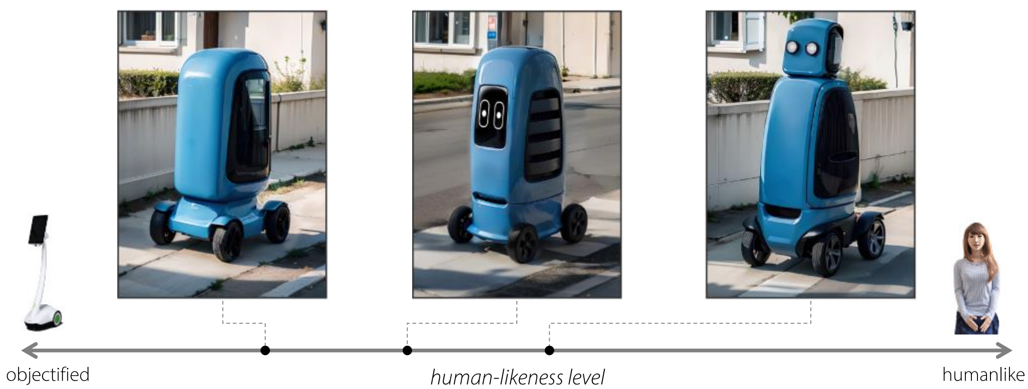

Figure 6. Final selection of delivery robot design concepts and their arrangement on the human-likeness scale

4. DISCUSSION

Figure 7 summarizes the AI-assisted workflow described in the previous section. It illustrates the iterative collaboration between a human designer and AI across the six steps, from context mapping to final design selection. In this process, Photoshop and Vizcom were used iteratively to integrate contextual backgrounds and manually refine images. This approach allowed greater control by the human designer, rather than relying solely on AI for the generation of renderings and risking unintended outcomes. Figure 7 also highlights how the strengths of AI, extracted from generative image AI literature, were applied to this workflow. In addition to its capacity for realistic visualization by converting hand-drawn sketches to renderings, the scenario creation aspect was adopted for context integration and the divergent ideation aspect was utilized for the creation and visualization of design variations. These all contributed to creating evaluation materials for a study of human perception of robot morphology.

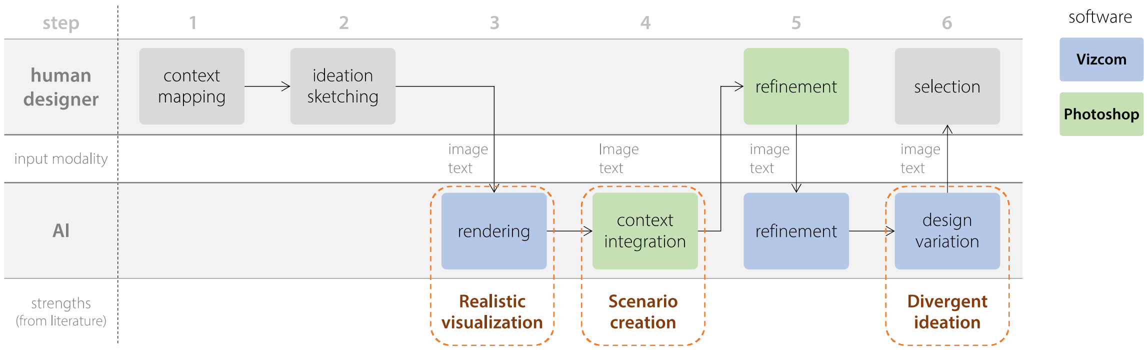

Figure 7. AI-assisted workflow

In this study, the use of generative image AI tools was found to be most effective during the early stages of ideation, especially when concepts are still relatively underdefined. One significant advantage of AI tools lies in their ability to swiftly progress from preliminary sketches to high-fidelity renderings. However, AI tools like Vizcom are limited when it comes to controlling design details and functional or interactive elements within design concepts. As mentioned earlier, unintended outcomes, such as misinterpreted sketch elements or obscure components, can undermine the original design intent. Not only do these outcomes negatively affect the accuracy of evaluation studies, but they also extend the time required to develop evaluation materials. To address these challenges, this paper introduced manual sketching and refinement into the workflow. By combining generative image AI tools with manual techniques, an array of design variations can be generated while preserving the original design intent and ensuring convincing material and lighting properties are reflected within the renderings. The primary challenge encountered with an AI-assisted workflow was maintaining alignment between concept development and the established design intent for the subject matter. These findings underscore the importance of foundational skills in industrial design, such as drawing, visualization, and aesthetics. These skills remain essential for discerning successful AI outputs that capture design intent and creating outcomes that fully reflect the designer's vision.

5. CONCLUSION

This study explored a novel application of generative image AI in HRI research to create and visualize design variations of service robots for evaluation studies related to robot morphology. Leveraging generative image AI's strengths in realistic visualization, scenario creation, and divergent ideation, the study introduced a six-step AI-assisted workflow. The workflow begins with sketches as input and incorporates manual refinement to maintain the original design intent while situating robots in context and creating variations. The study produced three design variations of a service robot in varying degrees of human-likeness for each of four selected service contexts. These designs were rendered in context and prepared for an evaluation study. However, this study is limited in that it lacked an actual evaluation study, leaving the outcomes from the proposed workflow unverified for their effectiveness. Future research should measure the human-likeness levels of these design variations based on this study's rendering outcomes.

