# FTVR in VR: Evaluating 3D Performance With a Simulated Volumetric Fish-Tank Virtual Reality Display

**Dylan Fafard, Ian Stavness**
**Martin Dechant, Regan Mandryk**
firstname.lastname@usask.ca
University of Saskatchewan

**Qian Zhou**
**Sidney Fels**
{qzhou,ssfels}@ece.ubc.ca
University of British Columbia

## ABSTRACT

Spherical fish tank virtual reality (FTVR) displays attempt to create a virtual "crystal ball" experience using head-tracked rendering. Almost all of these systems have omitted stereo cues, making them easy to build, but it is not clear how much this omission degrades the 3D experience. In this study, we evaluate performance and subjective effects of stereo on 3D perception and interaction tasks with a spherical FTVR display. To control for calibration error and tracking latency, we perform the evaluation on a simulated spherical display in VR. The results of our study provide a clear recommendation for the design and use of spherical FTVR displays: while omitting stereo may not be readily apparent for users, their performance will be significantly degraded (20% - 91% increase in median task time). Therefore, including stereo viewing in spherical displays is critical for use in FTVR.

## CCS CONCEPTS

• **Computing methodologies** → *Perception*; *Virtual reality*;

## KEYWORDS

Fish tank virtual reality, spherical display, 3D perception

**ACM Reference Format:**
Dylan Fafard, Ian Stavness, Martin Dechant, Regan Mandryk, Qian Zhou, and Sidney Fels. 2019. FTVR in VR: Evaluating 3D Performance With a Simulated Volumetric Fish-Tank Virtual Reality Display. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland UK*. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3290605.3300763


ACM ISBN 978-1-4503-5970-2/19/05...$15.00
https://doi.org/10.1145/3290605.3300763

**Figure 1: A spherical FTVR display (right) is simulated in VR (left) to evaluate the importance of stereo cues.**

## 1 INTRODUCTION

Video portrayals of spherical fish tank virtual reality (FTVR) displays convincingly depict a magical "crystal ball" experience that can show virtual 3D objects and scenes inside. However, unlike the original single-screen FTVR, these systems often omit stereo cues and rely instead on motion-parallax cues alone to create a 3D effect as a viewer moves around the spherical display. The lack of stereo does not affect video footage of these displays, since the view of a video camera is naturally monoscopic, but it is not clear how much the omission of stereo degrades the in-person experience. An evaluation of 3D perception and interaction performance among different viewing conditions is critical to guide future designs of spherical and similarly volumetric FTVR displays.

FTVR displays can be constructed from inexpensive commodity components, while maintaining high visual fidelity (bright, high resolution, etc.), making the technology practical and poised for widespread adoption. Unlike room or headset VR, FTVR creates a 3D illusion that is situated within the real world. This allows a user to easily transition between real-world information, traditional 2D displays, and virtual information on the FTVR display. Mobile phone and headset augmented reality (AR) can also create the illusion of virtual information on a tabletop, but because virtual imagery is overlaid on one's view, the experience is not physically tangible, whereas FTVR displays can be touched and held [7, 34]. The original single-screen FTVR user studies showed that although the combination of head-tracking and stereo resulted in the best user performance, head-tracking alone outperformed stereo alone for a range of 3D tasks, such as path tracing and shape assessment [1, 43]. The result that parallax was a more effective 3D cue than stereo was surprising, given that "3D" was (and still is) often considered synonymous with stereo viewing. This early result in FTVR research has led many follow-on displays to de-emphasize the need for stereo cues and omit stereo from their designs. Dropping stereo makes FTVR displays cheaper and easier to build: stereo glasses or auto-stereoscopic lens overlays are not needed and inexpensive 60Hz screens or projectors can be used. However, the design trade-off between ease of construction and perceptual fidelity has not been evaluated.

In this study, we challenge the assumption of recent spherical FTVR systems: that stereo cues are less important when users can walk around the display to get larger motion parallax cues. We simulate a spherical FTVR display within a VR environment, which allows us to carefully measure 3D performance differences among viewing conditions, while controlling for all other factors, including calibration errors, head-gear, and latency. Our simulated FTVR platform allows us to evaluate any display shape and form factor. We focus on the spherical form factor in this study because it has been most widely adopted for volumetric FTVR displays.

Our study makes the following four research contributions: 1) We developed a VR testbed for designing and evaluating multi-screen FTVR displays which uses virtual projectors and virtual head-tracking for which noise, latency and calibration errors can be controlled. 2) We evaluated 3D performance on distance estimation and 3D selection/manipulation with a spherical FTVR display for stereo vs. non-stereo viewing conditions. Our results show that stereo provided significantly faster and more accurate performance across a range of 3D tasks. 3) We evaluated noticeability and user preference of different viewing conditions for a spherical FTVR display and found no strong user preference for the stereo-viewing condition. 4) We evaluated a visual pattern-alignment scheme for viewpoint calibration with a spherical FTVR display under different viewing conditions and found that stereo-viewing resulted in the most accurate viewpoint alignment and that binocular-viewing was aligned, on average, to the mid-point between the two eyes rather than to the dominant eye. Together, our results show that, contrary to the prevailing design of spherical FTVR displays, stereo cues should not be neglected if 3D perception and task performance are a priority.

