

3D Dreams Are Made of This - Comparing Sketching on XR and Tablet Devices as Input Modalities for 3D Model Generation

Niklas Pfützenreuter
HCI Group
University of Duisburg-Essen
Essen, Germany
GENERIO
Essen, Germany
niklas.pfuetzenreuter@uni-due.de

Jonathan Liebers
Research Center Trustworthy Data
Science and Security
Dortmund, Germany
TU Dortmund University
Dortmund, Germany
jonathan.liebers@tu-dortmund.de

Roman Heger
HCI Group
University of Duisburg-Essen
Essen, Germany
roman.heger@uni-due.de

Henning Maly
HCI Group
University of Duisburg-Essen
Essen, Germany
henning.maly@stud.uni-due.de

Steven Tischer
HCI Group
University of Duisburg-Essen
Essen, Germany
steven.tischer@stud.uni-due.de

Anh Duy Dao
HCI Group
University of Duisburg-Essen
Essen, Germany
anh.dao@stud.uni-due.de

David Goedicke
HCI Group
University of Duisburg-Essen
Essen, Germany
david.goedicke@uni-due.de

Stefan Schneegass
HCI Group
University of Duisburg-Essen
Essen, Germany
stefan.schneegass@uni-due.de

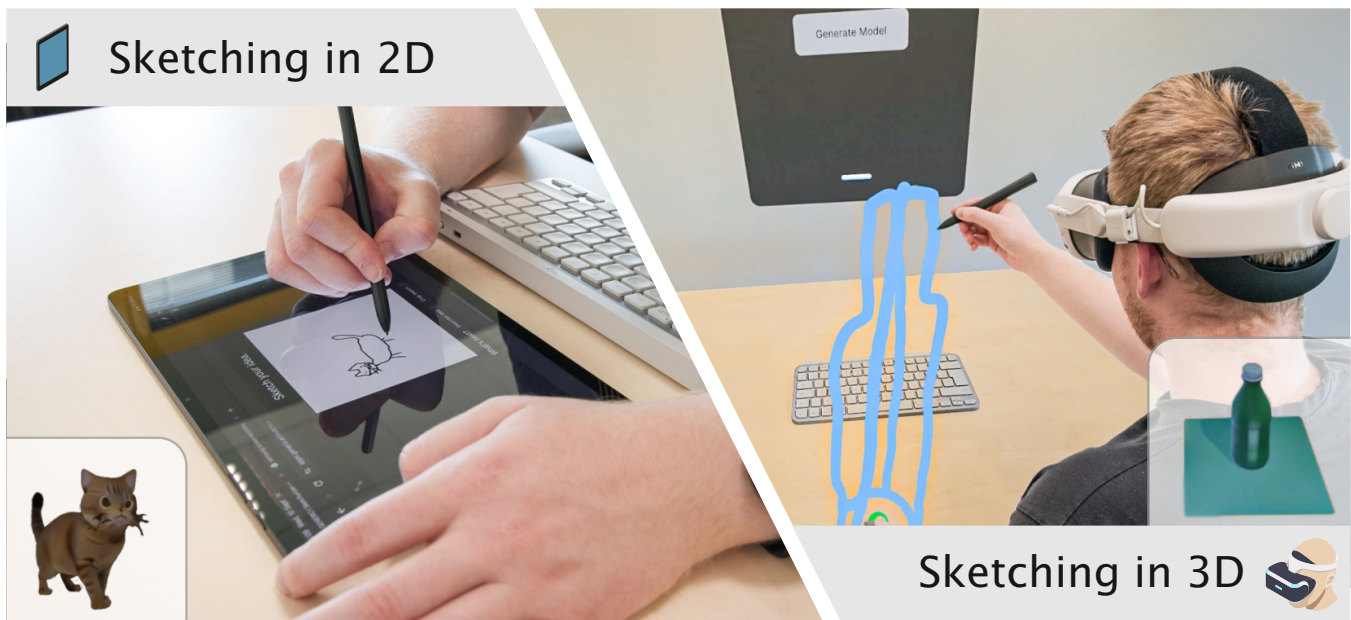


Figure 1: In 3D Dreams Are Made of This, our goal is to understand how the sketching dimension affects the user experience, control, and object quality when interacting with 3D object generation models. Therefore, we compare sketching in 2D and 3D as input modalities for a 3D object generation model in our user study.

Abstract

Generative AI (GenAI) enables the creation of 3D objects from textual descriptions that can be used in game development or 3D printing. However, users struggle to express their vision because control mechanisms are bound to 2D space, resulting in unsatisfyingly generated objects. To address this, we introduce an Extended Reality (XR) sketching application that offers a new way to use 3D sketches to control the generation of 3D objects with GenAI. We compare this approach to a 2D sketching application on a tablet in a study ($N=16$), both methods using pen-based input. Results show that 3D sketching enables faster specification than 2D sketching, due to reduced mental effort and shorter ideation times. However, perceived control decreases with increasing object complexity, and this negative effect is stronger for 3D sketching. These findings highlight the potential of 3D sketching as a control modality for generative 3D object creation, while identifying object complexity as a key challenge for 3D sketching interfaces for GenAI.

CCS Concepts

• **Human-centered computing** → **Empirical studies in HCI**.

Keywords

Generative Artificial Intelligence, Virtual Reality, 3D Model Generation

ACM Reference Format:

Niklas Pfützenreuter, Jonathan Liebers, Roman Heger, Henning Maly, Steven Tischer, Anh Duy Dao, David Goedicke, and Stefan Schneegass. 2026. 3D Dreams Are Made of This - Comparing Sketching on XR and Tablet Devices as Input Modalities for 3D Model Generation. In *Proceedings of Mensch und Computer 2026 (MuC '26)*, August 30–September 02, 2026, Duisburg, Germany. ACM, New York, NY, USA, 13 pages. <https://doi.org/10.1145/3820253.3820276>

1 Introduction

With the ongoing advancements of GenAI for text and image generation, we also see a growing number of approaches focusing on 3D object generation [39]. These approaches use images or text prompts to generate 3D meshes for use in video games or for 3D printing [33]. In such visual contexts, it has been shown that pure prompt-based generation lacks user control [11]. Hence, current approaches for 3D object generation aim to increase the users' control over the generation process and output. This is done by using reference images of the desired object or by user-created digital sketches of the object outlines [11]. Especially the latter enables a direct and intuitive approach to controlling object generation [25]. Further, it can support creativity, as users can sketch whatever comes to mind without being limited to a preexisting image of their vision [11]. However, existing control mechanisms, such as sketching or reference images, limit the user's input to 2D. The third dimension can only be simulated through isometric, i.e., 3D-projection-based, sketching, view rotation, or using multiple 2D perspectives [4, 26].

While such 2D abstractions work quite well as input for GenAI models to generate simple geometric shapes, the models tend to fail to transfer from 2D to 3D space as object complexity increases. This occurs because complex geometries often obscure important details when viewed from a single perspective [29]. Here, from a model perspective, enabling input in 3D space can improve the quality of 3D object generation by providing more relevant details. Thus, the GenAI models profit from a more detailed control input [11]. From a user perspective, this also limits users' control, as sketching in 3D space has been shown to be beneficial over 2D sketching due to increased spatial perception [19].

Therefore, in this paper, we introduce a control mechanism for 3D object generation with GenAI using 3D sketching in Extended Reality (XR) and compare this to a 2D sketching approach. We focus on 3D sketching in XR as it not only enables 3D input like tools that use view rotation [4], but also a real-world 3D perception of the sketched object [27]. We investigate the effects of 3D sketching on the perceived quality of generated 3D objects, user experience, and control, compared to 2D sketch-based object generation. We further reflect this effect in dependence on the geometrical complexity of the desired 3D object, as it was found that the complexity has an influence on the sketching performance [30, 36]. To address our research question, we developed an XR application that generates 3D objects from 3D sketching input. Users can draw outlines of objects using a 3D-tracked pen on a table surface and mid-air. The user-created 3D sketches are used to generate multi-view control images for the 3D generation model. We compare this XR application to a commercial 3D object generation application with 2D sketching input on a tablet in a user study ($N = 16$). We used pen-based input for both conditions, as this proved the most natural, easy to use, and accurate for sketching in 2D and 3D [7].

Our results show that users are faster at sketching in 3D than in 2D, due to a shorter ideation phase. Users reported needing less mental effort to draft sketches in their minds and thus started sketching objects more quickly. Yet users were more familiar with 2D sketching and had higher precision. Further, our results show that sketching simple objects leads to higher perceived control than sketching complex objects. However, we did not find significant differences in perceived user control or generated object quality between 2D and 3D, as both had their limitations. In 2D, on the one hand, the users struggled to express the three-dimensionality of the objects. In 3D, on the other hand, users reported challenges with perceiving sketch depth and drawing straight lines. Thus, further investigation is needed into how to increase user control in 2D and 3D for 3D object generation processes.

