26TH INTERNATIONAL CONFERENCE ON ENGINEERING AND PRODUCT DESIGN EDUCATION
5-6 SEPTEMBER 2024, ASTON UNIVERSITY, BIRMINGHAM, UNITED KINGDOM

COMPARING UNDERSTANDABILITY OF HAND SKETCHES VERSUS AI-GENERATED RENDERS FOR PRODUCT DESIGN

Kristin BARTLETT, Jonathan MILLS and Carly HAGINS
University of Kentucky, United States of America

ABSTRACT
Sketching to communicate design ideas is an important step in the design process. Image-generative artificial intelligence (AI) tools are increasing in prevalence and popularity, some of which focus specifically on sketch-based inputs. The primary aim of this study was to investigate whether an AI-based sketch tool would improve the understandability of students' design concepts. In the context of an undergraduate engineering and product design course, we introduced Vizcom, an AI tool which uses sketch-based inputs to generate renders. In this paper, we describe a quantitative analysis of a dataset of students' hand drawn concepts and the resulting Vizcom render outputs. The sketches and renders were rates separately using standardized rubrics. Our analysis of these ratings indicated that there was no significant difference in understandability between the hand sketches and the AI-generated renders. The characteristics of the hand sketches: line quality, proportionality, and understandability, were all positively correlated with the proportionality and understandability of the AI-generated renders. Our results suggest that the use of Vizcom did not reduce the need for strong hand-sketching skills in communicating design concepts, but Vizcom may offer some benefits in quickly conveying realistic materials and contexts. Results may be different for other types of design concepts and other populations.

Keywords: Artificial Intelligence, product design, design education, design sketching, understandability

1 INTRODUCTION
This paper describes an application of generative artificial intelligence (AI) to student work in a collaborative undergraduate design course. We describe an initial quantitative comparative analysis of a set of concept sketches generated by hand using digital sketching software (most students used Procreate with an Apple pencil on an iPad), and the analogous concept images generated using an AI-based tool called Vizcom. Vizcom software pairs a visual interface with a text-based input and is designed to create high-quality rendered images from a combination of sketch and text input. The primary aim of this analysis was to investigate whether Vizcom helped the students communicate their ideas more clearly than hand sketches. As far as we are aware, there is not yet any published work exploring the effectiveness of using AI to help designers represent their ideas at the concept sketching stage. Artificial intelligence tools are likely to change the way product designers sketch, and the present work aims to investigate how.

2 BACKGROUNDS
2.1 The role of design sketching in the design process
Design sketches used at different phases of the design process have different roles, and sketching is treated differently in engineering design and product/industrial design. In product design, *ideation sketches* are used to explore a wide variety of ideas and/or variations of the same idea. These sketches are typically looser and may resemble 'doodles' or 'thumbnails' [1]. *Concept sketches* are more detailed than ideation sketches and begin to explore more technical solutions and feasible ideas [1]. *Presentation sketches* are more detailed and realistic than concept sketches and are typically presented to a client or external audience to provide a clear and inspiring image of what the final product could look like [1]. In engineering design education, artistic sketching skills are not as highly emphasized. While sketching is

fundamental to design brainstorming and documenting ideas, engineering and technology education place little emphasis on the instruction of design sketching [2]. In this study, we analyse a collection of sketches that we will classify as “concept sketches.”

While art-based drawing courses might focus on training students to create realistic representations of existing objects, the primary aim of teaching designers to draw is to apply established techniques of drawing to representations of physical artifacts which do not currently exist and therefore must be imagined and communicated outwardly. Thus, design-based sketching instruction relies heavily on methods for “constructing” the new concept on the page, producing a representation of a 3-dimensional artifact using 2-dimensional media. A wide range of student prior background in drawing/sketching means first establishing a foundation of drawing elements — the “must haves” of line quality, plane, volume, perspective, and proportion — before advancing to the “nice to haves” of shading, shadow, materiality, context, sequence, and so on. Application of these basic elements to increasingly complex, existing target artifacts — real objects, products, or devices that increase in formal complexity from simple boxes to hand-held tools, allows students to build up their skill and confidence before approaching the challenges of visually representing non-existent artifacts. Once the basic “vocabulary” of drawing is established, focus shifts to fluency. Through practice and exploration, students begin to learn when, how, and where to employ differing levels of fidelity and resolution of their sketches.