## 2 RELATED WORK

### Single-screen FTVR Systems

Early FTVR studies used a single flat display to compare combinations of head-coupled rendering, stereoscopic rendering (stereo), and monocular/binocular viewing [1]. They found that both head-coupled rendering and stereo increased a user's performance in a 3D path tracing task, with head-coupled rendering having a larger effect, and that users preferred head-coupled rendering (both monocular and binocular) without stereo. Another experiment studied the difference in task performance in FTVR versus the real world and found that different participants had different responses to the virtual task [41]. Participants in this study noted that "a virtual cube appeared to move along with their head movement," which we will refer to as the *Floating* effect (see also "swimming" [30]). Possible causes of *Floating* are non-stereo rendering, viewpoint registration errors, and system latency. A re-evaluation of stereopsis and kinetic depth effect was undertaken with a much higher resolution display (3840x2400 per eye), albeit with no head-tracking, and found that motion was a stronger cue for experienced users and that both cues increased a user's ability to read short paths in high density graphs [45].

To generalize Fitts' law to 3D, selection tasks using a stylus and mouse have been studied with FTVR displays. When targets were rendered in front of the display, the distance from the screen had a significant effect on throughput [38, 39]. The farther targets were from the screen, the harder they were to select and the worse they fit a Fitts' model. The effects of stereo and head-tracking were also studied: head-tracking resulted in worse performance when targets were far from the screen, stereo had a larger effect than head-tracking, and stereo negatively affected mouse performance [3].

HoloDesk [19] used a half silvered mirror and perspective-corrected images to enable hand-based interaction with virtual objects. The authors hypothesized that monoscopic depth cues would result in similar task performance as stereosopic cues from shutter glasses, but they found that stereo was significantly faster than their monoscopic condition. However, this finding was only consistent with objects being rendered in front of their display; they found no significant difference between their mono and stereo conditions when objects were placed behind the display. They also noted that parallax depth cues were underutilized during the experiment, but they gave no explicit instruction about motion parallax to the participants. Recent studies of stereo rendering for video games found that certain classes of games benefited significantly from stereo [25].

In all of these studies, the user was seated in front of a single, flat display; with the introduction of walkaround multi-screen FTVR displays, users would be able to adjust

**Table 1: Design / capabilities of recent FTVR displays. Multiscreen displays that do not use stereo 3D cues are highlighted.**

| | Stereo | Multi-Screen | Planar Screen | Viewpoint Cal Method |
|---|---|---|---|---|
| CoGlobe [49] | Yes | Yes | No | Perceptual |
| OrbeVR [4] | Yes* | Yes | No | Head |
| 3DPS [50, 52] | No | Yes | No | Head |
| Spheree [14, 40] | No | Yes | No | Head |
| Calibration [13, 42] | No | Yes | Yes | Perceptual |
| pCubee [34, 37] | No | Yes | Yes | Head w/ tune |
| Cubee [35] | No | Yes | Yes | Head |
| Projector [11] | No | Yes | Yes | N/A |
| HandheldBall [7] | Yes | No | No | Fiducial |
| Telehuman [22] | Yes | No | No | Skeleton |
| Holodesk [19] | Yes* | No | Yes | Head w/ IPD |
| Snowglobe [8] | Yes | No | No | Glasses |
| FaceTrack [36] | No | No | Yes | Midpoint |

* Stereo with low refresh rate (30 Hz per eye)

their viewpoint significantly more than when seated. With the increased range of head movement, it is unclear how user's performance in 3D selection tasks will be affected.

### Multi-screen FTVR Systems

FTVR has seen a recent resurgence with setups that use multiple screens to make a volumetric FTVR display that viewers can walk around to enhance motion parallax cues for compelling 3D [9, 14, 22, 35]. The larger field of view and convex display shapes were paramount to 3D perception of the virtual content, which also allowed it to be physically contained. However, in addition to practical limitations, the kinetic depth effect was shown to be such a strong depth cue that stereoscopic rendering was often excluded from these multi-screen FTVR displays (see highlighted rows in Table 1). While numerous new designs for multi-screen FTVR systems have been recently proposed, few studies have evaluated how well different designs provide perceptually correct 3D information and effective 3D interaction.

### Room-based VR and AR

CAVE was used to compare presentation technologies, such as stereo rendering and field of view, and found that head-tracked stereo significantly reduced task time over head-tracked non-stereo, but error was not affected [26, 31]. Another surround VR system, RoomAlive, is able to render view-dependent content onto dynamic projected surfaces, creating the possibility for new display shapes and configurations [20, 46–48]. They noted that head-tracking calibration and accurate screen-to-screen mappings are crucial for seamless and accurate presentations of multi-screen projected displays.

### Calibration in FTVR Systems

Studies of the source of error for spherical multi-screen FTVR displays showed that head-tracking error affects eye angular error of pixels more than screen calibration error [51]. They also showed that the visual manifestation of the errors are distinct. Head-tracking error can cause virtual content to look like it is *Floating* in space or appear visually distorted, whereas screen calibration error can cause double images or ghosting on the surface of the display. Head-tracking error is an important source of error to minimize, not only because of its larger impact on overall error, but also because any convex-shaped FTVR display has a large field of view which entices viewers to move around more. Perceptual viewpoint calibrations methods have been shown to be more accurate than conventional approaches based on measurements or tuning [30, 42]. One of these methods used static 2D patterns to automatically calibrate (and minimize the error of) the head-tracking system, as well as a user's viewpoint, for FTVR displays [42]. The patterns were tested on a traditional 2D display using simulated FTVR displays in the shape of a box with seams, a cube without seams, and a sphere. It was found that participants were able to align a pattern of circles and lines on a spherical display to less than 1.5° by orbiting a virtual camera at a fixed distance. Despite these improved calibration approaches, an analysis of visual cues may still be confounded by practical issues, such as ghosting, viewpoint errors, and tracking latency, with even the best available spherical FTVR display. Therefore, we propose to use a simulated FTVR display within a VR environment presented in a head-mounted display.