Contribution Statement. Our work makes the following contributions: We implement an XR sketching application that uses 3D sketches to generate 3D objects using GenAI. We compare this XR sketching application to a commercial 3D object generation application with 2D sketch input in a user study ($N = 16$). Our findings indicate that 3D sketching is faster than 2D sketching, yet 2D sketching is more familiar and provides higher precision.

2 Related Work

Our work compares and is based on understanding 2D and 3D sketching as input modalities for 3D object generation with GenAI,



This work is licensed under a Creative Commons Attribution 4.0 International License. *MuC '26, Duisburg, Germany*

© 2026 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-2611-8/26/08

<https://doi.org/10.1145/3820253.3820276>

examining their effects on user experience and control, and object quality. Further, we examine metrics for measuring object complexity.

2.1 3D Object Generation

GenAI enables the creation of 3D objects from representations like textual descriptions or 2D images [35]. Thereby, the approaches transfer representations to the 3D space using techniques such as NeRFs or SLATs [24, 35]. The generated objects are used for game development, 3D printing, or to build research prototypes [23]. However, these approaches are lacking user control, especially regarding the geometry of the generated objects [11]. As a result, users often struggle to generate objects that align with their visions [18]. To increase user control, many approaches rely on control mechanisms for 2D image generation [9]. These include using reference images or sketches, which are mostly limited to the 2D space [26]. This, on the one hand, limits the amount of control users can exercise, and on the other hand, the amount of information the GenAI models can extract [9]. More recent approaches to increase user control over 3D object generation aim to use 3D sketching in virtual reality [10]. However, they focus on creating simple geometric shapes and do not address more complex objects [10]. Thus, it remains unclear how the control mechanism in 3D space influences users' perceived control and object quality for more complex objects.

2.2 Digital Sketching in 2D and 3D Space

Digital sketching can be done using mouse or touch input, digital pens, or controllers in Virtual Reality (VR) [12, 40]. The generated digital sketches are used in domains such as design, engineering, and art [38]. In 2D space, most applications work on desktop devices and use mouse input for drawing lines or also predefined shapes [5]. However, this is often found to be unintuitive and complicated when creating freehand sketches [16]. In contrast, tablet applications that focus on digital pens enable a more natural interaction, comparable to sketching on paper [16]. Therefore, users find sketching with digital pens on tablet devices to be more accurate and usable than on desktop devices [16]. To sketch three-dimensional content in 2D space, users can apply techniques such as isometric sketching, which shows an object from a specific perspective and visualizes its depth [3]. Furthermore, sketching applications can create a 3D view of the object by allowing the sketch to be rotated [37]. Using generative design algorithms, 2D sketches can also be converted to 3D sketches [20].

Instead of simulating the third dimension, sketching applications in VR enable complete three-dimensional sketching [12]. This has been shown to help users with the spatial perception of 3D content and sketches, thereby enhancing user control compared to approaches that simulate 3D in 2D [27, 32]. However, there are also challenges regarding the accuracy and the usability of the VR sketching applications [12]. To solve this, approaches use surfaces like tablets that support users while sketching [1, 12].

2.3 Virtual Object Complexity

To measure the complexity of virtual objects, different metrics exist. One metric by Globa et al. [14] considers the number of

components, i.e., geometric primitives such as spheres or cubes, in the objects. Further, the metric considers symmetry and the number of transformations, such as rotations or translations, to construct the virtual object [14]. A second metric by Gero et al. is based on the number and arrangement of faces, edges, and vertices of the virtual object [13]. All in all, the metrics define that the more components or faces a virtual object has, or the more unsymmetrical an object is, the higher its complexity [28, 29]. This object complexity also affects how easy it is to sketch the objects, both in 2D and 3D. For more complex objects, users find it more challenging to sketch them compared to simple objects [30, 36].

3 Experiment

In this work, we compare 2D and 3D sketching as input modalities for 3D object generation with GenAI. 2D sketching is a state-of-the-art modality to add additional control to 3D object generation processes [9]. While there are approaches that enable 3D sketching in 2D by using view rotation [4] or generative design algorithms [20], they are limited in the perception of the 3D content compared to XR [27]. Thus, our work focuses on 3D sketching in XR, which more recent approaches are also trying to implement as an input modality for 3D object generation models [10]. While 3D sketching has been shown to improve users' spatial thinking and user experience [32], it remains unclear how it influences perceived user control and object quality, including object complexity. Thus, we assess the following research question:

RQ: How do the input dimensionality (2D and 3D) and the object complexity (simple and complex) influence the interaction with 3D generation models?

3.1 Hypotheses

Considering the background of 3D object generation and sketching in 2D and 3D, we formulate and investigate the following three hypotheses:

H1: Sketching simple objects leads to better (a) user control and (b) perceived object quality compared to sketching complex objects.

H2: Sketching in 3D leads to better (a) user control and (b) perceived object quality compared to sketching in 2D.

H3: Sketching in 3D particularly increases (a) user control and (b) perceived object quality compared to 2D sketching for complex objects.

Our first hypothesis serves as a manipulation check to verify that our chosen objects have the intended level of complexity and that this complexity has the intended influence on sketching performance and object generation. Previous investigations regarding the influence of the object quality have shown that users perform significantly worse when sketching complex objects [30, 36]. Further, it has been shown that 3D object generation models struggle with the higher-level details of complex objects [26]. Thus, we can assume that user control and perceived object quality are higher for simple objects than for complex ones.

For our second hypothesis, we assume that sketching in 3D increases the user control over the object generation process. This is based on findings from previous comparisons between 2D and 3D sketching applications, which showed that sketching in 3D enhances users' spatial thinking [32]. Additionally, it has been



Figure 2: In our investigation, we assess the independent variables *Input Dimensionality* (2D and 3D) and *Object Complexity* (simple and complex), both with two levels each. This results in 4 study conditions: (a) 2D-Simple, (b) 2D-Complex, (c) 3D-Simple, and (d) 3D-Complex.

shown that a more embodied interaction, such as in VR, compared to simple 2D screen interaction, can enhance spatial cognition of virtual objects [22]. This can positively affect perceived control over generation, as users can express spatial geometry more effectively. Further, we expect that sketching in 3D can help to increase the perceived quality of the generated objects. Previous research has shown that more perspectives as input for GenAI object generation increase the quality of generated objects [9]. With 3D sketches, it is possible to generate as many perspectives as needed from each sketch, thereby increasing the amount of data for generative models. This can increase the level of detail available to the 3D generation model.

Regarding our third hypothesis, we assume that 3D sketching is particularly beneficial for complex objects compared to 2D sketching. This is motivated by previous investigations that have shown that sketching simple three-dimensional geometries is also possible in 2D [36]. However, the complexity of the object influences users' sketching performance, especially in 2D environments [30]. For more complex objects, users struggle more to sketch them with a 2D system than with a 3D system [17]. Further, a three-dimensional embodied interaction can increase the user control compared to a 2D screen interaction, especially for more complex shapes [21]. Thus, we assume that the advantages of sketching in 3D are particularly evident for complex objects and also positively affect perceived control and object quality during 3D object generation.

3.2 Study Design

We conducted a within-subject repeated-measures lab study, examining the independent variables *Input Dimensionality* and *Object Complexity*. For the *Input Dimensionality*, we chose two levels: 2D and 3D. In 2D, users sketch on a tablet with a digital pen, which is the most natural interaction for 2D digital sketching [3]. In 3D, users sketch in an XR passthrough environment, using a digital pen to keep interactions comparable to 2D sketching. The *Object Complexity* also has two levels: simple and complex. Based on the metrics described in subsection 2.3, we selected *laptop*, *book*, *mug*, and *bottle* as simple objects and *flower bouquet*, *seashell with pearl*, *cat*, and *action figure* as complex objects. Our simple objects were therefore defined as those that are symmetrical on at least one axis, are not self-occluded, and can be easily constructed with fewer than 1000 vertices. The complex objects do not fulfill one or more of these criteria. In our study, we explore all four combinations

of *Input Dimensionality* and *Object Complexity* as conditions. The conditions are displayed in Figure 2.

As dependent variables, we measured task completion time (TCT), workload using NASA-TLX [15], and user experience using the UEQ-short [31]. Further, we measure perceived object quality, sketch easiness, and control using 7-point Likert items, inspired by Wang et al. [34]. For qualitative feedback, we conducted semi-structured interviews with open questions ("What was good when sketching in 2D?", "What was bad when sketching in 2D?", "What was good when sketching in 3D?", "What was bad when sketching in 3D?", "What difference did the object complexity make?", "In which way did the objects differ from your original vision?", "What would you improve regarding the 2D sketching application?", "What would you improve regarding the 3D sketching application?"). All our measurements are shown in Table 1.