6. REFERENCES

Bartneck, C., Belpaeme, T., Eyssel, F., Kanda, T., Keijsers, M., & Šabanović, S. (2020). *Human-robot interaction: An introduction*. Cambridge University Press.

Chien, W. C., & Hassenzahl, M. (2017). Technology-mediated relationship maintenance in romantic long-distance relationships: An autoethnographical research through design. *Human–Computer Interaction*, *35*(3), 240-287.

Chiou, L. Y., Hung, P. K., Liang, R. H., & Wang, C. T. (2023). Designing with AI: An Exploration of Co-Ideation with Image Generators. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (pp. 1941-1954).

Clarkson, E., & Arkin, R. C. (2007). Applying Heuristic Evaluation to Human-Robot Interaction Systems. In *FLAIRS* (pp. 44-49).

DiSalvo, C. F., Gemperle, F., Forlizzi, J. & Kiesler, S. (2002). All robots are not created equal: the design and perception of humanoid robot heads. In *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques* (pp. 321-326).

Fink, J. (2012). Anthropomorphism and human likeness in the design of robots and human-robot interaction. In *Social Robotics: 4th International Conference, ICSR 2012, Chengdu, China, October 29-31, 2012. Proceedings 4* (pp. 199-208). Springer Berlin Heidelberg.

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and autonomous systems*, *42*(3-4), 143-166.

Goetz, J., Kiesler, S., & Powers, A. (2003). Matching robot appearance and behavior to tasks to improve human-robot cooperation. In *The 12th IEEE International Workshop on Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003.* (pp. 55-60). IEEE.

Hoggenmueller, M., Lupetti, M. L., Van Der Maden, W., & Grace, K. (2023). Creative AI for HRI design explorations. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 40-50).

IDSA (2023). *Innovation, 42*(1). Industrial Designers Society of America.

International Federation of Robotics. (n.d.). Service Robots. Available at: https://ifr.org/service-robots/

Ishiguro, H. (2006). Interactive humanoids and androids as ideal interfaces for humans. In *Proceedings of the 11th international conference on Intelligent user interfaces* (pp. 2-9).

Kalegina, A., Schroeder, G., Allchin, A., Berlin, K., & Cakmak, M. (2018). Characterizing the design space of rendered robot faces. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 96-104).

Kunold, L., Bock, N., & Rosenthal-von der Pütten, A. (2023). Not all robots are evaluated equally: the impact of morphological features on robots' assessment through capability attributions. *ACM Transactions on Human-Robot Interaction*, *12*(1), 1-31.

Lin, C. & Kim, B. (2023). Exploration of AI-Driven Furniture Design Education: A Workshop Case Study. In *IDSA Education Symposium*. Industrial Designers Society of America.

Lindblom, J., Alenljung, B., & Billing, E. (2020). Evaluating the user experience of human–robot interaction. *Human-Robot Interaction: Evaluation Methods and Their Standardization*, 231-256.

Lucero, A., Desjardins, A., Neustaedter, C., Höök, K., Hassenzahl, M., & Cecchinato, M. E. (2019). A sample of one: First-person research methods in HCI. In *Companion Publication of the 2019 on Designing Interactive Systems Conference 2019 Companion* (pp. 385-388).

Mori, M. (1970). The uncanny valley. *Energy, 7*(4), 33-35.

Phillips, E., Zhao, X., Ullman, D., & Malle, B. F. (2018). What is human-like? decomposing robots' human-like appearance using the anthropomorphic robot (abot) database. In *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction* (pp. 105-113).

Rafner, J., Zana, B., Dalsgaard, P., Biskjaer, M. M., & Sherson, J. (2023). Picture This: AI-Assisted Image Generation as a Resource for Problem Construction in Creative Problem-Solving. In *Proceedings of the 15th Conference on Creativity and Cognition* (pp. 262-268).

Randall, N. and Šabanović, S. (2023). A Picture Might Be Worth a Thousand Words, But It's Not Always Enough to Evaluate Robots. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 437-445).

Roesler, E., Naendrup-Poell, L., Manzey, D., & Onnasch, L. (2022). Why context matters: the influence of application domain on preferred degree of anthropomorphism and gender attribution in human–robot interaction. *International Journal of Social Robotics*, *14*(5), 1155-1166.

Tellez, F. A. (2023). Exploring A.I. in Design Education: Enhancing Ideation with A.I. Text-to-Image Generators in Digital Design Courses. In *IDSA Education Symposium*. Industrial Designers Society of America.

Wirtz, J., Patterson, P. G., Kunz, W. H., Gruber, T., Lu, V. N., Paluch, S., & Martins, A. (2018). Brave new world: service robots in the frontline. *Journal of Service Management*, *29*(5), 907-931.

Zhang, C., Wang, W., Pangaro, P., Martelaro, N., & Byrne, D. (2023). Generative Image AI Using Design Sketches as input: Opportunities and Challenges. In *Proceedings of the 15th Conference on Creativity and Cognition* (pp. 254-261).