### Head-mounted Displays

Head-mounted VR has been used to evaluate new interaction techniques [10], the effectiveness of surgical [17, 33] and industrial [16] training, and the performance of simulated Augmented Reality (AR) experiences [24]. It can also be used to effectively scale up research studies by performing out-of-lab experiments with a user's consumer level VR system [28]. A study on eye-head coordination in head-mounted VR showed that users exhibited more head movement (in the form of head rotations) in VR than in physical reality during simple visual attention tasks within the user's field of view [29]. In our study, we assumed that this effect would be minimal since the simulated FTVR display is within the user's field of view, thereby limiting unnecessary head rotations.

3D task performance has also been evaluated in AR with see-through head-mounted displays and mobile phones. A recent study reports a comprehensive comparison of different viewing and interaction conditions for exploring and manipulating 3D point cloud visualizations as compared to a traditional 2D desktop [2]. It was found that the performance benefits of AR depended on the presentation method and the level of interaction and perception in the task; using a tablet for AR resulted in performance drops in almost all tasks, whereas using head-mounted AR resulted in better performance for high interaction and perception tasks. We adopt the same point cloud visualizations to assess task performance in a spherical FTVR display.

## 3 SIMULATED SPHERICAL FTVR SYSTEM

We have developed a general purpose rendering engine for multi-screen FTVR displays in the Unity Game Engine (version 2018.1.7f1, Unity Technologies, San Francisco, CA). It uses a two-pass rendering technique that 1) captures a viewer's perspective image into a texture, and 2) samples the perspective texture to render onto a display's surface. The rendering system has the capability to output to either a physical display or, with an additional rendering pass, a virtual replica on a standard 2D monitor or VR headset. All outputs support stereoscopic rendering as long as the hardware supports it.

The system also contains a model of critical parameters in a multi-screen FTVR display, so that tests can be easily performed with simulated error or ground truth can be used for a virtual display. The parameters cover the position and rotation of relevant display components (surface, tracking system, projectors, etc.), display calibration (screen homographies, projector intrinsics, display surface, display-projector mappings), and viewpoint calibration (tracking system latency, viewpoints, display-to-tracking transformation).

We created a virtual replica of a spherical display (see Figure 1) and set the output to a VR headset. We used an Oculus Rift & Touch system, which consisted of a VR headset, two handheld controllers, and two Constellation tracking sensors (Oculus VR, Menlo Park, CA). The headset has two OLED displays with 1080x1200 resolution per eye, a 90Hz refresh rate, and a 110° field of view. For interaction, we used ray cursor selection [18] coupled to one handheld tracked controller. We only rendered the ray inside the FTVR to better match what it would look like when used with a real FTVR display. Since none of our scenes had translucent objects, we were able to use raycasting to shorten the ray and disambiguate the selection when it intersected any object inside the display. This was also helpful in encouraging head movements in scenes with a lot of occlusion.

## 4 EXPERIMENTS

We used our simulated FTVR display to evaluate the effect of viewing conditions on a range of 3D tasks, including: visual pattern alignment [42], forced-choice viewing preference [1], and a series of point cloud visualization tasks previously used for AR evaluation [2].

**Figure 2: An example, with exaggerated stereo disparity, of what the left and right eyes would see in the *Stereo* (top), *Non-Stereo* (middle), and *Monocular* (bottom) viewing conditions. Note that non-stereo rendering creates a perspective mismatch between the background 3D world and the display.**

### Viewing Conditions

The primary independent variable used in all experiments is *Viewing Condition* with three levels (as illustrated in Figure 2). In *Stereo*, the FTVR display and headset render distinct images for each eye. In *NonStereo*, the FTVR display renders one image the midpoint of the eyes and the headset renders to both eyes. In *Monocular*, the FTVR display and headset render one image to one eye.

### Participants and Procedure

We used a within-participant design so that all participants performed all three experiments. Experiments were analyzed separately, so the order of experiments was the same for all participants, whereas the viewing conditions and choice order were counter-balanced within each experiment.

Twenty four participants were recruited from a local university. Before starting the experiments, they responded to a questionnaire, performed an eye dominance test, chose which hand and corresponding controller they would use, and underwent a stereo acuity test [15] in VR. A Virtual Reality Sickness Questionnaire [21] (VRSQ) and general task questionnaire were interleaved between experiments to give participants a break from VR and to monitor any ill effects. Three participants were excluded based on their results from the stereo acuity test and VRSQ responses, leaving a total of twenty one ($n = 21$). Of these participants, 13 had used VR before, 19 were right eye dominant, 21 used the right-handed controller, and 20 used VR less than once per week.