3.3 Apparatus

For the 2D sketching and object generation application, we chose the commercial application Sketchurizer¹, which enables sketch-based generation. It is a web application that works on desktop and mobile devices and lets users sketch the outlines of objects on a digital canvas. The sketched lines have a defined thickness, which makes them comparable to pen-and-paper sketching and provides interpretable control input to the GenAI model [41]. Users can erase sketched lines with an eraser tool and can undo and redo actions. The sketches, along with a user-defined prompt, are used to generate an image of the desired object. The generated image is then used to generate the final 3D object. For both the 2D and 3D applications, we use a Logitech MX Keys Mini² keyboard to type in the prompts. The final 3D object can be viewed and rotated in a 3D viewer. In the study, the application is executed on a Samsung Galaxy Tab 6 with the associated pen. Our 3D sketching application in XR was implemented in Unity 6 for the Meta Quest 3 headset. In the application, we used the Quest 3 passthrough feature so users could see the table for sketching. We use the Logitech MX Ink pen³ to enable sketching in XR. The users start sketching by pressing a button on the pen. We set a fixed line thickness and did not provide

¹Sketchurizer, GENERIO, <https://apps.generio.ai/> (last accessed June 1, 2026)

²Logitech MX Keys Mini, Logitech, <https://www.logitech.com/de-de/shop/p/mx-keys-mini> (last accessed June 1, 2026)

³Logitech MX Ink MR Stylus, Logitech, <https://www.logitech.com/de-de/products/vr/mx-ink-mr-stylus.html> (last accessed June 1, 2026)

Table 1: Overview of dependent variables with corresponding measurements.

DV	Measurement
Task performance	Task completion time (TCT)
Workload	NASA TLX [15]
User experience	UEQ-S [31]
Perceived object quality: Correspondence to vision	Likert (7-point): The generated object corresponds to my original vision.
Perceived object quality: Correspondence to sketch	Likert (7-point): The generated object corresponds to the sketch.
Perceived object quality: Correspondence to prompt	Likert (7-point): The generated object corresponds to the prompt.
Sketch easiness: Ease	Likert (7-point): It was easy for me to sketch the object.
Sketch easiness: Quickness	Likert (7-point): I felt like I could sketch the object quickly.
Sketch easiness: Idea expression	Likert (7-point): I was able to express my design idea with the sketch.
Perceived control: Control level	Likert (7-point): I feel to have a good level of control over the generation.
Perceived control: Predictability	Likert (7-point): I was able to influence the generation process in a predictable manner.
Perceived control: Ownership	Likert (7-point): The model was created by me.
Qualitative feedback	Semi-structured interview

basic shape tools (such as spheres or cubes) to keep the application comparable to the 2D sketching application. The sketch can be rotated around its up axis using the controller’s thumbstick. Users can also undo and redo actions and erase sketched lines with an eraser tool. When the users finished sketching, they typed a prompt in a UI and started the generation. For the generation, we rendered 4 images from the users’ sketches (front, back, left, and right), as most multiview 3D generation approaches use 4 perspectives to capture all sides of an object [8]. Further, our used pipeline showed the best generation results using these 4 perspectives. Then we send the sketch images to the GENERIO API to use the 3D generation pipeline as for the 2D application. When the generation is complete, the object is displayed in the XR application at the sketch’s position, and the sketch is hidden. The object can be rotated using the thumbstick on the controller.

3.4 Power Analysis

To determine the sample size for our user study, we conducted an a priori power analysis for a repeated-measures within-subjects analysis of variance (RM-ANOVA) using G*Power 3.1.9.7. Given $\alpha = .05$, $\beta = .05$, effect size $f = .4$, one group (within-subject) and four measurements, we obtain a total suggested sample size of 15 ($\lambda = 19.20$, $F = 2.83$).

3.5 Participants

We recruited 16 participants (11 male, 4 female, and 1 agender) via university mailing lists for our study. They were 26 years old on average ($SD = 6.35$, $Min = 20$, $Max = 43$). When asking how often participants used GenAI for 3D object generation, 4 answered “never”, 8 “occasionally”, 3 “once a month”, and 1 “once a week”. Regarding the use of VR/XR, 5 answered “never”, 6 “occasionally”, 2 “once a month”, 2 “once a week”, and 1 “daily”. The participants

did not receive any monetary compensation for their participation in the study.

3.6 Procedure

Before the study, we briefed participants on the goals and procedures and informed them that their interactions with the applications would be recorded. After they provided written consent and their demographic information, we led a tutorial in which participants generated a table to familiarize themselves with both applications. Upon completion, the participants conducted the four conditions of our study, each with the following sequence. The participants started by sketching the first object. After that, they typed in the prompt and started the generation. The object was reviewed by the participants when the generation was finished. Lastly, the participants completed the NASA-TLX, the UEQ-short, and our custom Likert items. This was repeated for two objects per condition. We counterbalanced the order of *Input Dimensionality* and *Object Complexity* using a Latin square. After all conditions were finished, we conducted the semi-structured interviews. An entire session took about 60 minutes per participant. Examples of the generated objects are shown in Figure 3.

3.7 Ethics

Prior to conducting the study, we obtained approval from our local ethics committee. All participants provided informed consent and were explicitly informed that they could withdraw from the study at any time without any disadvantages. To ensure participants’ privacy, we implemented a pseudonymization approach using a randomized identifier. After completing the study, we permanently deleted the mapping between these identifiers, rendering any re-identification infeasible. Consequently, the data was fully pseudonymized.

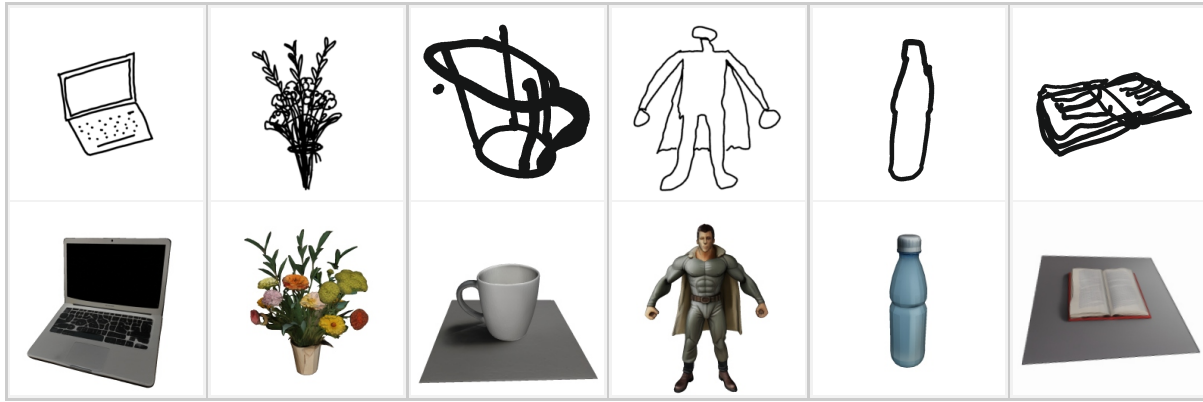


Figure 3: These are examples of objects generated with both systems in the study by our participants. The first row shows their sketches, and the second row shows the generated objects.

4 Results

We measured the TCT, workload, user experience, perceived object quality, sketch easiness, and perceived control. To obtain our TCTs, we measured the time elapsed from the start of the sketching phase in each application to its end in seconds. This also included the time for ideation, as participants needed to develop a strategy for sketching the object. We deliberately chose not to include the object generation time, as it is constant and similar for both 2D and 3D sketching and would have distorted the analysis. For the NASA TLX and UEQ short, we calculated the respective scores. We checked our data for normality using a Shapiro-Wilk test. When the data were not normally distributed, we calculated an Aligned Rank Transform (ART) before conducting a repeated-measures analysis of variance (RM-ANOVA).

4.1 Task Completion Time

First, we explore how *Input Dimension* and *Object Complexity* influence the TCT (in seconds). We found a significant main effect for *Input Dimension* showing that sketching in 2D (Med. = 136.23, IQR = 178.57) led to higher TCTs than sketching in 3D (Med. = 88.25, IQR = 45.38), $F(1, 45) = 5.1808$, $p = .0276$, $\eta_p^2 = .1032$. However, we could not find a significant main effect for *Object Complexity* between simple (Med. = 94.50, IQR = 60.22) and complex (Med. = 117.75, IQR = 117.31), $F(1, 45) = .0494$, $p = .8250$, $\eta_p^2 = .0011$, and no significant interaction effect, $F(1, 45) = .0710$, $p = .7911$, $\eta_p^2 = .0016$. Figure 4a and Table 3 show the results. Despite not finding a significant interaction effect, we found a significant contrast between Complex-2D (Med. = 158.47, IQR = 143.70) and Simple-3D (Med. = 84.75, IQR = 35.50), indicating a significantly higher TCT for Complex-2D compared to Simple-3D, $t(15) = 2.8421$, $p = .0403$.