Many design sketching courses begin with analogue sketching by hand on paper with simple tools (pencil or pen), before progressing to digital sketching programs. Digital sketching software applications like Procreate offer an overwhelming array of simulated analogue tools (paint, ink, texture, etc.) along with additional functions not easily available in analogue drawing such as layering, mistake-correction, and automation.

2.2 Generative AI in design

Authors García-Peñalvo & Vázquez-Ingelmo define generative AI as the “production of previously unseen synthetic content, in any form and to support any task, through generative modelling” [3]. The use of image-generative AI is increasingly being explored for design applications. Image-generative AI models rely on large datasets of images paired with text-based descriptions which are used to train the models. Diffusion models gradually add noise to image data, then run in a reverse process to denoise the data, resulting in a synthetic image [5].

So far, the applications of generative AI in design have mixed findings. In fashion design, generative AI created many 2D images of designs that were not distinguishable from human-created designs [4], and in architectural design, the generative AI created many designs that would be impossible to construct [5]. Cai et al. investigated the use of a generative AI design tool which creates inspiration mood boards for art and design, and participants found the generative tool more useful for inspiration compared to a traditional image search on Pinterest [6]. Generative AI tools should not necessarily be used uncritically by designers. Bartlett & Camba raised concerns about ethical issues of image-generative AI in product design education, including biased outputs, theft from artists whose work is included in the training data without consent, lack of originality in outputs or “AI copying,” hidden labour in generating the datasets used to train AI, and lack of copyright protection of generated works [7].

Vizcom is a web-based application that allows users to import sketches, 3D models, or draw directly into the interface. These image or model-based inputs are accompanied by a text prompt where the user describes what the model is supposed to be depicting. Figure 1 shows an example of a sketch-based input accompanied by a text prompt “white partial glove with green plastic part” on the left, and on the right is an output image generated by Vizcom. Vizcom outputs are typically high fidelity, similar to what a designer might be produce using detailed hand shading for a final presentation sketch or through computer rendering from a 3D model.

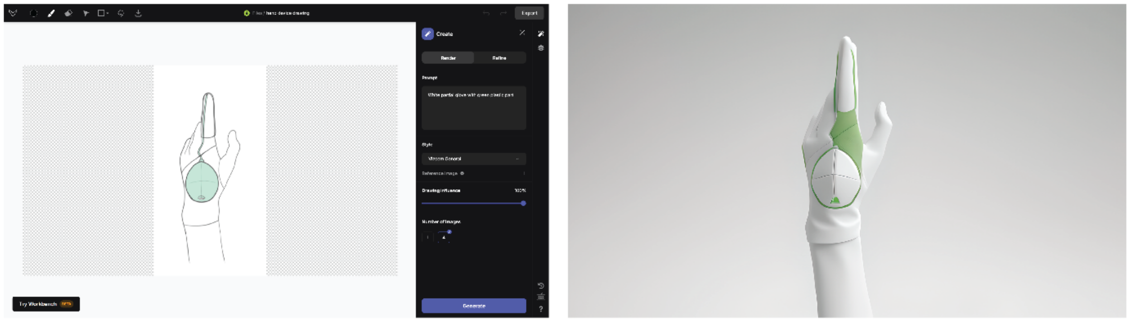

Figure 1. Example of Vizcom input interface (left) and output image (right)

3 METHODS

In our course, student teams comprised of both product design and biomedical engineering students interviewed physicians to identify problem areas in various medical disciplines and designed solutions around these problem areas. During this collaborative project, the product design students were tasked with creating concept sketches of the five final concepts each team devised. To aid the students in creating their concept sketches, the instructor introduced the use of a sketch-based generative AI tool, Vizcom. Using the Pro version of Vizcom software, the students transformed digital sketches into higher-fidelity renderings of medical device design concepts.

Nine third year product design students (five women, four men) depicted five concepts as both a hand sketch and a Vizcom render. Students were encouraged to use the digital sketching software Procreate to create their hand sketches, although one student opted to draw with pencil on paper. The students were provided one hour of in-class training in how to use Vizcom to aid in their design explorations. Thus, the students were novice users of the software.