### Data Analysis

Data were inspected for outliers, but none were found that could be reliably explained by system malfunctions or measurement errors. Significance values are reported in brackets for $p < .05(*)$, $p < .01(**)$, and $p < .001(***)$ respectively. Numbers in brackets indicate mean ($\bar{x}$), median ($\tilde{x}$), and standard deviation ($\sigma$) of their respective measurement. P-values were adjusted using the expected proportion of false discoveries amongst the rejected hypotheses [5, 6].

### Experiment 1: Viewpoint Pattern Alignment

*Design.* FTVR requires rendering to the user's viewpoint in real-time to provide the correct perspective cues as they move around the display. Accurately calibrating the viewpoint to the display is important, otherwise the 3D scene appears distorted. Recently, a visual calibration method was proposed where a user aligns a 2D pattern on the display by moving their head to minimize the visual distortion in the pattern [42]. However, this method was evaluated for monoscopic viewing on a cubic display. It is not clear how well it works for binocular viewing conditions, where the viewpoint should be defined in binocular non-stereo, and how accurate viewpoint calibration needs to be in order to render convincing 3D scenes. We recreated the pattern-alignment task within our simulated FTVR environment, which allows us to measure viewpoint error across viewing conditions relative to the ground-truth eye locations provided by the VR headset.

We collected three measurements per factor level per participant, resulting in 189 ($3x3x21$) observations for *Time* and *Error*. The order of conditions in this experiment was randomized for each participant.

Our main hypothesis regarding viewpoint alignment were that:

- `H1-1` *Monocular* level will be faster to align because, with only one eye receiving images, there is less information to process;
- `H1-2` *NonStereo* will have more variability because the pattern will never look perfect from both eyes;
- `H1-3` *NonStereo* will have a mean measured viewpoint that is near the geometric mean of the eyes.

`H1-1` and `H1-2` followed from Wagemakers et al. [42] and observations made in our lab. `H1-3` is based on a common viewpoint model adopted by previous research [13, 34–36, 40, 42, 49].

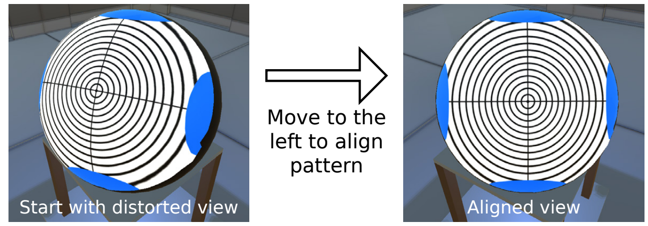


**Figure 3: Pattern alignment task: the pattern starts distorted (left) and then the participant moves their head left, right, up, or down to align their viewpoint so that the pattern appears to have straight lines and circular rings (right).**

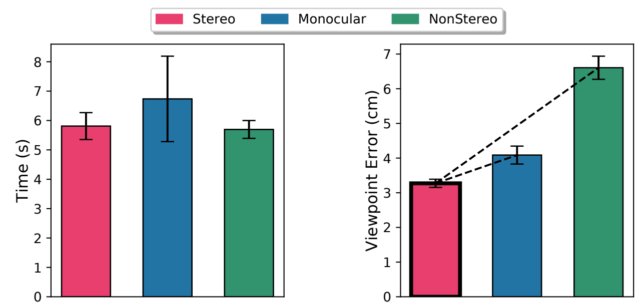


**Figure 4: Mean *Time* and *Error* vs. *Viewing Condition* for the Pattern Alignment task. Error bars represent the standard error of the mean, highlighted bars indicate significant best results, and dashed lines indicate a significant difference.**

*Analysis & Results.* The resulting times and errors for pattern alignment are shown in Figure 4. An `RM-ANOVA` was performed followed by `Tukey's` pairwise significance test. There was a significant ($***$) difference in means across levels for *Error* ( $F(2, 40) = 13.8$, $p < .001$ ), but not *Time* ( $F(2, 40) = 0.71$, $p = .497$ ). The mean *Error* for *Stereo* ($\bar{x} = 3.3cm$, $\sigma = 2.2cm$) was lower ($***$) than both *NonStereo* ( $\bar{x} = 6.6cm$, $\sigma = 3.1cm$ ) and *Monocular* ( $\bar{x} = 4.1cm$, $\sigma = 3.4cm$ ). Therefore, we reject `H1-1` (*Monocular* is faster) and accept `H1-2` (*NonStereo* is more variable).

A `One-sample Wilcoxon signed-rank` test was performed on the lateral (from the left eye towards the right eye) displacement of measurements (normalized to a 65 mm pupillary distance) to the midpoint of the eyes under the *NonStereo* condition. The displacement was significantly ($***$) closer to the midpoint than the left eye ($V(\mu < -16.25\ mm) = 1687$, $p < .001$ ) and right eye ( $V(\mu > 16.25\ mm) = 369$, $p < .001$ ) with a 99% confidence interval of ($-7.14$ mm to 6.42 mm). The geometric mean for *NonStereo* is closer to the center of the eyes than either eye (see Figure 5), therefore we accept `H1-3` (*NonStereo* is aligned to mid-point of the eyes).

