# Empirical Study of Virtual Reality and Desktop Systems for Qualitative Editing of 3D Meshes: Impacts of Expertise and Context

Jennifer Cieliesz Cremer
jcremer3@ufl.edu
University of Florida
Gainesville, Florida, USA

Connor Lausch
clausch@ufl.edu
University of Florida
Gainesville, Florida, USA

Jörg Peters
jorg@cise.ufl.edu
University of Florida
Gainesville, Florida, USA

Eric Ragan
eragan@ufl.edu
University of Florida
Gainesville, Florida, USA

## ABSTRACT

For editing 3D spatial data, 3D interaction through virtual reality (VR) can be a viable alternative to 2D interfaces: research indicates that model editing in VR provides a more enjoyable experience and is fast to learn. These advantages make VR an appealing option for training new users' spatial understanding before transitioning to standard 2D tools, like Blender and Maya. But how much does model editing in VR benefit the trained user? Our experiment compares the modeling accuracy of non-modelers, casual users, and formally trained artists for objects of varying complexity in desktop and VR. For users with no prior modeling experience, the study found significant improvements in qualitative accuracy and efficiency of aesthetic edits using VR. Importantly, improvements decreased with higher user experience and varied with types of editing for different surface features. The findings suggest that adaptation of traditional desktop modeling tools to VR should be situational decisions based on specific modeling scenarios.

## CCS CONCEPTS

• **Human-centered computing** → *Empirical studies in visualization*; • **Computing methodologies** → **Virtual reality**; Mesh models.

**ACM Reference Format:**
Jennifer Cieliesz Cremer, Connor Lausch, Jörg Peters, and Eric Ragan . 2025. Empirical Study of Virtual Reality and Desktop Systems for Qualitative Editing of 3D Meshes: Impacts of Expertise and Context. In *31st ACM Symposium on Virtual Reality Software and Technology (VRST '25), November 12–14, 2025, Montreal, QC, Canada.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3756884.3765967


*VRST '25, November 12–14, 2025, Montreal, QC, Canada*

ACM ISBN 979-8-4007-2118-2/2025/11
https://doi.org/10.1145/3756884.3765967

## 1 INTRODUCTION

Three-dimensional modeling is an integral part of design, animation, and simulation, and with the increasing utility of 3D reconstruction, modeling is spreading into areas such as medicine, archaeology, and journalism. However, it is a complex process that requires extensive training to perform spatial manipulations when restricted to 2D inputs and displays; introducing challenges such as distinguishing between vertices in close proximity and estimating the depth of surface features. A long-standing paradigm to address some of this restrictive flattening uses multiple orthogonal views paired with a perspective view to provide additional depth and orientation information [32]. While effective, the integration of multiple views into one mental 3D model is cognitively taxing, and the mental demand can be tied directly to the level of familiarity with the task and method of depiction [9]. Without this familiarity, leveraging 3D modeling becomes inaccessible to novices and novel domains seeking to advance their processes with 3D representations.

One way to avoid visual and interaction flattening, and provide a more intuitive interface, are virtual reality (VR) systems, matching 3D data with spatial visual cues and 3D interactions. By providing higher visual display fidelity and supporting natural forms of physical interaction that improve spatial understanding [21, 27, 37], VR obviates the use of multiple views. By promoting familiar body-based interactions via hand and head movement tracking, VR removes the 2D input restrictions (i.e., mouse and keyboard) and lowers the interface's learning curve. This enables users to easily adjust their view, and reach out as they would in the real world, e.g. to move vertices [5]. However, the performance of VR for 3D modeling has largely been studied with participant populations primarily comprised of *trained* artists or engineers (e.g., [10, 16, 19, 20, 33]). That is, much of our understanding of the practicality of VR for 3D modeling relies on participants contrasting it with their prior experience with traditional tools. This also means that we currently lack the data to make informed designs for these new, non-traditional groups of users.

Our research expands on prior work by studying a range of user expertise and editing cases for 3D modeling. We present a controlled experiment that evaluates qualitative accuracy and speed between desktop and VR systems for novice, occasional, and experienced modelers: participants performed vertex manipulations to edit provided meshes to match point clouds of distinct *surface complexity*.

## 2 RELATED WORK

Our experiment builds on prior research, comparing VR to desktop interfaces for spatial understanding and 3D modeling, while adding emphasis on the impact of user experience. In contrast to previous research of 3D modeling, focusing on the early stage of prototyping, our study of 3D modeling focuses on a later fine-tuning stage of modeling, in particular vertex editing, and when non-traditional users are more likely to be involved.

### 2.1 3D Viewing and Editing

Prior research from the computer graphics and human-computer interaction communities have studied different interfaces for 3D viewing and editing. For example, to better understand the challenges of spatial viewing styles for desktop displays, Tory et al. [32] investigated the efficacy of different viewing styles for computer-aided design (CAD) interfaces. They found that 3D views were only appropriate for relative positioning tasks, and the 2D views' accuracy were sensitive to task specifics. Their best performing configuration was thus a combination of three flat views paired with a perspective view of the object. This configuration is the basis for the industry-standard use of "quad-view" layouts, and thus is also the configuration used in the desktop portion of our study.

There have also been attempts to resolve 2D input restrictions via input tools ranging from a wand [36] to a 3D mouse [3] to hand-tracking [34], with major outcomes including improved task time [30] and correctness [12]. While these input techniques show benefit even without the use of a stereoscopic display [23], users prefer the combination of high fidelity input *and* display [30]. This preference is likely a result of users feeling a disconnect between the spatial freedom for moving objects and still being axis-restricted for view manipulations; and one of the underlying motivators for most of the application-specific research for 3D modeling largely utilize the complete system (controllers/hands and head-tracking).

### 2.2 Virtual Reality Modeling

Based on prior research about VR's display and interaction components' effects on spatial comprehension [5, 21, 27], VR is a well-suited candidate for the spatial tasks involved in 3D modeling. Visual stereoscopy creates binocular disparity that helps clarify spatial depths and object relations such as occlusion and connectivity [21, 27]. Spatial understanding and manipulation of 3D objects further benefit from stereoscopic vision when paired with body-based inputs for view adjustment [4, 22, 26, 35]. Studies have also generated design guidelines for how to facilitate workflows in the VR space. For instance, an analysis of the interface features by Mendes et al. [24] provides a taxonomy to assist comparison of desktop and VR for 3D object manipulations.

Generating 3D models in VR comes in a variety of forms. One of the most popular is sketch-based, where users draw using ribbons. Prime examples of this method are Google's Tilt Brush [13], CASSIE [38], StripBrush [28], Gravity Sketch [11], and 3DBrushVR [40]. These methods are accompanied by equally prevalent work to convert sketch models into surfaces via methods like Lift-Off [16], RodMesh [31], SurfaceBrush [29], and ScaffoldSketch [39]. While these works are pivotal in their own right, sketch-based modeling and subsequent parameterization is in direct contrast to the rigid, topology-based mesh modeling in which this study is grounded.

There are common findings for both styles of modeling: users found these systems easy to learn [15, 31, 33], gained clearer spatial insights [10, 33], and felt more freely creativity [16, 20]. Many of these studies highlight that VR provides a richness to the modeling process. For example, Vlah et al. [33] determined that for product design, freeform and geometrically complex objects, VR was advantageous for identifying and manipulating surface control points.