The final set of images for analysis contained 41 sketches and 41 corresponding Vizcom renders, indicating that a few students submitted individually incomplete assignments. Each sketch depicted a medical device concept idea, and each sketch was used as an input to generate the Vizcom render. Examples of sketches and corresponding Vizcom renders are shown in Figure 2. If the sketches included any hand-written notes, these notes were obscured prior to analysis.

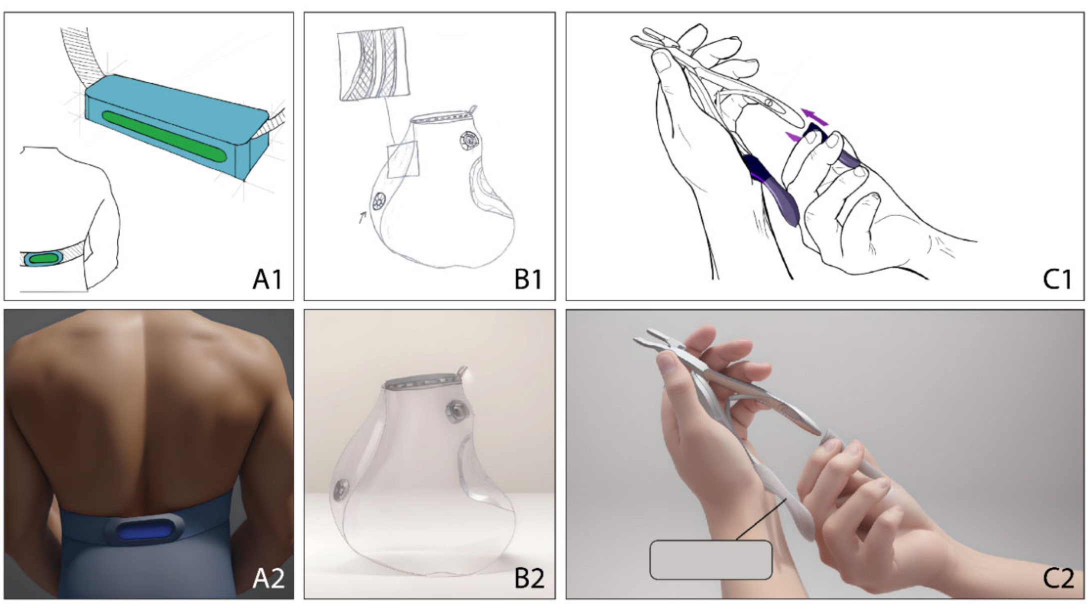

Figure 2. Examples of hand sketches and accompanying Vizcom renders analysed in this study

To provide a basis for comparing the sketches and Vizcom renders, we used a rubric devised by Das & Yang for evaluating design sketches [8]. Das & Yang devised this rubric to assess primarily early-stage sketches from the brainstorming stage. The rubric assesses the sketches on three criteria: line smoothness, proportion/accuracy, and understandability [8]. For each criterion, drawings are rated on a scale of 1 to 5, 1 being the lowest, and 5 being the highest.

In this study, three raters independently evaluated the drawings. Rater 1 was the product design professor who taught the collaborative course for which the sketches were created, and thus knew the context of each sketch. The other two raters were product design professors with extensive experience teaching design sketching. Of these other raters, rater 2 was somewhat familiar with the concepts being sketched, and rater 3 was not familiar with the concepts. Each hand-drawn sketch was rated on all three criteria in Das & Yang's rubric: line smoothness, proportion/accuracy, and understandability. Because the Vizcom renderings are more photorealistic and do not include cartoon-like outlines, the Vizcom renderings were rated on only two criteria: proportion/accuracy and understandability. Hand sketches and Vizcom renders were assessed separately. The order of the sketches and Vizcom renders was randomized so that it would not be immediately obvious to the raters which hand sketches had been used to create which Vizcom renders. Following the ratings, we calculated the interrater reliability, the correlations between the different dimensions scored on the rubric, and we compared the means of the understandability scores for the hand sketches and corresponding Vizcom renders.

4 RESULTS

Intraclass Correlation Coefficients were calculated to determine interrater reliability. We used a two-way mixed model because every rater coded every drawing, but the raters were not random in that they were somewhat familiar with the student projects [9]. Average measures were used rather than single measures since average measures will be used in the rest of the analysis. The Intraclass Correlation Coefficients are reported in Table 1.