4.2 NASA TLX

Second, we assess the effects of *Input Dimension* and *Object Complexity* on the NASA TLX scores. For this measure, we could not find any significant main effects between simple (Med. = 5.00, IQR = 16.25) and complex (Med. = 10.00, IQR = 17.50), $F(1, 45) = 3.4393$, $p = .0702$, $\eta_p^2 = .0710$, and 2D (Med. = 5.00, IQR = 14.38) and 3D (Med. = 8.75,

IQR = 18.12), $F(1, 45) = .1216$, $p = .7289$, $\eta_p^2 = .0027$. Further, we could not find a significant interaction effect, $F(1, 45) = .0985$, $p = .7551$, $\eta_p^2 = .0022$. This was also the case for all subscales of the NASA TLX. The main results for the overall scores and all subscales are reported in Table 2.

4.3 User Experience

Next, we explore the effects of *Input Dimension* and *Object Complexity* on the user experience using UEQ short. However, we could not find significant main effects between simple (Med. = 1.59, IQR = 1.20) and complex (Med. = 1.41, IQR = .75), $F(1, 45) = .4707$, $p = .4962$, $\eta_p^2 = .0104$, and 2D (Med. = 1.50, IQR = 1.16) and 3D (Med. = 1.50, IQR = .91), $F(1, 45) = .2114$, $p = .6479$, $\eta_p^2 = .0047$. Further, we could not find a significant interaction effect, $F(1, 45) = .0017$, $p = .9669$, $\eta_p^2 < .0001$. This was also true of the pragmatic and hedonic values. The main results are displayed in Table 3.

4.4 Perceived Object Quality

Following, we measure the effects of *Input Dimension* and *Object Complexity* on the perceived object quality. The ANOVA revealed no significant main effects regarding the correspondence to the vision between simple (Med. = 4.00, IQR = 1.62) and complex (Med. = 3.50, IQR = 1.00), $F(1, 45) = 2.1415$, $p = .1503$, $\eta_p^2 = .0454$, and 2D (Med. = 3.50, IQR = 1.12) and 3D (Med. = 4.00, IQR = 2.00), $F(1, 45) = 2.9365$, $p = .0935$, $\eta_p^2 = .0613$. Further, we could not find a significant interaction effect, $F(1, 45) = .0206$, $p = .8866$, $\eta_p^2 = .0005$. This was also the case for the correspondence to the prompt and the correspondence to the sketch. The main results are displayed in Table 4.

4.5 Sketch Easiness

Next, we assess the effects of *Input Dimension* and *Object Complexity* on the sketch easiness. However, we could not find significant main effects between simple (Med. = 6.00, IQR = 1.12) and complex (Med. = 5.00, IQR = 2.00), $F(1, 45) = 3.3890$, $p = .0722$, $\eta_p^2 = .0700$, and 2D (Med. = 5.50, IQR = 1.75) and 3D (Med. = 5.50, IQR = 1.62), $F(1, 45) = .0099$, $p = .9213$, $\eta_p^2 = .0002$. Further, we could not find a

Table 2: RM-ANOVA results for NASA-TLX overall and subscales. We could not find significant differences for the NASA-TLX scores.

Metric	Effect	F(1, 45)	p-value	η_p^2	Median	IQR
Overall	Complexity	3.44	.070	.071	Com. = 10.00 / Sim. = 5.00	Com. = 17.50 / Sim. = 16.25
	Dimension	0.12	.729	.003	2D = 5.00 / 3D = 8.75	2D = 14.38 / 3D = 18.12
	Interaction	0.10	.755	.002	—	—
Mental	Complexity	1.78	.189	.038	Com. = 20.00 / Sim. = 16.25	Com. = 31.25 / Sim. = 23.12
	Dimension	1.52	.224	.033	2D = 15.00 / 3D = 20.00	2D = 28.12 / 3D = 26.25
	Interaction	0.03	.860	.001	—	—
Physical	Complexity	2.18	.146	.046	Com. = 7.50 / Sim. = 6.25	Com. = 15.00 / Sim. = 11.25
	Dimension	1.38	.246	.030	2D = 7.50 / 3D = 7.50	2D = 10.00 / 3D = 17.50
	Interaction	1.97	.168	.042	—	—
Temporal	Complexity	0.70	.408	.015	Com. = 7.50 / Sim. = 7.50	Com. = 15.00 / Sim. = 11.25
	Dimension	0.02	.884	.000	2D = 7.50 / 3D = 7.50	2D = 15.00 / 3D = 8.75
	Interaction	3.79	.058	.078	—	—
Performance	Complexity	1.64	.207	.035	Com. = 63.75 / Sim. = 70.00	Com. = 23.12 / Sim. = 26.25
	Dimension	0.00	.976	.000	2D = 66.25 / 3D = 63.75	2D = 33.12 / 3D = 26.25
	Interaction	0.25	.622	.006	—	—
Effort	Complexity	1.51	.225	.032	Com. = 23.75 / Sim. = 17.50	Com. = 38.12 / Sim. = 30.62
	Dimension	0.70	.409	.015	2D = 17.50 / 3D = 22.50	2D = 31.25 / 3D = 33.12
	Interaction	0.31	.583	.007	—	—
Frustration	Complexity	3.44	.070	.071	Com. = 10.00 / Sim. = 5.00	Com. = 17.50 / Sim. = 16.25
	Dimension	0.12	.729	.003	2D = 5.00 / 3D = 8.75	2D = 14.38 / 3D = 18.12
	Interaction	0.10	.755	.002	—	—

significant interaction effect, $F(1, 45) = .1236$, $p = .7268$, $\eta_p^2 = .0027$. This was also the case for the sketch quickness and idea expression. The main results for the overall scores and all subscales are reported in Table 4.

4.6 Perceived Control

Following, we explore the effects of *Input Dimension* and *Object Complexity* on the perceived control. We found a significant main effect for *Object Complexity* showing that sketching simple objects (Med. = 4.50, IQR = 1.25) led to a higher perceived control than sketching complex objects (Med. = 3.50, IQR = 1.50), $F(1, 45) = 9.1669$, $p = .0041$, $\eta_p^2 = .1692$. This supports **H1a**. However, we could not find a significant main effect between 2D (Med. = 4.00, IQR = 1.00) and 3D (Med. = 4.25, IQR = 2.12), $F(1, 45) = 1.8124$, $p = .1850$, $\eta_p^2 = .0387$, and no significant interaction effect, $F(1, 45) = 1.1120$, $p = .2973$, $\eta_p^2 = .0241$. The results can be seen in Figure 4b and Table 3. Despite not finding a significant interaction effect, we found a significant contrast between Complex-2D (Med. = 3.50, IQR = 1.12) and Simple-3D (Med. = 4.75, IQR = 2.00), indicating a significantly higher perceived control for Simple-3D compared to Complex-2D, $t(15) = -2.9407$, $p = .0309$.

4.7 Outcome Predictability

Next, we assess the effects of *Input Dimension* and *Object Complexity* on the predictability of the generated outcome. We found a significant main effect for *Object Complexity* showing that sketching

simple objects (Med. = 4.25, IQR = 2.00) led to a higher predictability of the outcome than sketching complex objects (Med. = 3.50, IQR = 1.12), $F(1, 45) = 7.6109$, $p = .0084$, $\eta_p^2 = .1447$. This supports **H1a**. However, we could not find a significant main effect between 2D (Med. = 4.00, IQR = 2.00) and 3D (Med. = 4.00, IQR = 1.62), $F(1, 45) = .1408$, $p = .7093$, $\eta_p^2 = .0031$, and no significant interaction effect, $F(1, 45) = .1555$, $p = .6952$, $\eta_p^2 = .0034$. Figure 5a and Table 3 show the results.

4.8 Perceived Ownership

Lastly, we explore how *Input Dimension* and *Object Complexity* affect the perceived ownership of the generated objects. We found a significant main effect for *Object Complexity* showing that sketching simple objects (Med. = 4.00, IQR = 1.62) led to a higher perceived ownership of the generated objects than sketching complex objects (Med. = 3.00, IQR = 2.00), $F(1, 45) = 11.7934$, $p = .0013$, $\eta_p^2 = .2077$. This supports **H1a**. However, we could not find a significant main effect between 2D (Med. = 3.50, IQR = 1.75) and 3D (Med. = 4.00, IQR = 2.50), $F(1, 45) = .0130$, $p = .9096$, $\eta_p^2 = .0003$, and no significant interaction effect, $F(1, 45) = .6378$, $p = .4287$, $\eta_p^2 = .0140$. The results can be seen in Figure 5b and Table 3. Despite not finding a significant interaction effect, we found a significant contrast between Complex-3D (Med. = 3.00, IQR = 2.00) and Simple-3D (Med. = 4.00, IQR = 1.50), indicating a significantly higher perceived ownership for Simple-3D compared to Complex-3D, $t(15) = -2.9585$, $p = .0295$.