However, VR should not be seen as a complete modality replacement. The intricacies of 3D modeling are not a uniform perfect fit—as demonstrated in both Mine et al.'s [25] adaptation of SketchUp into VR and Ladwig et al.'s [20] proposed system—and participants often comment on a lack of precision [16, 20, 33]. For sketching, this stems from a lack of grounding references that provide spatial scale [17]. For primitive modeling, it is either a case of attempting fine-grain motor control [10] or exact geometric placement (i.e., dimensioning or parallelism) [33], functionality that does not inherently translate to VR [20, 25].

### 2.3 Novices versus Expert Users

In addition to accuracy for 3D inspection and manipulation tasks, the study of 3D modeling can also be affected by the nature of the modeling task, the amount of freedom for 3D creation, and the participants' experience with modeling. Studies involving full start-to-finish modeling of 3D objects using primitive-geometry requires a level of fluency in the modeling process, such as those by Vlah. et al. [33]. This is also the case for the more freeform studies such as by Jackson and Keefe [16]. In the pilot study of CaveCAD by Hughes et al. [15], where experts and novices recreated Disney's Magic Castle, the intuitive design of the interface and tools could not offset the prior knowledge and experience with design principles necessary to successfully complete the given task.

This prior experience would be expected to influence any qualitative feedback, as experienced users are accustomed to the extensive set of tools of desktop software that they have tailored workflows for, creating a comfort bias. Thus, we designed our study to bypass the need for existing familiarity with the modeling process by testing the late-stage task of vertex editing on existing geometry. Moreover, this task can be equally completed by novice and expert with only a mouse and on-screen widgets.

## 3 EXPERIMENT

Our mixed-factorial user experiment compares qualitative mesh editing efficacy of VR against desktop interfaces for participants with different levels of 3D modeling experience. Participants utilized two separate hardware systems to edit surface meshes to match point clouds of three unique shapes to elicit different types of editing behaviors.

### 3.1 Goals

While the overall performance of VR for modeling has been studied, e.g., [15, 16, 20, 31, 33], they focused primarily on the ideation and prototyping tasks of 3D modeling. This requires extensive knowledge of the modeling process, putting the occasional novice participant at a disadvantage [15], while feedback from experts is biased by their prior experience in traditional interfaces. The primary goal of this study was to elucidate any differences in the benefits of VR when stratified by levels of user expertise. Consequently, we focused on the later stage of modeling, mesh alteration, which does not require the extensive tools needed for generating model topology. Focusing on editing existing models can not only inform artistic communities for design review alternatives but, also fields that do not regularly interact with CAD and are more likely to have untrained, or low-trained, users, such as a surgeon editing a liver reconstruction to include tumorous regions.

## 3.2 Experimental Design

To address the research goals, we designed an experiment to assess qualitative accuracy and interface efficiency between desktop and VR applications for different levels of modeling training. The experiment is thus a $3 \times 2 \times 3$ mixed-factorial design.

EXPERIENCE : modeling experience—None, Low, or High
DISPLAY : interface—Desktop and VR
SHAPE : trial geometry—Concavity, Protrusion, and Twist

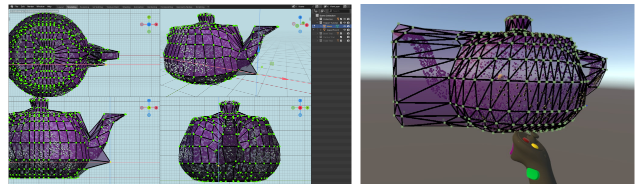
EXPERIENCE is the between-subject factor, DISPLAY and SHAPE are the within-subject factors. The starting DISPLAY was counterbalanced to mitigate any learning bias due to ordering effects.

The study **task** required participants to move *individual* vertices of a convex hull so that the faces and vertices of that mesh tightly encapsulate the outer surface of a co-located, ground-truth point cloud, see Fig. 3 . We define the *qualitative accuracy* of this task to refer to how closely the *entire* mesh adheres to the surface of the ground-truth point cloud. There was no single, perfect location that each vertex was supposed to be moved to; the participant could move each vertex to any location they deemed appropriate so that the mesh's faces and vertices adhered to as much of the point cloud as possible.

*3.2.1 Prior Modeling Experience:* To offset the complication of nonlinear skill accumulation when measuring expertise, we opted for a descriptive categorization that relied on how modeling skill was acquired in addition to the user's current usage patterns. Since the procedure would be unaffected by the EXPERIENCE level, participant designation was set after the study using self-reported values of learning experience and 3D modeling frequency from the demographics questionnaire. This allowed for indiscriminate recruitment, with some later targeted recruiting to balance the groups where random chance failed. For the purposes of this study, we classified participant experience using the following categories:

None = No prior experience modeling 3D objects.
Low = Users who were self-taught and had sporadic usage.
High = Users with experience via courses or tutorials who model 3D objects at least once a month.

While this comes with the reliability limitation that accompanies self-reported data, the desired alternative of a longitudinal study was outside our available resources to conduct.



**(a) Desktop scene** **(b) VR scene**

**Figure 1: Tutorial scene (Utah teapot challenge) for both DISPLAY types presented to the participant.**

*3.2.2 Displays Systems:* To compare DISPLAYs across different levels of EXPERIENCE interactions and interface features were simplified to the necessities: perspective viewing, zoom, rotation, and single-vertex translation. Notably, addition and removal of vertices were disabled to force all participants to work around the same constraints of the provided vertex connectivity.

*The* Desktop *condition was implemented using Blender [7] in the standard quad-view configuration.* This provides the primary perspective view along with the complimentary three orthogonal views to provide additional spatial awareness cues as recommended by Tory et al. [32] and part of industry-standard practice. Each of the four viewports could be rotated to a different perspective view or reset to an orthogonal projection, allowing for investigation of the SHAPE from multiple angles at once. Blender was locked into "Edit Mode: Vertex" with all other panels hidden, other than the "Outliner", as seen in Fig. 1a . The "Outliner" panel was provided to allow participants to independently toggle the visibilities of the convex hull and the point cloud. Panning and zooming a viewport were available using the middle mouse button and scrollwheel, and viewport rotation was performed using the right mouse button or the viewport orientation gizmo in the top right corner of each viewport. The left mouse button was used for selecting a vertex and interacting with the vertex's translation gizmo to move the selected vertex along any and all three axes if the participant desired.

*The* VR *condition used a headset and controllers to interact with the trial in a bare environment, as seen in Fig. 1b .* The headset's head-tracking was used for rotation and translation of the scene camera, where the translation accomplishes the same goals as pan and zoom on the Desktop. By providing these controls via head-tracking, it enhances the effects of stereoscopic viewing with motion parallax for depth perception by binocular disparity. Grabbing and rotating the co-located mesh-cloud pair, as well as scaling via the "stretch-and-pinch" gesture, utilized the controllers' grip buttons. These controls were provided to allow participants to comfortably situate the scene with respect to their physical posture, similar to adjusting mouse and chair locations in front of the Desktop monitor. Vertex selection and movement were performed using the left or right trigger button(s). Visibility toggling for the convex hull mesh and point cloud were mapped to the main buttons on the right controller.