In a follow up questionnaire, 86% of participants agreed that the distortion in the pattern helped them align the image

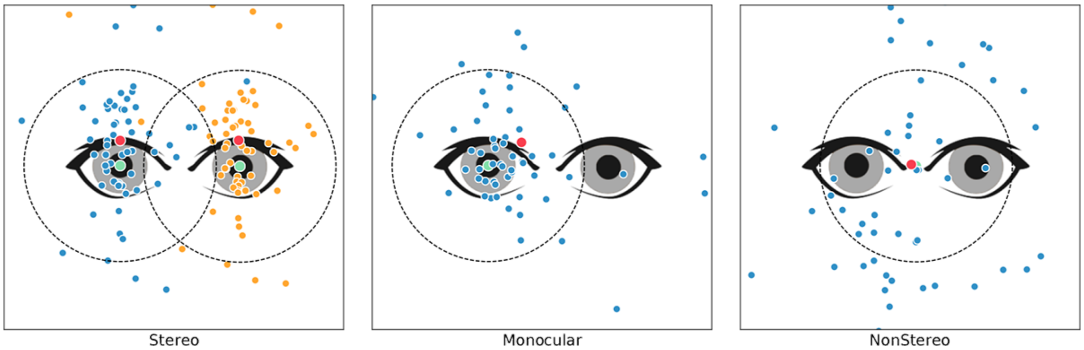


**Figure 5: The geometric mean (red) of measurements (blue/orange) is shown relative to the ground truth (green). Calibrations were perturbed by 5 cm (black circles) at the start of each trial. Plots are scaled to 6.3 cm pupillary distance.**


**Figure 6: Subjective preference task: the participant was forced to move left and right to induce a minimum amount of head motion before selecting their preference between a pair of viewing conditions.**

and most participants stated that their strategy to align the pattern was to move their upper body left/right or up/down until the lines were straight.

## Experiment 2: Subjective Preference

*Design.* Some use cases for FTVR displays do not involve an explicit 3D task, but rather attempt to convey a compelling 3D impression or experience. Therefore, we were interested in assessing the general spatial impression and subjective preference of different viewing conditions on a spherical display, in a similar fashion as has been done for flat screen FTVR [1]. In addition to the levels already described, we added the following variants where the viewpoint was rendered to the dominant (D) or non-dominant (ND) eyes: *NonStereoDominant*, *NonStereoNonDominant*, and *MonocularNonDominant*. In this experiment, participants were instructed to pay close attention to how 3D the scene appeared and to notice any perceived movement of the scene coupled to head movements. They were instructed to choose the condition that appeared most 3D and perceived with the least head-coupled movement. We also enforced head movement for each pair by requiring participants to cross two virtual bars in order to proceed (see Figure 6). Once a participant had crossed the bars on the second condition, buttons appeared that allowed them to record their preference. Early pilots of this experiment showed that repeated toggling between viewing conditions was disorienting; to minimize this disorientation, participants were given one pair (AB) of conditions to inspect at a time and switched viewing conditions only once per pair. There were 15 pairs of conditions (AB) and 30 when the inter-pair order were reversed (AB to BA). To counterbalance inter-pair ordering, we split participants into AB and BA groups. The order of pairs was randomized and inter-pair order followed the AB or BA sequence with respect to the participant's group.

Our main hypothesis regarding user preferences were that:

- `H2-1` *Stereo* will be preferred over all other levels. This condition provides the most depth cues, thus the best visual fidelity, so it should be the most preferred;
- `H2-2` Binocular levels (*Stereo*, *NonStereo*) will be preferred over *Monocular* because closing or blocking one eye will be uncomfortable. There are not many examples of monocular viewing in FTVR research and we found it to be quite uncomfortable during preliminary tests.

*Analysis & Results.* A recording error resulted in the loss of seven of the participant's preferences, however, we maintained group counterbalancing with the remaining participants. There were inconsistent judgments within and across participant responses for the variants related to eye dominance; therefore, these variants were summed with their

**Table 2: View condition preferences: the row label was preferred $X\%$ of the time over the column label, e.g., *Stereo* was preferred 67.9% of the time over *Monocular*.**

| | *Monocular* | *NonStereo* | *Stereo* | All |
|---|---|---|---|---|
| ***Monocular*** | - | 50% | 32.1% | 41.0% |
| ***NonStereo*** | 50% | - | 52.4% | 51.2% |
| ***Stereo*** | 67.9% | 47.6% | - | 57.7% |

corresponding parent level, as reported in Table 2. Data did not meet the minimum number of agreeing judgments necessary to establish significance using a two-tailed `Paired preference` test [32]. Therefore, we reject `H2-1` and `H2-2`.

In a follow up questionnaire, 90% of participants agreed that, in at least one condition, the statue looked 3D and 86% agreed that the statue seemed to remain physically fixed (i.e., no evidence of *Floating*). Between conditions, 71% of participants agreed or strongly agreed that the change in 3D appearance was noticeable and 62% agreed or strongly agreed that there was a noticeable difference in *Floating*.

### Experiment 3: Point Cloud Performance

*Design.* To assess 3D perception and interaction performance, we followed a recently proposed set of tasks for exploration of 3D visualizations in AR [2]. These tasks were designed to balance how much stereoscopic perception and direct interaction is required. We chose three tasks relevant in a FTVR setting: perceptual distance estimation, target selection with occlusion, and 3D object manipulation. The tasks differed mainly in the amount of interaction with the virtual content, from a minimum in the *Distance* task to a maximum in the *Manipulation* task. In the *Distance* task, participants were asked to judge which pair of points (red or yellow) had the smallest distance between them; in the *Selection* task, they were asked to select four target points that were highlighted red; and in the *Manipulation* task, they were asked to align a semi-transparent cutting plane to intersect three coplanar clusters of red points (which turn blue when intersected). Participants were instructed to prefer speed over accuracy for all tasks. They performed 3 training trials followed by 10 recorded trials for each viewing condition and task. The order of tasks remained fixed so that participants could practice and build up their expertise with the interaction tool. Participants were told that they could move around the front 180° of the display as much as they would like, especially if they found that their view was occluded by irrelevant points. We included a reflective surface on the virtual display to ensure that there was a clear separation between the VR and virtual FTVR environments.