Table 3: RM-ANOVA results for TCT, perceived control, and UEQ scores. We found that users are significantly faster in 3D and that the perceived control is higher for simple objects. However, we could not find significant differences for UEQ scores.

Metric	Effect	F(1, 45)	p-value	η_p^2	Median	IQR
TCT [s]	Complexity	0.05	.825	.001	Com.=117.75; Sim.=94.50	Com.=117.31; Sim.=60.22
	Dimension	5.18	.028	.103	2D=136.23; 3D=88.25	2D=178.57; 3D=45.38
	Interaction	0.07	.791	.002	—	—
Perceived Control	Complexity	9.17	.004	.169	Com.=3.50; Sim.=4.50	Com.=1.50; Sim.=1.25
	Dimension	1.81	.185	.039	2D=4.00; 3D=4.25	2D=1.12; 3D=2.12
	Interaction	1.11	.297	.024	—	—
Outcome Predictability	Complexity	7.61	.008	.145	Com.=3.50; Sim.=4.25	Com.=1.62; Sim.=2.00
	Dimension	0.14	.709	.003	2D=4.00; 3D=4.00	2D=2.00; 3D=1.62
	Interaction	0.16	.695	.003	—	—
Perceived Ownership	Complexity	11.79	.001	.208	Com.=3.00; Sim.=4.00	Com.=2.00; Sim.=1.62
	Dimension	0.01	.910	.000	2D=3.50; 3D=4.00	2D=1.75; 3D=2.50
	Interaction	0.64	.429	.014	—	—
UEQ Overall	Complexity	0.47	.496	.010	Com.=1.41; Sim.=1.59	Com.=0.75; Sim.=1.20
	Dimension	0.21	.648	.005	2D=1.50; 3D=1.50	2D=1.16; 3D=0.91
	Interaction	0.00	.967	0.000	—	—
UEQ Pragmatic	Complexity	0.83	.366	.018	Com.=1.31; Sim.=1.69	Com.=1.22; Sim.=1.31
	Dimension	0.30	.588	.007	2D=1.25; 3D=1.62	2D=1.31; 3D=1.31
	Interaction	0.10	.753	.002	—	—
UEQ Hedonic	Complexity	2.10	.154	.045	Com.=1.75; Sim.=2.00	Com.=1.03; Sim.=1.09
	Dimension	0.10	.759	.002	2D=1.81; 3D=1.88	2D=1.28; 3D=1.12
	Interaction	0.03	.858	.001	—	—

4.9 Qualitative Results

To understand why sketching in 2D took longer and why sketching simple objects improved perceived control, we analyzed the interview responses. Using the thematic analysis [6], the lead researcher analyzed the interview responses. We identified six main themes: Impact of 3D Representations, Ease of Use, Precision Limitations, Generated Object Quality, Impact of Object Complexity, and Application Improvements.

Impact of 3D Representations. Many participants reported that creating 3D representations in the 2D sketching application was difficult. They tried to simulate the third dimension with their sketches, but this was often misinterpreted by the GenAI model, “So, for example, you can suggest that the table is somehow three-dimensional, but that might then be understood as a fifth and sixth table leg, perhaps.” (P10). This negatively influenced the perceived object quality of the objects generated with the 2D application. Further, it made the participants need longer to draft and finally sketch their objects.

Participants liked having a 3D view of the sketch in the XR application, which allowed them to see it from multiple angles. This made it easier to review the spatial properties of the sketch and to express their vision, “Exactly, you don’t have to have abstractions of it or anything like that, you don’t have to project it yourself in 2D from 3D, you can just draw the object as it really is.” (P3). Participants stated that this was the main reason they were faster in 3D than in

2D, as the mental effort required to draft the sketch in their minds was reduced.

Ease of Use. Regarding ease of use, many participants found the 2D sketching application easier to use. They explained this primarily by the familiarity with pen-and-paper sketching, which is very similar to sketching with a digital pen on a tablet device, “Otherwise, what I like about the tablet is that drawing is easier than in 3D because you’re used to drawing things in 2D with a pen.” (P10). This also made it feel faster for the participants to sketch the objects in 2D.

By contrast, sketching in 3D was perceived as more difficult by many participants. This was due to 3D sketching being a novel experience, and the participants needed to get used to it first, “But I think because it’s so unfamiliar and inexperienced, it’s difficult.” (P2). However, they also said that with more practice, this would become easier for them.

Precision Limitations. Some participants noted that both applications had limitations in sketching precision. What was missing for most participants was the ability to change the size or diameter of the sketching tool, “With only one line thickness, and such a thick one at that, it’s really tedious to draw in smaller details.” (P14). This made it difficult for participants to sketch fine details or large surfaces.

In 3D, participants also reported difficulties with line alignment and maintaining the correct depth. This meant that participants were often unable to connect lines successfully on the first try, “But I basically had that somehow, so the lines didn’t really connect the

Table 4: RM-ANOVA results for perceived object quality and sketch easiness metrics. We could not find significant differences for these metrics.

Metric	Effect	F(1, 45)	p-value	η_p^2	Median	IQR
Correspondence Vision	Complexity	2.14	.150	.045	Com.=3.50; Sim.=4.00	Com.=1.00; Sim.=1.62
	Dimension	2.94	.094	.061	2D=3.50; 3D=4.00	2D=1.12; 3D=2.00
	Interaction	0.02	.887	.001	—	—
Correspondence Prompt	Complexity	0.41	.527	.009	Com.=4.00; Sim.=4.50	Com.=1.25; Sim.=2.00
	Dimension	1.29	.262	.028	2D=4.00; 3D=4.50	Com.=1.50; Sim.=2.00
	Interaction	0.43	.515	.010	—	—
Correspondence Sketch	Complexity	2.02	.162	.043	Com.=4.00; Sim.=4.50	Com.=1.50; Sim.=2.12
	Dimension	3.31	.076	.068	2D=4.00; 3D=4.25	Com.=1.50; Sim.=2.50
	Interaction	0.23	.631	.005	—	—
Sketch Easiness	Complexity	1.36	.250	.029	Com.=5.00; Sim.=5.50	Com.=2.00; Sim.=1.62
	Dimension	0.95	.335	.021	2D=5.50; 3D=5.00	Com.=1.62; Sim.=2.00
	Interaction	0.74	.394	.016	-	-
Sketch Quickness	Complexity	0.35	.556	.008	Com.=5.50; Sim.=5.75	Com.=2.50; Sim.=2.00
	Dimension	1.39	.245	.030	2D=5.75; 3D=5.00	Com.=1.62; Sim.=2.12
	Interaction	0.02	.890	.000	-	-
Sketch Idea Expression	Complexity	3.39	.072	.070	Com.=5.00; Sim.=6.00	Com.=2.00; Com.=1.12
	Dimension	0.01	.921	.000	2D=5.50; 3D=5.50	Com.=1.75; Sim.=1.62
	Interaction	0.12	.727	.003	-	-

way I wanted them to, the way I wanted them to.” (P14). This was also amplified by the struggle to draw straight lines when switching from the surface-based sketching on the table to mid-air sketching.

Generated Object Quality. Regarding the perceived quality of the generated objects, most participants felt that those generated with the 2D sketching application matched their original vision more closely. They said that the GenAI model was able to convert even simple sketches into high-quality objects, “He got a lot out of my very simple sketches for the 2D stuff.” (P5). However, sometimes this interpretation of the sketches also led to unwanted details in the generated object.

In the 3D application, participants reported that the generated objects often deviated from their original vision. The generated objects often contained additional layers and planes or other artifacts when generating complex geometries like human faces, “So there were many additional layers and surfaces in the 3D object, i.e., in the 3D sketch, that I didn’t really want to have in there and hadn’t drawn in either.” (P1). This made the generated objects look less realistic and less accurate for the 3D application.

Impact of Object Complexity. The object complexity was perceived to have a major impact on sketching ease and the quality of the generated objects in both applications. Most participants mentioned that complex objects are much harder to sketch and that the quality of the generated objects was much lower, “The more complex the sketch task was, the worse the generation was, because you just couldn’t do it. I couldn’t do it very well. If I couldn’t draw it very well myself, then I think it’s logical that the generation wasn’t very good either.” (P2). The drop in quality was explained by the

fact that the GenAI models had more difficulty in understanding the complex sketches.

Specifically for the 3D sketching application, participants said that complexity also depended on how natural or imprecise the object’s shape was. This means that more precise shapes, like a bottle, were perceived as more difficult to sketch than more natural shapes like a flower bouquet, “But it was much more difficult in 3D, for example, to make this bottle or the cup as complex objects.” (P9). The reason was that natural shapes forgive more mistakes when sketching than more precise shapes do.

Application Improvements. Regarding both applications, the participants had many suggestions for improvement. Many participants wished to have an adjustable line thickness to improve sketching accuracy, “And the option to make the pen smaller or larger. That would be very helpful.” (P11). More feedback for improvements to both applications included the ability to sketch with colors and zoom functionality for sketches.