*3.2.3 Trial Shapes:* The study was designed to use a small variety of shapes to cover different spatial perceptions and promote specific types of vertex manipulations, with three anticipated levels

**Figure 2: A subset of commonly used test models reviewed for feature patterns: concave spaces (Teapot and Fertility), protrusions (Spot the Cow and Bunny), and twists (Bunny and Dragon).**

of surface complexity. These SHAPEs demonstrated categories of object features as distilled from the author's review of the graphics community's commonly used test models (Fig. 2 ). A similar set of shape characteristic groupings have been used by Arora et al. [1] to test drawing projected curves on objects of various difficulties of spatial complexity.

We generalized the topological similarities and used coarse meshes for the convex hulls to maintain study sustainability. One such generalization was the exclusion of holes as a category—as any tweaking of a hole's perimeter is no different from the Concavity scenario. The use of the point cloud format for the ground-truth shape was to force participants to work with an imaginary surface, preventing possible surface details of the models from influencing the participants' vertex placements. The decision to use convex hulls was to simulate exaggerated versions of common reconstruction algorithm errors, as this is a main use-case doctors and museum curators would be interacting with the 3D modeling discipline.

*The* Concavity *challenge was designed to be the easiest of the three* SHAPE*s, with a single, uniform area of interest.* It was designed to address filling in concave spaces (e.g., teapot interior or the inset of an eye socket), and also provided a crisp-edge scenario, where face normals have a sudden, sharp change in direction to form the rim. The Concavity is also the most symmetric of the three SHAPEs, with two planes of symmetry.

*The* Protrusion *challenge required vertices to be redistributed to capture surface details.* This scenario is common in surface reconstruction algorithms, which prioritize adherence to the outer-most edges and will stretch a single face to span the adjacent areas. These connecting spaces include smooth changes in surface normals (e.g., bunny), linear details (e.g., teapot spout, bunny ears), and connection space between other features (e.g., space between fingers or between the ears and back of bunny). We decided to shrink-wrap a quadrant of Spot's head, instead of using a trebol, so that the scale of smooth regions and the sharpness of the protrusions varied. This made for a medium-difficulty object, as it had a single degree of symmetry, two main areas of correction, and provides multiple locations a participant can pull vertices.

*The* Twist *challenge was the most spatially complex of the three* SHAPE*s, involving concavity and vertex redistribution.* The shape represents the modeling of thin features (e.g., fingers) and multi-planar changes in orientation (e.g., bunny, dragon), with no planar symmetry or unique landmark features with which to orient oneself. It is also an example of an extraordinary point, another common occurrence for surface reconstruction, where a single vertex has an unusually high or low valence compared to the surrounding area.

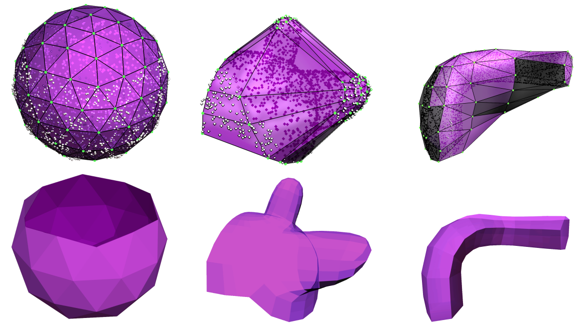



**Figure 3: SHAPEs. Top row: starting state (convex hull); Bottom row: ground-truth (GT) point clouds as meshes to better illustrate the target. Participants freely move the green vertices of the purple mesh to encase the gray point cloud of the target as tightly as possible.**

## 3.3 Hypotheses

In order to isolate the impact of DISPLAY and SHAPE for different EXPERIENCE, we tested along three hypotheses:
Mesh editing in VR...

- **H1** ...yields greater qualitative accuracy than Desktop.
- **H2** ...has greater efficiency than Desktop.
- **H3** ...is, for experienced users, on par with Desktop in accuracy and efficiency.

Notably, **H1** refers specifically to qualitative accuracy, which permits multiple correct solutions, as a distinction from quantitative precision, where exact localization in space is the goal. **H3** posits that the benefits predicted in **H1** and **H2** will be less pronounced in the experienced groups (Low and High) due to a well-established set of skills and training mitigating the challenges of model editing on the desktop.

## 3.4 Measures

In order to test **H1** and **H2**, three quantitative measures were recorded for the trials:

- *Error* : the unweighted average between the normalized *face distance deviation*, (1) and the overall *shape deviation*, (2);
- *Time* : Time taken to complete the trial;
- *Actions* : Number of vertex movements performed.

We use a combination of face distance deviation and shape deviation to measure error for each SHAPE × DISPLAY trial; thus, each participant generated six different error values. This approach to error, instead of a process that incorporates measures of "good geometry" (e.g., no intersecting or equally sized faces) was used because a third of the target population needed to have no prior modeling experience and it was decided that explaining these principles would unnecessarily burden these participants. The usage of shape, as well as distance, to calculate error, allows for the distinction between a result that had a tight fit but many jagged faces, and a result that closely followed the intended topology but was further away from the point cloud.

The surface distances used for the *face distance deviation* part of the error metric were computed using the Hausdorff Distance filter in MeshLab [6], and normalized to the distance deviation of the trial's starting mesh configuration. This value was then converted into a percentage. A participant who made no improvements to an SHAPE would have a *face distance deviation* value of 100 for that trial, while a participant who achieved a perfect match of the SHAPE to the GT point cloud would have a *face distance deviation* of 0 for that trial. The normalization indicated for when no tightening occurred, and allowed for the distances to have equal magnitudes despite the SHAPEs occupying different sized bounding boxes.

The *face distance deviation* uses

- $S$ starting mesh configuration for a given trial ($n$ faces),
- $P$ participant's edited configuration,
- $T$ fine-grain target mesh (GT point cloud).

We denote the root-mean square of the Unidirectional Hausdorff Distance from the mesh to the GT as $RMS_{HD}(X \rightarrow T)$, using the Euclidean distance between centroids, where $X$ takes the place of either $S$ or $P$ and $S = P \Leftrightarrow \forall f(f \in S \Leftrightarrow f \in P)$. Then the *face distance deviation* component of the error metric is represented as

$$\Delta Dist(P \rightarrow T; S) = \frac{RMS_{HD}(P \rightarrow T)}{RMS_{HD}(S \rightarrow T)} \times 100. \quad (1)$$

The *shape deviation* component of the error metric sums the surface normal deviations between the trial mesh and the GT point cloud's mesh equivalent , multiplied by the surface area percentage of that face for the trial mesh. This prevented extreme local scenarios, such as a face with 1% coverage and a large oscillation, from unfairly biasing the overall fit that was achieved.

More formally, for each face $f_i \in P$, let:

- $\vec{\boldsymbol{n}}_i^P$ : unit normal vector of face $f_i$
- $\boldsymbol{A}_i^P$ : area of $f_i$
- $t_i^*$ : the face in $T$ whose centroid is nearest to the centroid of $f_i$
- $\vec{\boldsymbol{n}}_i^T$ : unit normal vector of $t_i^*$

The total surface area of $P$ is defined as $\boldsymbol{A}_{total}^P = \sum_{j=1}^{n} \boldsymbol{A}_j^P$ and the angle between the unit normal vectors is defined as $\theta_i = arccos(\vec{\boldsymbol{n}}_i^P \cdot \vec{\boldsymbol{n}}_i^T)$. Thus the shape deviation for a given trial is

$$\Delta Shape(P \rightarrow T) = \sum_{i=1}^{n} (\frac{\theta_i}{\pi} \cdot \frac{\boldsymbol{A}_i^P}{\boldsymbol{A}_{total}^P}) \times 100. \quad (2)$$

The angle between the normal vectors is divided by $\pi$ to convert from radians to $[0, 1]$ as for the surface area ratio.