Table 1. Intraclass Correlation Coefficients between raters

Rubric Dimension	Intraclass Correlation – Avg. Measures	95% CI Lower Bound	95% CI Upper Bound	Sig.
Hand sketch line qual.	.820	.697	.898	<.001
Hand sketch prop.	.875	.790	.929	<.001
Hand sketch under.	.875	.789	.929	<.001
Vizcom prop.	.861	.766	.921	<.001
Vizcom under.	.882	.802	.933	<.001

The Intraclass Correlation measures were all above .80, indicating good interrater reliability [9]. Rater 1, who was the instructor of the class, gave the drawings highest average scores for all categories, rater 2, who was somewhat familiar with the projects, gave drawings middle ratings, and rater 3, who was least familiar, gave lowest averages for all categories.

Correlations, based on averages of the three raters' scores, were calculated between each of the different rubric categories. The correlations are reported in Table 2. All correlations were statistically significant at the 0.01 level (2 tailed). All the rubric categories were positively correlated with one another: hand sketch line quality, hand sketch proportion, hand sketch understandability, Vizcom render proportion, and Vizcom render understandability.

Table 2. Pearson Correlations Coefficients between rubric criteria

	Hand sketch prop.	Hand sketch under.	Vizcom prop.	Vizcom under.
Hand sketch line qual.	.65**	.68**	.73**	.66**
Hand sketch prop.		.79**	.73**	.57**
Hand sketch under.			.65**	.74**
Vizcom prop.				.70**

** $p < .001$

A paired samples t-test was used to evaluate whether there was a difference in the understandability ratings between Vizcom renders and hand sketches. The results indicated that there was no significant difference between the hand sketch understandability (M = 3.53, SD = 1.08), and the Vizcom render understandability (M = 3.62, SD = 1.00), t (40) = .761, p = .226.

5 DISCUSSIONS

Hand sketches that had good line quality, proportion, and understandability were positively correlated with proportional, understandable Vizcom renderings. Vizcom renders were not significantly different than the root hand sketches in terms of understandability. Taken together, these results suggest that Vizcom, in its' present form, is not likely to replace or diminish the value of being able to sketch by hand. Regarding the strengths of Vizcom, we see a strength in depicting the human form in much greater detail than what the students' hand sketches contained. For example, the person's back shown in image A2 of Figure 2 contains much more detail conveyed through colour and shading than the simple outline of the person's back contained in image A1 of Figure 2. As showing context is very important for design communication, and many concepts in product design relate to the human body, enhancing this imagery may be a valuable strength of using Vizcom in product design. However, it is well-known that image-generating AI can also generate uncanny or disturbingly inaccurate depictions of the human body, although newer models are working on improving this. Image B2 of Figure 2 demonstrates another strength of using Vizcom, which is showing materials in higher fidelity. While Image B1 is not a very clear drawing, and the clarity of the lines is not necessarily improved in B2, what is improved in B2 is that the viewer gets a sense that the bag is made of a clear material. The clarity of materials was not something that is captured in the rubric that we used to assess drawings in this study.

We do not believe that the raters' familiarity with the concepts necessarily impacted their ratings on the "understandability" category, because although rater 3, who was least familiar with the concepts, gave the lowest ratings for understandability, rater 3 also gave the lowest ratings in the other categories. Thus, this discrepancy could be related to the interpretation and application of the rubric rather than familiarity with concepts. The rubric we used was developed in the context of engineering design, a discipline in which students receive less training in sketching and the purposes of sketching are fundamentally different than in product design. We chose to use this rubric because it was the most relevant published rubric available. Were we to adapt the rubric for product design, we might add a wider range of scores, since this rubric may lead to a ceiling effect when used with product design sketches.

One limitation of this study is that the participants were third year product design students who had one hour of formal training using Vizcom software. Beyond this training, any learning in Vizcom was done independently, and some students likely spent more time than others exploring the software. However, none of the students would be classified as experts in using Vizcom. It is possible that expert users of Vizcom might produce different results than what was found in this study. Another limitation is the fact that this data set contained only medical device concepts. Medical device concepts are likely to be very underrepresented in the tagged image data used to train the AI models that Vizcom runs, in comparison to more commonly represented objects like cars and shoes. It is possible that a different type of project that focuses on the design of a more common object would yield different results. Another limitation is the fact that the raters were all product design professors with some experience teaching design sketching. Results may be different with different populations of raters.