Specifically for the 2D sketching application, the participants requested a larger canvas and the ability to move or rotate the canvas, “Actually, I found it rather difficult to have the smaller sheet as well, because then you are limited, and with the cat, I found there was a bit too little space to draw.” (P12).

For the 3D sketching application, it was requested to have a multi-axis rotation and a tool for drawing straight lines, “I think it would be cool if you had some kind of tool, like a ruler or something, that you could use, or something like that, an analogy to a ruler in 3D, or some kind of thing that you could use to draw straight lines and stuff like that.” (P3). Further, the participants wished for tools to create surfaces and predefined geometric shapes such as cubes or spheres.

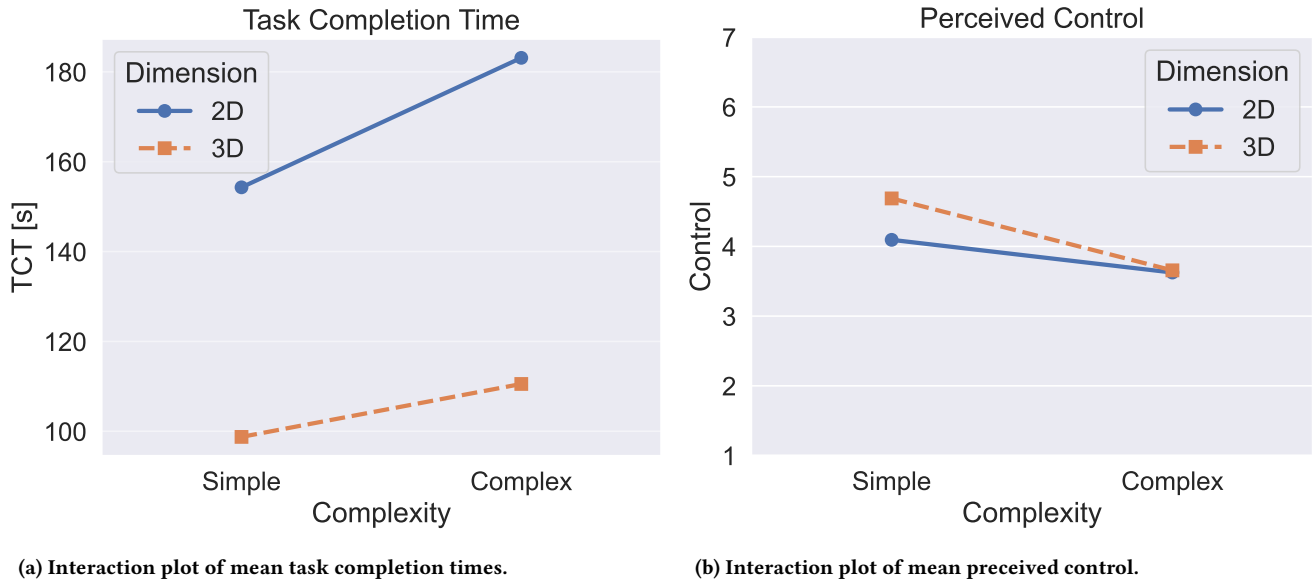


Figure 4: We found that the task completion time is higher in 2D than in 3D and that the perceived control is higher for simple objects than for complex objects.

5 Discussion

We discuss the findings of our assessment below. While we found that simple objects lead to higher perceived control and that users are faster when sketching in 3D, we did not find significant differences in perceived control or in the quality of generated objects between the two dimensions.

5.1 Influences on the Perceived Control

Our results show that the perceived control is significantly higher for simple objects than for complex objects. The predictability of the outcome and perceived ownership confirm this finding, as both are higher for simple compared to complex objects. Therefore, we can accept **H1a**. However, we did not find significant differences in perceived control between 2D and 3D sketching. Thus, we cannot accept **H2a**. Further, we found that perceived control is significantly higher for Simple-3D than for Complex-2D, indicating that sketching simple 3D objects yields the highest perceived control, whereas sketching complex 2D objects yields the lowest. The much larger drop in perceived control in the 3D application from simple to complex, compared to the 2D application, indicates that sketching in 3D is more negatively affected by increasing complexity than sketching in 2D. We also found this trend in perceived ownership, which was significantly higher for Simple-3D than for Complex-2D. This contradicts **H3a**, and thus we cannot accept it.

The differences between simple and complex objects are in line with previous findings, showing that users struggle more in sketching complex objects compared to simple objects [30]. However, the non-presence of differences between 2D and 3D and the higher drop for complex objects in 3D compared to 2D are unexpected in light of previous findings, which found that sketching in 3D can help with the spatial perception, especially for complex objects [32]. A possible explanation for the different results in our

investigation is the slightly different sketching task compared to previous studies. Where previous investigations defined sketching as a more coarse task for solid shapes and where users used tools like brushes or shaping tools for ideation, in our investigation, users were asked to sketch the exact outlines of the objects they wanted to generate. This requires higher precision and relies solely on pen-based drawing. Thus, the participants mentioned struggling to be as precise as required, especially in the 3D application. A further limitation in 3D sketching is that the XR prototype did not explicitly handle depth-consistent occlusion between the pen and already drawn strokes. This may have weakened important visual depth cues, thereby making it difficult to perceive sketch depth and accurately place lines. As a result, lines often did not connect properly, requiring additional corrections. This effect was also caused by users sketching in mid-air rather than relying on the table surface, resulting in a decreased precision [12]. In the end, the generated sketches were perceived as harder for the GenAI model to interpret. For the 2D sketching application, this was not as much of a problem as the participants stated. This was due to being used to pen-and-paper sketching, which is similar to our 2D sketching task on a tablet with a digital pen. However, we did not explicitly measure or control participants' prior sketching expertise. Differences in drawing ability and familiarity with pen-based 2D or spatial 3D sketching may therefore have influenced sketch quality, task completion time, and perceived control, suggesting that future studies should assess sketching experience or use repeated sessions to gain experience with both systems.

5.2 No Differences Regarding Perceived Object Quality

Regarding perceived object quality, we did not find any significant differences between simple and complex objects, or between 2D and

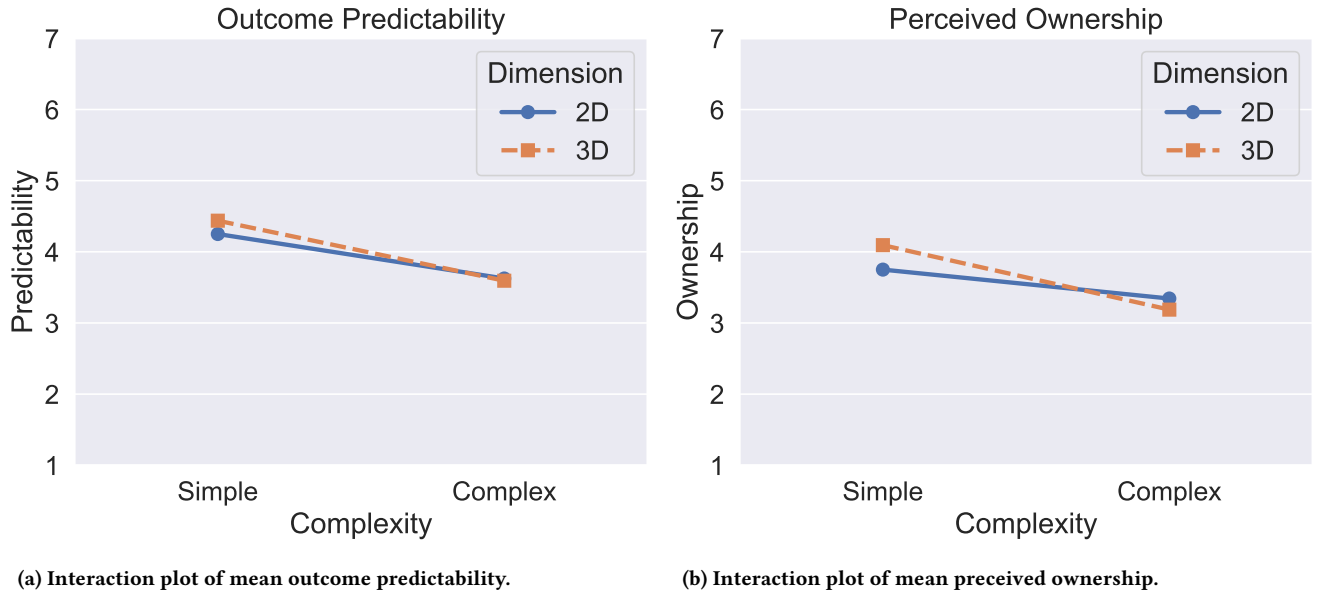


Figure 5: We found that the predictability of the outcome and the perceived ownership are higher for simple than for complex objects.