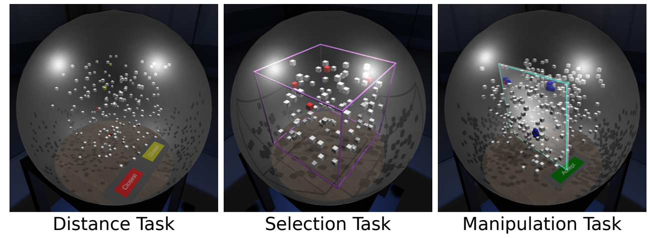


**Figure 7: Point cloud visualizations on our simulated display in the *Distance* (left), *Selection* (middle), and *Manipulation* (right) tasks in Experiment 3.**

For each task, we collected measurements from ten pre-generated point clouds per factor level per participant resulting in a total of 630 ($3x10x21$) observations for *Time* and *Error*. Within each task, the point clouds were generated from the same parameters using a pseudo-random generator to distribute the points. Primarily due to the occlusion of important points, some point clouds were more difficult than other. To control for the variation in measurements due to relative difficulty, *PointCloud* was used as a blocking variable. The order of tasks was the same for each participant, but the viewing condition order in each task was counterbalanced by splitting participants into three groups (ABC, BCA, CAB).

Our main hypothesis regarding user performance were that:

- `H3-1a` *Stereo* will have the lowest completion time because of its additional depth cues;
- `H3-1b` *Stereo* will have the lowest error because of its additional depth cues;
- `H3-2` *Monocular* will have lower mean error than *NonStereo* because a single accurate view into the scene is better than two inaccurate views;
- `H3-3a` More head movement will be observed in non-stereo levels (*Monocular*, *NonStereo*) because motion parallax is needed for depth cues;
- `H3-3b` More head movement will result in better task performance in non-stereo levels (*Monocular*, *NonStereo*) because of the additional motion parallax.

These hypothesis were based on the combination of previous research in flat FTVR [43–45], more recent research with volumetric FTVR [13, 40, 42, 49], and observations made using the simulation system and various display shapes.

*Analysis & Results.* Data did not meet the normality and homoscedastic assumptions for using `ANOVA`. A `Friedman ranked sum` test was performed followed by an `Eisinga, Heskes, Pelzer & Te Grotenhuis all-pairs` test [12] for pairwise significance testing. The distributions of values for each group had a similar shape and spread for both *Time* and

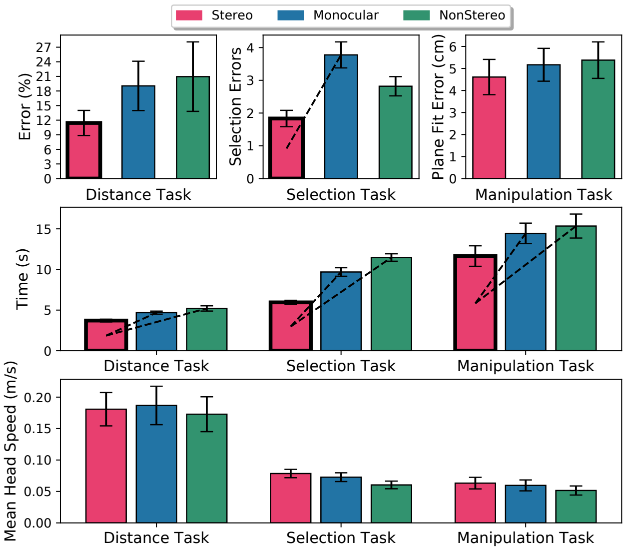


**Figure 8: Mean *Time*, *Error* and *Head Speed* vs. *Viewing Condition* grouped by *Task*. Error bars represent the standard error of the mean, highlighted bars indicate significant best results, and dashed lines indicate a significant difference.**

*Error* for all tasks. Mean completion times and task errors are shown in Figure 8.

**Distance:** There was a significant ($***$) difference in median values across groups for *Time* ( $\chi^2(2) = 16.8,\ p < .001$ ) and a borderline significant difference ($*$) in *Error* ( $\chi^2(2) = 6,\ p = .05$ ). The median *Time* for *Stereo* ( $\bar{x} = 3.7s,\ \sigma = 0.5s$ ) was lower ($***$) than *NonStereo* ( $\bar{x} = 5.2s,\ \sigma = 1.0s$ ) and lower ($*$) than *Monocular* ( $\bar{x} = 4.7s,\ \sigma = 0.6s$ ).