The error for any of the six trials for a given participant is then defined as

$$Error = \frac{\Delta Dist(P \rightarrow T; S) + \Delta Shape(P \rightarrow T)}{2}. \quad (3)$$

Pilot studies of the procedure revealed that task fatigue was a nontrivial concern, and something novice users might not easily identify. Thus, the task time was constrained within a 7-minute window to protect participants from this phenomenon and maintain a sustainable study duration. To evaluate if increased modeling experience resulted in higher consistency in performance across interfaces, as **H3** claims, all three measures were compared across the EXPERIENCE groups.

To check if the quantitative results translate to user experience, participants completed a post-questionnaire that collected the perceived demands with a separate section for each SHAPE. The participant was asked to rate each DISPLAY's mental load, perceived performance, effort, and frustration for that section's SHAPE on a 10-point scale adapted from the NASA-TLX [14]. All scales used a negative-positive correlation except for performance, which had an inverse relationship with the scale. These values were collected at the end to ensure self-relativity of the responses within the participant for each SHAPE. The post-questionnaire also asked for preference and confidence ratings for each DISPLAY.

## 3.5 Procedure

This study was approved by our Institutional Review Board as an exempt study. After informed consent was obtained, participants were given a brief description of the study's goal and their task for each trial—to redistribute and move the vertices of a loosely-fitted mesh to tightly enclose the given point cloud. We relied on the casual language of "fit like a second skin" to carry this intention, as well as providing a demonstration of what the mesh looked like when above, on, and below the point cloud. Since one of the target populations needed to have no prior modeling experience, no explicit instructions of what constituted "good mesh quality" were provided. So while the Protrusion was designed to encourage vertex redistribution, the participants were not measured on if that distribution was evenly spaced. Participants then completed a demographics questionnaire including experience with 3D modeling software and virtual reality. They were shown rotating previews of each ground truth object used in the trials to assist with task clarification and equalize SHAPE familiarity between interfaces.

At the start of each interface, the correspondence between the point cloud and mesh, and the available tools, were demonstrated to participants using an instructional video. During this time, the experimenter highlighted the pre-fit portions of the mesh as an example of the goal, as well as what the mesh looked like when a vertex was moved far away, beneath the point cloud, and intersecting other faces. The participants then used the same scene to practice using each of the controls, as guided by the experimenter, and fit the mesh to the underside of the spout of the Utah Teapot, as seen in Fig. 1 . The instructions included identifying the provided reference diagram of the controls, starting and stopping the timer, toggling mesh and cloud visibility, vertex manipulations, and camera/view manipulations. Additional freeform practice was permitted for the remainder of the 7-minutes of training.

Both the practice scene and each of the six trials had a 7-minute restriction to reduce the chances of task fatigue without inducing undue temporal stress for the participant. It also allowed for qualitative accuracy measures to be more clearly skill-dependent across the EXPERIENCE groups; as, given infinite time, a None user could achieve the same accuracy as a High user. The same time limit was used for the training scene to allow the participant to adjust to the feeling of 7 minutes for this task. To reduce mistakes from rushing behavior, as well as discourage undue time usage from over-prioritization of accuracy, participants were reminded at the start of trials that they were not expected to finish in the 7 minutes, nor were they required to use the full time if they decided they had finished the task early. Upon each trial completion the experimenter saved the action logs and final mesh, giving participants a break before starting the next trial. After completing all six trials, the participant filled out a final questionnaire asking about their perceptions of their experience for each SHAPE.

## 3.6 Interfaces and Systems

The computer used had 32gb of RAM, an NVIDIA GeForce RTX2080 SUPER GPU, and an Intel(R) Core(TM) i7-6700k CPU @ 4.00Gz. The Desktop trials were performed at a two-monitor workstation with a 3-button mouse and no keyboard. The primary monitor contained the preset Blender window, while the secondary monitor, situated to the left, contained the Windows Timer widget floating on top of a reference diagram of the controls. The VR trials used a tethered Meta Quest 2 headset and both controllers; the application was developed using Unity 2022.3.5f1 with Universal Render Pipeline graphics and OpenXR. Participants were seated in an ergonomic office chair at least 3 feet from all obstructions with the headset cable suspended to prevent entanglement. Participants with corrected vision maintained their glasses while in the headset. The headset's built-in head-tracking was used to control the scene camera and the inside-out tracking was used for controller position. Models of the controllers with color-coded buttons were used for the virtual hands. A tool guide with matching colors was attached to the left controller. The timer was centered in the lower third of the view, preventing task occlusion while allowing constant monitoring by the participant. Visibility toggling for the controls diagram, as well as starting and stopping the timer, were mapped to the main buttons on the left controller. The exact control diagrams provided, scene settings, SHAPE material specifics, and workstation set-up can be found in the Supplemental Material for reproducibility. Experimenters were able to observe progress and the trial timer on the desk monitor.

The SHAPE point clouds were generated using geometry nodes to replace the mesh faces with randomly placed spheres and the start meshes for the Protrusion and Twist were generated using Blender's built-in shrink-wrap deformation modifier [8]. As we did not want the Concavity to necessitate vertex redistribution, we used an unaltered icosphere, instead of a shrink-wrapping—which pushed all of the vertices-of-interest out to the rim of a flattened top. The convex hulls contained 92, 101, and 98 vertices, as ordered in Section 3.2.3.

## 3.7 Participants

Of the 61 participants who completed the study, 4 were excluded from analysis based on observation of extreme outlier behavior during the experiment that suggested a lack of effort or understanding of the task. All four excluded participants were from the Desktop starting condition: 3 from None and 1 from High. The remaining 57 participants self-reported as 20 female, 35 male, and 2 non-binary. The majority of participants were graduate or undergraduate students with computer science or digital arts degrees. Participants were recruited through the computer science department and poster advertising. A total of 3 participants had never used VR before, 1 in None and 2 in Low. All other participants had a semi-annual interaction with VR, with 4 High using VR on a weekly basis. All but 7 None participants had at least experienced 3D data visualizations for fun. When reporting experience with 3D modeling, 21 participants were in the category None, 21 were Low, and 15 were High. Of those participants, 35 had prior experience with Blender or Maya.

# 4 RESULTS

To assess the qualitative accuracy (**H1**) and efficiency (**H2**) of VR for editing of 3D meshes, mesh error, task time, and vertex movement actions were analyzed for VR and Desktop display systems across three SHAPEs. These results were then compared between the three EXPERIENCE groups to assess if a user's prior modeling experience would diminish the potential benefits from using VR (**H3**). Additionally, a correlation analysis was performed to compare cognitive demands and the impacts on performance. Due to the skewed distributions of the data, an aligned rank transform was applied using ARTool [18] before running a Type III, mixed-effect 4-way ANOVA for each of the quantitative measures. Summaries of the main factors and significant interaction tests can be seen in Table 1.