3D sketching. There were only tendencies in the data that sketching in 3D led to higher perceived object quality than sketching in 2D. However, we cannot accept **H1b** and **H2b**. This is not in line with previous findings, as it was found that simpler geometries, as well as more detailed control input through more perspectives, can increase the quality of generated objects [11]. The reasons are that the study may have been underpowered to detect small or moderate effects, and both applications had limitations. In 2D, on the one hand, the perceived quality was limited by the challenge of conveying three-dimensionality, and thus, sketches were often misinterpreted. In 3D, on the other hand, the aforementioned limited precision and difficulty in sketching straight lines resulted in lower-quality sketches and, consequently, lower-quality objects. With that, the advantage of having control images from multiple perspectives in 3D was also neutralized, and it was even counterproductive, as the participants mentioned. A limitation is that object quality was assessed only through participants' subjective Likert-scale ratings, which capture perceived correspondence to the vision, sketch, and prompt but do not provide an objective measure of generation quality. Future work should complement these ratings with external evaluations, such as expert judgments, independent, blind ratings, or computational similarity metrics between the intended object, the sketch input, and the generated 3D model.

5.3 Users are Faster in 3D

Although we did not find differences between 2D and 3D in perceived control and object quality, users are significantly faster when sketching in 3D than in 2D. Further, we found that users were significantly faster in Simple-3D compared to Complex-2D. This indicates that users were especially fast at sketching simple 3D objects but struggled more with complex 3D than with complex 2D objects,

contradicting **H3b**. This is not in line with previous findings that users spend more time sketching in 3D than in 2D [17]. The reason sketching in 3D was faster than in 2D, in our case, was that we included the ideation phase in our timing. The participants noted that they needed more mental effort and time in 2D to understand how to express the third dimension within the 2D environment. In contrast, they reported simply starting to sketch in 3D, as they could sketch the object as it is. This saved a lot of time and, in the end, could compensate for the challenges with the 3D sketching interaction. However, further analysis regarding behavioral or motion-level measures, such as movement trajectories, idle time, or interaction patterns, could provide more detailed insights and should be examined more directly in future work. Further, our findings should be interpreted not as isolating the effect of dimensionality alone, but as comparing two concrete implementations of sketch-based control for 3D generation. Future work should use more closely matched and sophisticated systems, including comparable drawing aids and precision support, to uncouple the effects of input dimensionality from implementation-specific limitations.

5.4 Potential of Sketching in 3D

All in all, we identified a potential that sketching in 3D may improve perceived control and object quality during 3D object generation. While our investigation was influenced by the concrete implementation of the tools and found no significant differences between 2D and 3D for perceived control and object quality, participants' qualitative feedback revealed tool-independent insights. Participants explicitly stated that the biggest difference between 2D and 3D sketching was that they were more familiar with 2D sketching due to the pen-and-paper-like interaction. With more practice, however, they said, they would become better at sketching in 3D

and sketch more precisely. Thus, it would be interesting to investigate how additional practice in 3D sketching by the participants would influence our results in a long-term study. Additionally, with more practice, the biggest advantage of 3D sketching would become even more apparent. Participants said they liked that they could sketch the objects with a reduced mental effort to figure out how to generate the three-dimensionality. As shown above, this has a positive effect on the task completion times. What could further improve sketching in 3D would be tools that support users while sketching [2]. We chose to let users sketch on a table in 3D, as previous research has found that mid-air sketching is less usable [12]. However, users often started sketching in mid-air anyway when setting the height of the objects. Furthermore, the surface could not solve the problems with the perception of the object's depth. Tools that can help create straight lines, like having a movable surface, as in the approach of Drey et al. [12], or even simple shapes, could drastically improve this. Additionally, broadening the design space beyond our tablet-2D and XR-3D conditions by comparing additional input modalities, such as 3D sketching on 2D displays with view rotation and other established sketch-based workflows, could solve some of these challenges [4, 20].

In addition to the advantages for users, using 3D sketches directly rather than images can further improve control inputs for 3D object generation. We focused on using multi-perspective image control, as this is a state-of-the-art control mechanism [11]. However, the 3D sketches can potentially also be used as input to initialize 3D generation procedures such as NERFs or SLATs [24, 35].

6 Conclusion

This work explored how the input dimension (2D vs. 3D) and the object complexity (simple vs. complex) influence interactions with 3D object generation models. In our user study ($N = 16$), participants sketched objects of varying complexity in 2D and 3D and rated the interaction and the outcome on workload, user experience, perceived object quality, sketch ease, and perceived control. Further, we measured the TCT as the time participants needed to complete the sketches. We show that users are significantly faster when sketching in 3D, mainly because ideation time is shorter, suggesting the mental effort to draft 3D sketches is lower in a 3D environment. However, users reported greater precision in 2D sketching due to its familiarity. Further, we found that perceived control is significantly higher for simple than for complex objects, confirming that complex geometries are harder to express through sketching. However, we did not find any differences in perceived control or object quality between 2D and 3D. This was due to challenges with expressing three-dimensionality in 2D on the one hand and with sketch precision in 3D on the other. Nevertheless, the participants liked the low mental effort required for 3D sketching. Thus, future investigation should focus on improving precision when sketching in 3D or on using the actual 3D representation as a control input to better leverage the advantages of 3D sketching.

References

- [1] Rahul Arora, Rubaiat Habib Kazi, Tovi Grossman, George Fitzmaurice, and Karan Singh. 2018. SymbiosisSketch: Combining 2D & 3D Sketching for Designing Detailed 3D Objects in Situ. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–15. doi:10.1145/3173574.3173759
- [2] Rahul Arora, Rubaiat Habib Kazi, Fraser Anderson, Tovi Grossman, Karan Singh, and George Fitzmaurice. 2017. Experimental Evaluation of Sketching on Surfaces in VR. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, Denver Colorado USA, 5643–5654. doi:10.1145/3025453.3025474
- [3] Sriram Karthik Badam, Senthil Chandrasegaran, Niklas Elmquist, and Karthik Ramani. 2014. Tracing and Sketching Performance Using Blunt-Tipped Styli on Direct-Touch Tablets. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces*. ACM, Como Italy, 193–200. doi:10.1145/2598153.2598182
- [4] Seok-Hyung Bae, Ravin Balakrishnan, and Karan Singh. 2008. ILoveSketch: As-Natural-as-Possible Sketching System for Creating 3d Curve Models. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology (UIST '08)*. Association for Computing Machinery, New York, NY, USA, 151–160. doi:10.1145/1449715.1449740
- [5] Sukanya Bhattacharjee and Parag Chaudhuri. 2020. A Survey on Sketch Based Content Creation: From the Desktop to Virtual and Augmented Reality. *Computer Graphics Forum* 39, 2 (2020), 757–780. doi:10.1111/cgf.14024
- [6] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. doi:10.1191/147808706qp063oa
- [7] Alberto Cannavò, Davide Calandra, Aidan Kehoe, and Fabrizio Lamberti. 2021. Evaluating Consumer Interaction Interfaces for 3D Sketching in Virtual Reality. In *Interactivity and Game Creation*, Anthony Brooks, Eva Irene Brooks, and Duckworth Jonathan (Eds.). Springer International Publishing, Cham, 291–306. doi:10.1007/978-3-030-73426-8_17
- [8] Minghao Chen, Roman Shapovalov, Iro Laina, Tom Monnier, Jianyuan Wang, David Novotny, and Andrea Vedaldi. 2025. PartGen: Part-level 3D Generation and Reconstruction with Multi-View Diffusion Models. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Nashville, TN, USA, 5881–5892. doi:10.1109/CVPR52734.2025.00552
- [9] Minglin Chen, Weihao Yuan, Yukun Wang, Zhe Sheng, Yisheng He, Zilong Dong, Liefeng Bo, and Yulan Guo. 2024. Sketch2NeRF: Multi-view Sketch-guided Text-to-3D Generation. arXiv:2401.14257 [cs]
- [10] Tianrun Chen, Chaotao Ding, Shangzhan Zhang, Chunan Yu, Ying Zang, Zejian Li, Sida Peng, and Lingyun Sun. 2024. Rapid 3D Model Generation with Intuitive 3D Input. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Seattle, WA, USA, 12554–12564. doi:10.1109/CVPR52733.2024.01193
- [11] Yang Chen, Yingwei Pan, Yehao Li, Ting Yao, and Tao Mei. 2023. Control3D: Towards Controllable Text-to-3D Generation. In *Proceedings of the 31st ACM International Conference on Multimedia*. ACM, Ottawa ON Canada, 1148–1156. doi:10.1145/3581783.3612489
- [12] Tobias Drey, Jan Gugenheimer, Julian Karlbauer, Maximilian Milo, and Enrico Rukzio. 2020. VRSketchIn: Exploring the Design Space of Pen and Tablet Interaction for 3D Sketching in Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–14. doi:10.1145/3313831.3376628
- [13] John Gero and Vladimir Kazakov. 2003. On measuring the visual complexity of 3D solid objects. *e-Activities in Design and Design Education, Europia, Paris* 1, 1 (2003), 147–156.
- [14] A A Globa, M Donn, and O A Ulchitskiy. 2016. Metrics for measuring complexity of geometric models. *Scientific visualization* 8, 5 (2016), 74–82.
- [15] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology*, Peter A. Hancock and Najmedin Meshkati (Eds.). Human Mental Workload, Vol. 52. North-Holland, Los Angeles, CA, USA, 139–183. doi:10.1016/S0166-4115(08)62386-9
- [16] Richard Helps and Clarissa Helps. 2016. Measuring Stylus and Tablet Performance for Usability in Sketching. In *Proceedings of the 5th Annual Conference on Research in Information Technology*. ACM, Boston Massachusetts USA, 19–24. doi:10.1145/2978178.2978186
- [17] Samory Houzangbe, Dimitri Masson, Sylvain Fleury, David Antonio Gómez Jáuregui, Jeremy Legardeur, Simon Richir, and Nadine Couture. 2022. Is Virtual Reality the Solution? A Comparison between 3D and 2D Creative Sketching Tools in the Early Design Process. *Frontiers in Virtual Reality* 3 (Oct. 2022), 16 pages. doi:10.3389/frvir.2022.958223
- [18] Tianyu Huang, Yihan Zeng, Zhilu Zhang, Wan Xu, Hang Xu, Songcen Xu, Rynson W. H. Lau, and Wangmeng Zuo. 2023. DreamControl: Control-Based Text-to-3D Generation with 3D Self-Prior. arXiv:2312.06439 [cs]
- [19] J. H. Israel, E. Wiese, M. Mateescu, C. Zöllner, and R. Stark. 2009. Investigating Three-Dimensional Sketching for Early Conceptual Design—Results from Expert Discussions and User Studies. *Computers & Graphics* 33, 4 (Aug. 2009), 462–473. doi:10.1016/j.cag.2009.05.005
- [20] Rubaiat Habib Kazi, Tovi Grossman, Hyunmin Cheong, Ali Hashemi, and George Fitzmaurice. 2017. DreamSketch: Early Stage 3D Design Explorations with Sketching and Generative Design. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. ACM, Québec City QC Canada, 401–414. doi:10.1145/3126594.3126662