As future work, we are interested in analysing the drawings at an individual student level. While there is no significant difference in the whole population between the understandability of hand sketches and Vizcom renders, it is possible that some students deviated from this trend depending on their level of sketching skill and style. It would also be helpful to analyse the text-based prompts.

6 CONCLUSIONS

At present, the sketch-based AI image-generating software Vizcom does not appear to replace hand sketching skills, though it does offer important benefits. We found no significant difference in understandability between the concept sketches produced by hand and the corresponding Vizcom renders. An individual's ability to produce high quality Vizcom renders appears to be related to their ability to sketch by hand, as the traits of line quality, proportionality, and understandability of hand sketches were positively correlated with proportionality and understandability of Vizcom renders. We did find that Vizcom provided some benefits in representing materials more realistically, and in providing more realistic context for devices used on the body. Our results were obtained in the specific context of novice-Vizcom-user students sketching medical device concepts, and the results may not hold true with other populations or in other contexts.

In the product design industry, the traditional craft of sketching by hand is still highly valued by some professionals and educators. The results of this study suggest that sketch enthusiasts do not necessarily need to be concerned that sketch-based AI tools like Vizcom will fully replace the need for traditional sketching skills. At present, individuals who want to master sketching with Vizcom will likely still need to practice both hand sketching and control of the software tool. However, Vizcom soon may replace the need for creating high-fidelity presentation sketches, which require realistic shading and materials, by hand. Creating such sketches by hand is quite time consuming, so the time savings afforded by AI-based tools like Vizcom is likely to promote their adoption.

ACKNOWLEDGEMENT

The authors would like to thank Vizcom for providing free Pro licenses for the product design students and instructor to use in this course.

REFERENCES

[1] Eissen K. and Steur R. *Sketching: The Basics*, 2011 (BIS Publishers).

[2] Kelley T. R. Design Sketching: A Lost Skill. *Technology and Engineering Teacher*, 2017, 76(8), pp.8–12.

[3] García-Peñalvo F. and Vázquez-Ingelmo A. What Do We Mean by GenAI? A Systematic Mapping of The Evolution, Trends, and Techniques Involved in Generative AI. *International Journal of Interactive Multimedia and Artificial Intelligence*, 2023, 8(4), 7-16. doi: 10.9781/ijimai.2023.07.006.

[4] Sbai O., Elhoseiny M., Bordes A., LeCun Y. and Couprie C. DesIGN: Design Inspiration from Generative Networks. In *Computer Vision – ECCV 2018 Workshops*, Vol. 11131, L. Leal-Taixé and S. Roth, Eds., in Lecture Notes in Computer Science, 2019, pp.37–44 (Cham: Springer International Publishing). doi: 10.1007/978-3-030-11015-4_5.

[5] Zhang C., Wang W., Pangaro P., Martelaro N. and Byrne D. Generative Image AI Using Design Sketches as input: Opportunities and Challenges. In *Creativity and Cognition*, Virtual Event USA, June 2023, pp.254–261 (ACM). doi: 10.1145/3591196.3596820.

[6] A. Cai et al. DesignAID: Using Generative AI and Semantic Diversity for Design Inspiration. In *Proceedings of The ACM Collective Intelligence Conference*, Delft Netherlands, November 2023, pp.1–11 (ACM). doi: 10.1145/3582269.3615596.

[7] Bartlett K. A. and Camba J. D. Generative Artificial Intelligence in Product Design Education: Navigating Concerns of Originality and Ethics. *International Journal of Interactive Multimedia and Artificial Intelligence*, 2024, 8(5), pp.55-64. doi: 10.9781/ijimai.2024.02.006.

[8] Das M. and Yang M. C. Assessing Early Stage Design Sketches and Reflections on Prototyping. *Journal of Mechanical Design*, 2022, 144(4), pp.1-11. doi: 10.1115/1.4053463.

[9] Koo T. K. and Li M. Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 2016, 15(2), pp.155–163. doi: 10.1016/j.jcm.2016.02.012.