## 4.1 Modeling Errors

By using the compound error metric defined in (3) we could directly compare the influence of EXPERIENCE and SHAPE on the efficacy of the DISPLAYs. A lower error for a particular SHAPE in VR would indicate that there was unique gain from the added fidelity VR provides. Similarly, a comparatively high *difference* of error between the DISPLAYs for that particular SHAPE for a given EXPERIENCE would indicate that the population has a greater benefit from using VR.

From the results in the Mesh Error column in Table 1, DISPLAY *ordering* had no substantial effect on error, meaning any impacts of DISPLAY on error were not a result of learning. Similarly, the significant interaction between *starting condition* and the DISPLAY was directly correlated to DISPLAY comparison rather than different starting conditions for that DISPLAY. DISPLAY had a positive effect on all SHAPEs, reducing the error results in VR and supporting **H1**—VR will yield higher qualitative accuracy. We visualize the full breakdown of the error results in Fig. 4 . Despite no statistically significant impact of EXPERIENCE on error—possibly from the large effect size of DISPLAY and SHAPE—there is a distinct trend to support **H3**—VR has less influence on the performance of experienced users. This is seen more clearly in Fig. 5 , which shows the per-participant *difference* in error between Desktop and VR; each EXPERIENCE group has progressively tighter and lower change in error as their modeling experience increases. The exception is the Twist trials, where VR had a strong improvement from all experience groups to be discussed later.

## 4.2 Time and Actions

The combination of time and vertex actions describe the efficiency the system has to enable the skill of the user. The initial investigation of the task times yielded strong significant main effects for DISPLAY (VR is faster than Desktop) and SHAPE, where comparisons against the Concavity trials were the main contributor, as shown in the second column of Table 1. The interaction between *starting condition* and DISPLAY, especially when further separated

**Table 1: Type-III 4-way ANOVA results for each of the three quantitative measures showing each factor and significant interaction effects. Effect size was calculated using partial Eta-squared and Cohen's d.**

| | Mesh Error | | | Task Time | | | Vertex Move. Actions | | |
|---|---|---|---|---|---|---|---|---|---|
| ***Factor*** | $F$ | $p$ | $\eta_p^2$ | $F$ | $p$ | $\eta_p^2$ | $F$ | $p$ | $\eta_p^2$ |
| *Starting* DISPLAY $F(1, 51)$ | 0.782 | 0.381 | 0.02 | 2.601 | 0.113 | 0.05 | 1.328 | 0.254 | 0.03 |
| EXPERIENCE $F(2, 51)$ | 1.493 | 0.234 | 0.06 | 1.317 | 0.277 | 0.05 | 4.703 None < High | 0.013* 0.012* | 0.16 d = 0.41 |
| DISPLAY $F(1, 51)$ | 119.629 VR < Desktop | < 0.001* | 0.32 | 138.53 VR < Desktop | < 0.001* | 0.35 | 0.137 | 0.711 | < 0.01 |
| SHAPE $F(2, 255)$ | 31.196 > Twist | < 0.001* < 0.001* | 0.20 d > 0.90 | 52.14 Concavity < | < 0.001* < 0.001* | 0.29 d < -1.10 | 84.419 Concavity < | < 0.001* < 0.001* | 0.40 d < -1.40 |
| *Start* × DISPLAY $F(1, 255)$ | 18.988 | < 0.001* | 0.07 | 19.871 | < 0.001* | 0.07 | — | — | — |
| EXPER. × DISPLAY $F(2, 255)$ | — | — | — | 39.557 | < 0.001* | 0.05 | — | — | — |
| DISPLAY × SHAPE $F(2, 255)$ | 27.055 | < 0.001* | 0.18 | 39.557 | < 0.001* | 0.24 | 30.401 | < 0.001* | 0.19 |
| *Start* × EXPER. × DISP. $F(2, 255)$ | — | — | — | 3.364 | 0.036* | 0.03 | — | — | — |
| EXPER. × DISP. × SHAPE $F(4, 255)$ | 1.418 | 0.228 | 0.02 | 2.454 | 0.046* | 0.04 | 2.553 | 0.039* | 0.04 |

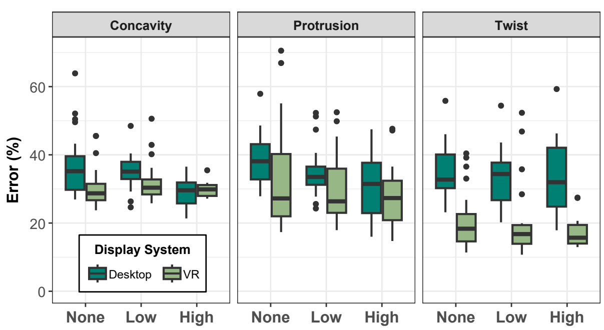



**Figure 4: Error percentage for both DISPLAYs per SHAPE across all EXPERIENCE groups using 0.95 confidence interval.**

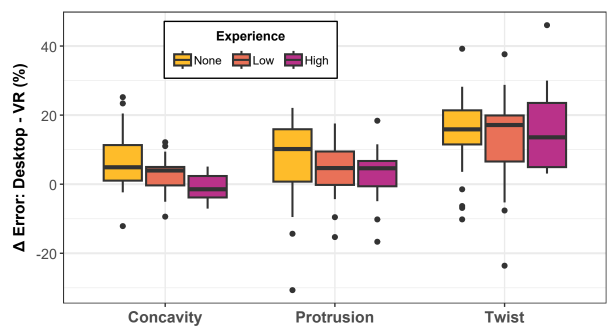



**Figure 5: The difference in error percentage between Desktop and VR for each EXPERIENCE across SHAPE, using 0.95 confidence interval. Improvement reduces as EXPERIENCE increases except for Twist. The inverse relationship observed in the Twist is likely a result of the shape's multi-planar nature.**

by EXPERIENCE, shows that participants starting in Desktop had faster times for the VR trials. However, Fig. 6 shows that this interaction does not hold for the VR-condition's Desktop trials. In Fig. 7 we see a trend for VR being the faster of the two DISPLAY, regardless of the *starting condition*, and is pronounced in participants with higher experience. However, for most of the trial conditions (both including and excluding the starting condition), the majority of the data points were at the 7-minute limit, preventing conclusions about exactly how much faster VR actually is.

For the vertex movement actions, the strongest factor was SHAPE. However, since the SHAPEs were fundamentally different, this obscures the statistical findings from the third column of Table 1 and thus results are viewed per SHAPE in Fig. 8 for contextualization. EXPERIENCE was the other main factor to have a reportedly statistically significant impact, showing an increase in actions taken in the more experienced groups. The exception is the Twist, which had an overall decrease in actions made in VR, but maintained the trend of higher actions for greater EXPERIENCE and will be discussed later. When isolating for DISPLAY, although it does not report as statistically significant as a main effect, there is a trend for a higher number of actions in VR.