- [21] Ben Leduc-Mills and Michael Eisenberg. 2011. The UCube: A Child-Friendly Device for Introductory Three-Dimensional Design. In *Proceedings of the 10th International Conference on Interaction Design and Children (IDC '11)*. Association for Computing Machinery, New York, NY, USA, 72–80. doi:10.1145/1999030.1999039
- [22] Serena Lee-Cultura and Michail Giannakos. 2020. Embodied Interaction and Spatial Skills: A Systematic Review of Empirical Studies. *Interacting with Computers* 32, 4 (July 2020), 331–366. doi:10.1093/iwcomp/iwaa023
- [23] Xiaoyu Li, Qi Zhang, Di Kang, Weihao Cheng, Yiming Gao, Jingbo Zhang, Zhihao Liang, Jing Liao, Yan-Pei Cao, and Ying Shan. 2024. Advances in 3D Generation: A Survey. arXiv:2401.17807 [cs] doi:10.48550/arXiv.2401.17807
- [24] Gal Metzger, Elad Richardson, Or Patashnik, Raja Giryes, and Daniel Cohen-Or. 2023. Latent-NeRF for Shape-Guided Generation of 3D Shapes and Textures. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, Vancouver, BC, Canada, 12663–12673.
- [25] Aryan Mikaeili, Or Perel, Mehdi Safaei, Daniel Cohen-Or, and Ali Mahdavi-Amiri. 2023. SKED: Sketch-guided Text-based 3D Editing. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE, Paris, France, 14561–14573. doi:10.1109/ICCV51070.2023.01343
- [26] Yeongtak Oh, Jooyoung Choi, Yongsung Kim, Minjun Park, Chaehun Shin, and Sungroh Yoon. 2024. ControlDreamer: Stylized 3D Generation with Multi-View ControlNet. arXiv:2312.01129 [cs]
- [27] Daniel Paes, Eduardo Arantes, and Javier Irizarry. 2017. Immersive Environment for Improving the Understanding of Architectural 3D Models: Comparing User Spatial Perception between Immersive and Traditional Virtual Reality Systems. *Automation in Construction* 84 (Dec. 2017), 292–303. doi:10.1016/j.autcon.2017.09.016
- [28] Jeanne Pellerin, Guillaume Caumon, Charline Julio, Pablo Mejia-Herrera, and Arnaud Botella. 2015. Elements for measuring the complexity of 3D structural models: Connectivity and geometry. *Computers & Geosciences* 76 (March 2015), 130–140. doi:10.1016/j.cageo.2015.01.002
- [29] Jessica Reichert, André R. Backes, Patrick Schubert, and Thomas Wilke. 2017. The power of 3D fractal dimensions for comparative shape and structural complexity analyses of irregularly shaped organisms. *Methods in Ecology and Evolution* 8, 12 (2017), 1650–1658. doi:10.1111/2041-210X.12829 _eprint: <https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/2041-210X.12829>
- [30] Ryan Schmidt, Azam Khan, Gord Kurtenbach, and Karan Singh. 2009. On Expert Performance in 3D Curve-Drawing Tasks. In *Proceedings of the 6th Eurographics Symposium on Sketch-Based Interfaces and Modeling (SBIM '09)*. Association for Computing Machinery, New York, NY, USA, 133–140. doi:10.1145/1572741.1572765
- [31] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. 2017. Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence* 4, 6 (2017), 103. doi:10.9781/ijimai.2017.09.001
- [32] Yu-Hsin Tung and Chun-Yen Chang. 2024. How Three-Dimensional Sketching Environments Affect Spatial Thinking: A Functional Magnetic Resonance Imaging Study of Virtual Reality. *PLOS ONE* 19, 3 (March 2024), e0294451. doi:10.1371/journal.pone.0294451
- [33] Chen Wang, Hao-Yang Peng, Ying-Tian Liu, Jiatao Gu, and Shi-Min Hu. 2025. Diffusion Models for 3D Generation: A Survey. *Computational Visual Media* 11, 1 (Feb. 2025), 1–28. doi:10.26599/CVM.2025.9450452
- [34] Jiarui Wang, Huiyu Duan, Guangtao Zhai, and Xiongkuo Min. 2025. Quality Assessment for AI Generated Images with Instruction Tuning. arXiv:2405.07346 [cs] doi:10.48550/arXiv.2405.07346
- [35] Jianfeng Xiang, Zelong Lv, Sicheng Xu, Yu Deng, Ruicheng Wang, Bowen Zhang, Dong Chen, Xin Tong, and Jiaolong Yang. 2025. Structured 3D Latents for Scalable and Versatile 3D Generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, Nashville, TN, USA, 21469–21480.
- [36] Chufeng Xiao, Wanchao Su, Jing Liao, Zhouhui Lian, Yi-Zhe Song, and Hongbo Fu. 2022. DifferSketching: How Differently Do People Sketch 3D Objects? *ACM Trans. Graph.* 41, 6 (Nov. 2022), 264:1–264:16. doi:10.1145/3550454.3555493
- [37] Pengfei Xu, Hongbo Fu, Youyi Zheng, Karan Singh, Hui Huang, and Chiew-Lan Tai. 2019. Model-Guided 3D Sketching. *IEEE Transactions on Visualization and Computer Graphics* 25, 10 (Oct. 2019), 2927–2939. doi:10.1109/TVCG.2018.2860016
- [38] Peng Xu, Timothy M. Hospedales, Qiyue Yin, Yi-Zhe Song, Tao Xiang, and Liang Wang. 2023. Deep Learning for Free-Hand Sketch: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, 1 (Jan. 2023), 285–312. doi:10.1109/TPAMI.2022.3148853
- [39] Yilin Ye, Jianing Hao, Yihan Hou, Zhan Wang, Shishi Xiao, Yuyu Luo, and Wei Zeng. 2024. Generative AI for Visualization: State of the Art and Future Directions. *Visual Informatics* 8, 2 (June 2024), 43–66. doi:10.1016/j.visinf.2024.04.003
- [40] Robert C. Zeleznik, Kenneth P. Herndon, and John F. Hughes. 2006. SKETCH: An Interface for Sketching 3D Scenes. In *ACM SIGGRAPH 2006 Courses on - SIGGRAPH '06*. ACM Press, Boston, Massachusetts, 9. doi:10.1145/1185657.1185770
- [41] Lvmin Zhang, Anyi Rao, and Maneesh Agrawala. 2023. Adding Conditional Control to Text-to-Image Diffusion Models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. IEEE, Paris, France, 3836–3847. https://openaccess.thecvf.com/content/ICCV2023/html/Zhang_Adding_Conditional_Control_to_Text-to-Image_Diffusion_Models_ICCV_2023_paper.html