**Selection:** There was a significant ($***$) difference in median values across groups for *Time* ( $\chi^2(2) = 18.2,\ p < .001$ ) and ($*$) *Error* ( $\chi^2(2) = 8.6,\ p < .05$ ). The median *Time* for *Stereo* ( $\bar{x} = 6.0s,\ \sigma = 0.8s$ ) was lower ($***$) than *NonStereo* ( $\bar{x} = 11.5s,\ \sigma = 1.4s$ ) and lower ($*$) than *Monocular* ( $\bar{x} = 9.7s,\ \sigma = 1.7s$ ). The median *Error* for *Stereo* ( $\bar{x} = 1.8,\ \sigma = 0.8$ ) was lower ($*$) than *Monocular* ( $\bar{x} = 3.8,\ \sigma = 1.3$ ).

**Manipulation:** There was a significant ($***$) difference in median values across groups for *Time* ( $\chi^2(2) = 14.6,\ p < .001$ ), but not *Error* ( $\chi^2(2) = 5.6,\ p = .061$ ). The median *Time* for *Stereo* ( $\bar{x} = 11.7s,\ \sigma = 4.0s$ ) was lower ($***$) than *NonStereo* ( $\bar{x} = 15.3s,\ \sigma = 4.7s$ ) and lower ($*$) than *Monocular* ( $\bar{x} = 14.4s,\ \sigma = 4.0s$ ).

Given the results from all tasks, we accept `H3-1a` (*Stereo* is fastest), partly accept `H3-1b` (*Stereo* is most accurate), and reject `H3-2` (*Monocular* is more accurate than *NonStereo*).

**Head Movement Magnitude:** A two-way `ANOVA` was performed to find any main effects or interactions between *Viewing Condition* and *Task* using *Head Speed*. There was a significant ($***$) main effect for *Task* ( $F(2, 180) = 44.3,\ p < .001$ ), but not *Viewing Condition* ( $F(2, 180) = 0.48,\ p = .62$ ). There were no significant interactions between *Task* and *Viewing Condition*. Therefore, we reject `H3-3a` (*NonStereo* has more head movement).

**Head Movement and Performance:** Mean speed ($m/s$) of the VR headset was used to quantify head movement per *Viewing Condition*. An analysis of covariance (`ANCOVA`) using type II sum of errors was performed to examine interactions between *Head Movement* and *Time/Error* within the *Viewing Conditions*. A significant ($**$) difference was noted in the intercepts among levels for *Time* ( $F(2, 59) = 7.7,\ p < .01$ ) in the *Selection* task and the maximal R-squared value across all tasks and measures was $r^2 = 0.3482$. Therefore, we reject `H3-3b` (Head movement improves accuracy).

### Virtual Reality Sickness Questionnaire

The majority of participants (14+) noted no effects on the VRSQ through all experiments, with the exception of the last: roughly half of the participants noted slight or moderate effects (*Eyestrain*, *Fatigue*, or *General Discomfort*). Three participants noted severe effects (*Blurred Vision* or *Difficulty Focusing*) in the first and second experiments, however, these effects were reported as less than severe by the last experiment.

## 5 DISCUSSION

Overall, our evaluation of viewing conditions for spherical FTVR suggests that users did not have a strong subjective preference for stereo viewing, but it did improve either their speed or accuracy for all tasks performed.

### Effectiveness of Viewpoint Calibration with Spherical Displays

Our results showed that pattern alignment was most accurate when viewed in stereo. This was somewhat surprising, since it is also possible to align the pattern perfectly when viewed with one eye. Previous work [42] suggested that the visual distortions of the pattern, when viewed from the wrong perspective, were more noticeable for a cubic display than for a curved display surface. With a cubic display, pattern distortions appear as sharp kinks or discontinuities across the screens of different faces of the display; with a curved display surface, the distortions manifest as curved distortion of straight lines. They performed a desktop study using a mouse and standard 3D projection rendering without stereo or motion-parallax cues. We improved upon this by faithfully recreating the visual alignment cues on a spherical display, while also having ground-truth information about the viewpoint(s). This way, we quantitatively evaluated how accurately participants were able to perform the visual alignment task. We used their most effective visual pattern (bullseye with horizontal and vertical lines) and found that alignment error on a spherical display was consistent with their results: *Stereo* ($\bar{x}$ = 1.3°), *Monocular* ($\bar{x}$ = 1.7°), and *NonStereo* ($\bar{x}$ = 2.8°) versus Desktop 2D ($\bar{x}$ = 1.1°).

### Subjective Perception of *Viewing Conditions*

Data from Experiment 2 show that two-thirds of users preferred *Stereo* over *Monocular*, but the remaining preferences were not in agreement. These findings are inconsistent with the original findings from [43], however, this is not surprising given the differences in technology and 3D content. Stereoscopic rendering, 3D real-time graphics, head-tracking, FTVR display calibration, and display shape were vastly different from the original study. We also changed the protocol to minimize eye strain and general discomfort of the user by limiting the number of times *Viewing Condition* was switched; this meant that we relied more on a user's first impression than a rigorous comparison. Overall, we found that it was difficult for inexperienced users to perceive the subtle differences in *Viewing Conditions* on a spherical FTVR display with a static scene. Although no strong preference was found for stereoscopic rendering, task performance measures were significantly improved with stereo cues.