## 4.3 Self-reported Effort and Confidence

To assess if the quantitative results translated into perceivable changes in the user's experience, an adaptation of the NASA-TLX was used to collect participants' mental demand, performance, effort, and frustration for each trial. All four metrics were found to have statistically significant effects from DISPLAY and SHAPE ($p < 0.001$) using the Aligned Ranking Transform and Type-III

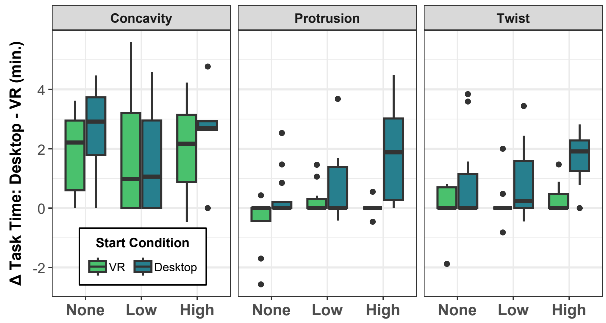



**Figure 6: The difference in task time between Desktop and VR by *start condition* for each EXPERIENCE and SHAPE, using 0.95 confidence interval. The Desktop *starting condition* consistently shows faster VR times, especially for higher levels of experience.**

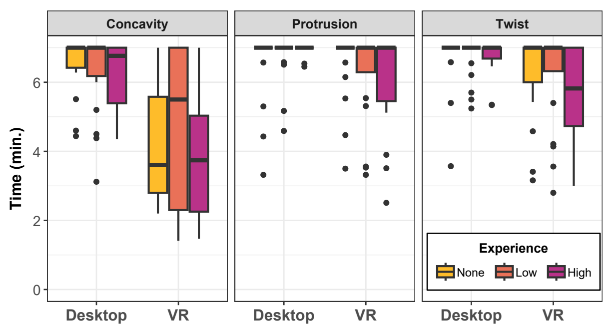



**Figure 7: Trial time for each EXPERIENCE across DISPLAY per SHAPE, using 0.95 confidence interval. Most participants used the full allotment of 7 minutes but, there is a trend for participants to complete trials faster in VR.**

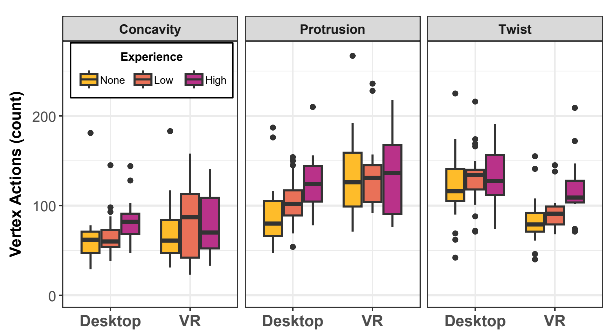



**Figure 8: Total vertex movements for each EXPERIENCE across DISPLAY per SHAPE, using 0.95 confidence interval. A higher action count in VR, despite the same time limit as Desktop, indicates action efficiency.**

4-way ANOVA. Performance ratings were positively influenced by using VR which decreased mental demand, effort, and frustration. They also had significant ($p < 0.001$) correlations to the error values using the Spearman correlation test and supporting the alternative hypotheses—effort, mental demand, and frustration were positively correlated ($\rho = 0.238, 0.252, 0.241$, respectively) and performance was negatively correlated ($\rho = -0.46$). Additionally, participants were asked to rate their preference and confidence for each of the DISPLAYs, with VR scoring significantly higher across all EXPERIENCE.

# 5 DISCUSSION

We summarize the study's findings and discuss implications for applications.

## 5.1 Summary of Quantitative Findings

The study's main findings show that VR outperforms Desktop in editing via qualitative accuracy, where VR trials produced lower error, supporting **H1** and aligning with prior research. Thus, in cases where a mesh already exists, VR is likely the better system to use for allowing subjective spatial adjustments. This is relevant for user populations, such as museum curators or radiologists, with infrequent interaction with computer modeling. For experienced modelers (High) and meshes with distinct features (Concavity and Protrusion) this improvement is minimal, with most High achieving less than 5% improvement in performance, supporting **H3** and making Desktop the more favorable modality in these cases. In contrast, for spatially ambiguous forms (i.e., Twist), all EXPERIENCE groups had a significant improvement when using VR, with High having the greatest improvement. In this instance, the study failed to support hypothesis **H3**.

The change in results between the Twist and other SHAPEs suggests that geometric features can significantly influence the impact of VR as a modeling modality. This supports Ladwig et al.'s [20] and Mine et al.'s [25] recommendation that VR not be treated as a uniform replacement for 3D modeling, but rather as an enhancement for parts of the process. One possible explanation is that shapes with ambiguous orientations gain greater benefit from the natural body-based movements and depth cues that VR provides due to the higher probability of complex visual inspection. This is supported by prior studies showing benefits of physical movement and stereoscopy for spatial judgments of small details [27] and structures with high levels of visual complexity [2]. In contrast, the planar nature of the Concavity and Protrusion did not require intricate view navigation and likely benefited from Desktop's discrete input handling, especially if multi-vertex selection had been enabled. Given that isolating impacts of different surface features was not an explicit focus of this study and the variety of these features was limited to a single sample of three broad categories, our findings concerning behaviors dependent on shape type are restricted, prompting us to recommend further investigation into when and for whom VR benefits 3D modeling.

## 5.2 Implications for Modeling Efficiency

By analyzing the quantitative results holistically—editing time and model error decreases while action frequency increases in VR—the data supports **H2**. This increased efficiency in VR allows for more work to be completed in the same amount of time, resulting in reduced overall error. The results from the Twist conditions suggests that this efficiency stems from the increased spatial understanding that VR provides. The decrease in actions alongside the decrease in error leads us to suspect that less time was used over-correcting prior edits, as observed in the Desktop trial, and instead chaining edits together, as observed for the other two SHAPE conditions.

However, this data is specific to overall vertex movements and a constrained time limit. Thus, this study can draw conjectures regarding the total efficiency of each DISPLAYs, but not an exact metric for when and how the efficiency interacts with the error produced. Inclusion of annotated action logs and progressive error measurements in future studies could further inform us of what stage and scope of the modeling process benefits from VR. This data will guide software development for VR to focus on key areas of the workflow that inherently benefit from the high visual and interaction fidelity VR offers, allowing for the platform to grow in tandem with new CAD features and changes in training paradigms.

## 5.3 Balancing Design for Levels of Experience

A governing component for this study's design was to obtain a clear comparison of how modeling EXPERIENCE impacts utility of the two DISPLAY modalities in isolation—irrespective of advanced software tools that would obscure the findings of these two variables. Prior research has focused primarily on multi-stage tasks, disadvantaging novice participants unfamiliar with the CAD process, and biasing experienced users to compare workflow mechanisms [33]. Thus, to engage and evaluate novices and experts equally, this study's task had a narrow scope, specifically targeting one of the quintessential activity in computer modeling—individual vertex manipulation—which we could distill as a single action type. The task used existing models, removing any need for understanding topological construction, akin to the scenario of editing volumetric reconstructions.

In addition to proper task scope, the metric used for evaluation also needed to be robust and appropriate. Prior studies have demonstrated that for fine-grain dimensional precision Desktop-based numerical input and restricted degrees-of-freedom outperforms VR's spatial placement [25]. This meant the measure for error could not rely on exactness of coordinate placement. Nor could it use traditional measures of "good mesh quality", such as even vertex distribution, due to the wide skill-range of intended participants. Thus, to equally score novices and experts, our error measure included total shape adherence in addition to traditional distance. While the minimalist toolset did allow for a one-to-one comparison of what novices and trained users were able to accomplish, it reduces the study's ability to investigate exactly how the more advanced CAD toolkits compare to the spatial benefits of VR.

## 5.4 VR for Spatial Training

The reduction of error and faster times in the VR trials also supports the reports that VR is easy to adjust to and quick to learn [10, 15, 31]. However, this alone is not enough to support the idea of using VR as a training platform for spatial and surface-structure understanding—as was postulated by some participants in Vlah et al.'s product design study [33]. If just the learnability of VR was sufficient, we would have expected to see an improvement for the None and Low groups that started in VR when compared to their counterparts that started on Desktop. Instead, we observed that *all* participants in the Desktop *starting condition* were more likely to take less time in the later VR session for that same SHAPE. Additionally, there was no meaningful impact from either starting condition on a participants error results. This suggests that participants starting on Desktop had to work harder to understand the SHAPEs on Desktop and thus were able to save time getting familiar with them in VR. Conversely, the initial exploratory phases by the VR starting condition did not transfer sufficiently to the Desktop trials to make a notable difference. These behaviors add formative support to the collected ratings for mental demand and effort, which were reported to be lower for VR. This contrasts with the results from Dadi et al. [9]—where it was the amount of experience a user had with the format, rather than the format itself, that affected the mental workload of different technical diagram formats.

These observations also indicate that in scenarios where the geometry already exists or is familiar to the user, VR *after* Desktopis more beneficial. In scenarios such as corrections to geometry from reconstruction processes for digitizing cultural artifacts or surgical simulation, then a full transition to a VR interface may be warranted. This would empower experts to have direct control over final outputs without needing to consult specialized technicians. For existing workflows that rely on the complex efficiency afforded by cascading tools, starting in Desktop is the more appropriate, and VR is better suited for finalization, such as design reviews.

# 6 CONCLUSION

With the rise of commercial VR headsets there has been a surge in development of 3D modeling applications to capitalize on the increase in input fidelity. Primarily based on data from experienced users, evaluations of 3D modeling in VR have found that it is easier to be creative but also that these systems may lack the precision necessary for proper start-to-finish modeling [16, 20, 33]. This study sought to clarify whether, if initial modeling is removed, there is a place for VR in the 3D modeling pipeline that does not rely on a full platform replacement. The study focused on comparing different training levels to see whether concerns regarding hardware-enabled precision was specific to the more experienced users. By analyzing error as a metric of surface distance and shape deviation, the analysis found that users had a significant improvement in error for modeling qualitative accuracy of protrusions and space curves when presented with a VR interface compared to using more traditional desktop software. For experienced users, these improvements diminished for objects with distinct orientations. The study found that VR was faster and more efficient for users of all experience levels, especially for spatially ambiguous objects. This increase in speed was most prevalent in the group that started with the desktop trials and then moved to VR, suggesting that some of the benefit comes from VR's body-based movements. Coupled with the findings that VR reduced frustration and cognitive demand, we recommend the use of VR for subjective edits and non-modeler communities unfamiliar with CAD-like software. These results encourage further investigation on the role user experience and surface features have on when to use VR for 3D modeling.

# REFERENCES


[1] Rahul Arora and Karan Singh. 2021. Mid-Air Drawing of Curves on 3D Surfaces in Virtual Reality. *ACM Transactions on Graphics (TOG)* 40, 3 (2021), 1–17. https://doi.org/10.1145/3459090

[2] Felipe Bacim, Eric D Ragan, Siroberto Scerbo, Nicholas F Polys, Mehdi Setareh, and Brett D Jones. 2013. The effects of display fidelity, visual complexity, and task scope on spatial understanding of 3D graphs.. In *Graphics Interface*, Vol. 2. 25–32.

[3] Amy Banic, Evan Suma, Frank Steinicke, Wolfgang Stuerzlinger, Vanessa Vuibert, Wolfgang Stuerzlinger, and Jeremy R Cooperstock. 2015. Evaluation of Docking Task Performance Using Mid-air Interaction Techniques. *Proceedings of the 3rd ACM Symposium on Spatial User Interaction* (2015), 44–52. https://doi.org/10.1145/2788940.2788950

[4] Woodrow Barfield, Claudia Hendrix, and Karl Bystrom. 1997. Visualizing the structure of virtual objects using head tracked stereoscopic displays. In *Proceedings of IEEE 1997 Annual International Symposium on Virtual Reality*. IEEE, 114–120.

[5] Doug A. Bowman, Ryan P. McMahan, and Eric D. Ragan. 2012. Questioning naturalism in 3D user interfaces. *Commun. ACM* 55, 9 (2012), 78–88. https://doi.org/10.1145/2330667.2330687

[6] Paolo Cignoni, Claudio Rocchini, and Roberto Scopigno. 1998. Metro: measuring error on simplified surfaces. In *Computer Graphics Forum*, Vol. 17. Blackwell Publishers, 167–174.

[7] Blender Online Community. 2024. *Blender - a 3D modelling and rendering package*. Blender Foundation, Stichting Blender Foundation, Amsterdam. http://www.blender.org

[8] Blender Online Community. 2024. *Blender - a 3D modelling and rendering package*. Blender Foundation, Stichting Blender Foundation, Amsterdam. https://docs.blender.org/manual/en/latest/grease_pencil/modifiers/deform/shrinkwrap.html

[9] Gabriel B. Dadi, Paul M. Goodrum, Timothy R. B. Taylor, and C. Melody Carswell. 2014. Cognitive Workload Demands Using 2D and 3D Spatial Engineering Information Formats. *Journal of Construction Engineering and Management* 140, 5 (5 2014), 04014001. https://doi.org/10.1061/(asce)co.1943-7862.0000827

[10] Sander Van Goethem, Regan Watts, Arno Dethoor, Rik Van Boxem, Kaz van Zegveld, Jouke Verlinden, and Stijn Verwulgen. 2020. The Use of Immersive Technologies for Concept Design. 698–704. https://doi.org/10.1007/978-3-030-51828-8_92

[11] Gravity Sketch Ltd. [n. d.]. *Gravity Sketch*. https://www.gravitysketch.com/ A virtual reality 3D design and collaboration tool.

[12] Kenny Gruchalla. 2004. Immersive Well-Path Editing: Investigating the Added Value of Immersion. *IEEE Virtual Reality 2004* (2004), 157–164. https://doi.org/10.1109/vr.2004.1310069

[13] Patrick Hackett and Drew Skillman. 2015. *Tilt Brush*.

[14] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, Peter A. Hancock and Najmedin Meshkati (Eds.). Advances in Psychology, Vol. 52. North-Holland, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

[15] Cathleen E. Hughes, Lelin Zhang, Jürgen P. Schulze, Eve Edelstein, and Eduardo Macagno. 2013. CaveCAD: Architectural Design in the CAVE. *2013 IEEE Symposium on 3D User Interfaces (3DUI)* (2013), 193–194. https://doi.org/10.1109/3dui.2013.6550244

[16] Bret Jackson and Daniel F. Keefe. 2016. Lift-Off: Using Reference Imagery and Freehand Sketching to Create 3D Models in VR. *IEEE Transactions on Visualization and Computer Graphics* 22, 4 (4 2016), 1442–1451. https://doi.org/10.1109/tvcg.2016.2518099

[17] J Joundi, Y Christiaens, J Saldien, P Conradie, and L De Marez. 2020. AN EXPLORATIVE STUDY TOWARDS USING VR SKETCHING AS A TOOL FOR IDEATION AND PROTOTYPING IN PRODUCT DESIGN. *Proceedings of the Design Society: DESIGN Conference* 1 (2020), 225–234. https://doi.org/10.1017/dsd.2020.61

[18] Matthew Kay, Lisa A. Elkin, James J. Higgins, and Jacob O. Wobbrock. 2021. *ARTool: Aligned Rank Transform for Nonparametric Factorial ANOVAs*. https://doi.org/10.5281/zenodo.594511

[19] Daniel F. Keefe, Robert C. Zeleznik, and David H. Laidlaw. 2007. Drawing on Air: Input Techniques for Controlled 3D Line Illustration. *IEEE Transactions on Visualization and Computer Graphics* 13, 5 (2007), 1067–1081. https://doi.org/10.1109/tvcg.2007.1060

[20] Philipp Ladwig, Jens Herder, and Christian Geiger. [n. d.]. Towards Precise, Fast and Comfortable Immersive Polygon Mesh Modelling: Capitalising the Results of Past Research and Analysing the Needs of Professionals. In *ICAT-EGVE 2017 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*. The Eurographics Association. https://doi.org/10.2312/egve.20171360

[21] Bireswar Laha, Doug A. Bowman, and John J. Socha. 2014. Effects of VR System Fidelity on Analyzing Isosurface Visualization of Volume Datasets. *IEEE Transactions on Visualization and Computer Graphics* 20, 4 (2014), 513–522. https://doi.org/10.1109/tvcg.2014.20

[22] Bireswar Laha, Kriti Sensharma, James D Schiffbauer, and Doug A Bowman. 2012. Effects of immersion on visual analysis of volume data. *IEEE transactions on visualization and computer graphics* 18, 4 (2012), 597–606.

[23] Ryan P. McMahan, Doug Gorton, Joe Gresock, Will McConnell, and Doug A. Bowman. 2006. Separating the effects of level of immersion and 3D interaction techniques. ACM Press, New York, New York, USA, 108–111. https://doi.org/10.1145/1180495.1180518

[24] D. Mendes, F. M. Caputo, A. Giachetti, A. Ferreira, and J. Jorge. 2019. A Survey on 3D Virtual Object Manipulation: From the Desktop to Immersive Virtual Environments. *Computer Graphics Forum* 38, 1 (2019), 21–45. https://doi.org/10.1111/cgf.13390

[25] Mark Mine, Arun Yoganandan, and Dane Coffey. 2015. Principles, interactions and devices for real-world immersive modeling. *Computers & Graphics* 48 (2015), 84–98. https://doi.org/10.1016/j.cag.2015.02.004

[26] Mark R. Mine, Frederick P. Brooks, and Carlo H. Sequin. 1997. Moving objects in space. *Proceedings of the 24th annual conference on Computer graphics and interactive techniques - SIGGRAPH '97* (1997), 19–26. https://doi.org/10.1145/258734.258747

[27] Eric D. Ragan, Regis Kopper, Philip Schuchardt, and Doug A. Bowman. 2013. Studying the Effects of Stereo, Head Tracking, and Field of Regard on a Small-Scale Spatial Judgment Task. *IEEE Transactions on Visualization and Computer Graphics* 19, 5 (2013), 886–896. https://doi.org/10.1109/tvcg.2012.163

[28] Enrique Rosales, Jafet Rodriguez, Chrystiano Araújo, Nicholas Vining, Dongwook Yoon, and Alla Sheffer. 2021. StripBrush: A Constraint-Relaxed 3D Brush Reduces Physical Effort and Enhances the Quality of Spatial Drawing. arXiv:2109.03845 [cs.HC]

[29] Enrique Rosales, Jafet Rodriguez, and ALLA SHEFFER. 2019. SurfaceBrush. *ACM Transactions on Graphics (TOG)* 38, 4 (2019), 1–15. https://doi.org/10.1145/3306346.3322970 arXiv:1904.12297

[30] Udo Schultheis, Jason Jerald, Fernando Toledo, Arun Yoganandan, and Paul Mlyniec. 2012. Comparison of a Two-Handed Interface to a Wand Interface and a Mouse Interface for Fundamental 3D Tasks. *2012 IEEE Symposium on 3D User Interfaces (3DUI)* 1 (2012), 117–124. https://doi.org/10.1109/3dui.2012.6184195

[31] Sorkine-Hornung, Floor Verhoeven and Olga. 2019. RodMesh: Two-handed 3D Surface Modeling in VR. In *Proceedings of the Symposium on Vision, Modeling and Visualization (VMV)*. Eurographics Association. https://doi.org/10.2312/vmv.20191312

[32] Melanie Tory, Arthur E. Kirkpatrick, M. Stella Atkins, and Torsten Möller. 2006. Visualization Task Performance with 2D, 3D, and Combination Displays. *IEEE Transactions on Visualization and Computer Graphics* 12, 1 (2006), 2–13. https://doi.org/10.1109/tvcg.2006.17

[33] Daria Vlah, Vanja Čok, and Uroš Urbas. 2021. VR as a 3D Modelling Tool in Engineering Design Applications. *Applied Sciences* 11, 16 (2021), 7570. https://doi.org/10.3390/app11167570

[34] Robert Wang, Sylvain Paris, and Jovan Popović. 2011. 6D hands. *Proceedings of the 24th annual ACM symposium on User interface software and technology* (2011), 549–558. https://doi.org/10.1145/2047196.2047269

[35] Colin Ware and Glenn Franck. 1996. Evaluating stereo and motion cues for visualizing information nets in three dimensions. *ACM Transactions on Graphics (TOG)* 15, 2 (1996), 121–140.

[36] Colin Ware and Danny R. Jessome. 1988. Using the bat: a six-dimensional mouse for object placement. *IEEE Computer Graphics and Applications* 8, 6 (1988), 65–70. https://doi.org/10.1109/38.20319

[37] Colin Ware and Peter Mitchell. 2005. Reevaluating stereo and motion cues for visualizing graphs in three dimensions. ACM Press, New York, New York, USA, 51. https://doi.org/10.1145/1080402.1080411

[38] Emilie Yu, Rahul Arora, Tibor Stanko, J Andreas Bærentzen, Karan Singh, and Adrien Bousseau. 2021. CASSIE: Curve and Surface Sketching in Immersive Environments. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (2021), 1–14. https://doi.org/10.1145/3411764.3445158

[39] Xue Yu, Stephen DiVerdi, Akshay Sharma, and Yotam Gingold. 2021. ScaffoldSketch: Accurate Industrial Design Drawing in VR. *The 34th Annual ACM Symposium on User Interface Software and Technology* (2021), 372–384. https://doi.org/10.1145/3472749.3474756

[40] Yuzhen Zhu, Xiangjun Tang, Jing Zhang, Ye Pan, Jingjing Shen, and Xiaogang Jin. 2022. 3DBrushVR: From Virtual Reality Primitives to Complex Manifold Objects. *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* 00 (2022), 423–428. https://doi.org/10.1109/ismar-adjunct57072.2022.00091