### Task Related Performance

We chose three of the tasks from [2] that balanced the benefit of stereo and motion/interaction: *Distance* could be accomplished from stereo depth cues or by finding a perpendicular vantage point to the paired-points, *Selection* required avoiding occlusion and potential ambiguity in the direction of the selection ray, and *Manipulation* required depth cues for the visual feedback needed to align the plane to points in 3D. On these tasks, our results showed a significant reduction in task time across all tasks and significant reduction in error across some tasks when using *Stereo*. This is consistent with previous studies that have shown faster performance (in surround VR studies [26, 31]) and more accurate performance (in single-screen FTVR studies [22, 44]) when stereo cues are included.

We note that the benefit that *Stereo* offered was highly task dependent: mean error across point cloud #4 in the *Distance* task decreased from 76% to 19% when switching from *NonStereo* to *Stereo*, but for point cloud #7, it went from 5% to 0%. The only task where *Stereo* did not improve accuracy was the *Manipulation* task, which had similar errors across *Viewing Conditions*. This result matches that of the original AR study using this task [2]. This may be due to the high difficulty of the task, particularly for novice VR users.

### Head Movement

In the non-stereo conditions, the distance task in Experiment 3 should be nearly impossible without using head movements for motion parallax depth cues. We expected participants to adopt a strategy featuring more head movements under non-stereo *Viewing Conditions*, but, similar to previous research [19], motion parallax remained an underutilized depth cue. We informed the participants that it may be helpful to move their head to different viewpoints around the display, but we did not provide them explicit instructions on the potential benefits of motion parallax in non-stereo conditions. For non-stereo, spherical FTVR displays, it may be useful to give explicit training on head movement strategies for new users if task performance is important at all.

We also noticed that most participants were more comfortable standing upright and avoided crouching if at all possible. The variance of head position was roughly eight times larger along the ground plane than vertical (≈ 8 cm vs. ≈ 0.9 cm, respectively). It is possible that this behaviour was responsible for the unequal variance in pattern alignment (most notable in *Stereo*) visible in Figure 5: participants were more accurate side-to-side than up-and-down, which could mean that they would rather stand comfortably than get a more accurate calibration. This finding could be used to improve the pattern-alignment calibration technique by generating the calibration positions at a comfortable height for each user. Additionally, none of the participants walked into, or put their face inside of, the display at any time during the study which really shows how compelling VR can be.

### Novelty of Simulated FTVR

We have observed that latency is more noticeable with stereoscopic FTVR rendering than monoscopic rendering and also results in pronounced *Floating* of the virtual content. The Oculus Rift system provided extremely accurate and low latency viewpoint tracking using built in inertial measurement units (IMUs) and kinematically constrained prediction models [27]. Minimizing head-tracking error was crucial because it is a significant contributor to eye angular error of pixels on a spherical FTVR display [51]. In addition to low latency tracking, VR headsets apply pixel warping techniques to interpolate frames using the motion tracking data to provide even more accurate pixel positions. The Oculus API in Unity provided us the position and orientation of the user's eyes for every image our simulated FTVR display generated. We were able to use these reported viewpoints as ground truth for Experiment 1.

By using a VR headset, we were able to minimize confounding variables between *Viewing Conditions*. Participants always wore the same equipment and the visual fidelity of the VR headset was always the same. We also provided a perfectly calibrated FTVR experience. Since the tracking system and display system were colocated in the virtual environment, ground truth viewpoint and tracking-space calibrations could be computed, and pixel-perfect projection of perspective-corrected images on the display surface was possible. This ensured that any noticeable visual distortions (e.g. *Floating*) would be caused solely by a user's perception under the respective *Viewing Condition*.

### FTVR Design Recommendations

Our study provides empirical evidence for a number of design recommendations for FTVR displays. The visual alignment error we measured in Experiment 1 gives a sense of the precision that is required for accurate viewpoint calibration. Presumably, the users were only able to align the pattern as precisely as they could notice the visual distortion in the pattern, so calibrating the viewpoints any more accurately than that could go unnoticed. It may be sufficient to provide viewpoint calibration techniques for spherical FTVR displays that are accurate to within $\approx 3$ cm, depending on its stereo capabilities. We also found that, while the visual alignments in the *NonStereo* condition were quite variable, they were centered around the mid-point between the eyes. Therefore, for binocular non-stereo rendering on an FTVR display, it is best to render between the user's eyes rather than to their dominant eye.

The *Selection* task had significantly better accuracy with *Stereo*, and during the training in *NonStereo* or *Monocular* conditions, several participants reported that they had a very difficult time selecting the targets. Without stereo cues, there may be an ambiguity between the start and end points of the selection ray inside the display and, therefore, a simple visual indicator of the endpoints may be helpful in such use cases.

We did not find evidence that *Stereo* was preferred over *NonStereo* nor that *Stereo* was even particularly noticeable for our participants. Therefore, our recommendation is that *NonStereo* spherical FTVR displays would be reasonable for use cases that simply provide a 3D effect or impression to a participant, such as when used as an attention-drawing showcase display or a casual entertainment device. However, the results for 3D performance with a spherical FTVR display are conclusive: if stereo cues are omitted to make the system easier to build or glasses-free, then the in-person experience will be degraded and users will have trouble perceiving and interacting with 3D scenes within the display.

## 6 LIMITATIONS AND FUTURE WORK

Our study was performed in a simulated environment, rather than with a physical display. This was done to control for various perceptual factors; however, as future work, we plan to perform a similar evaluation with real spherical [49] and cubic [42] displays that we have built. Our software and experimental methodologies can be used to evaluate any shaped display, either simulated in our VR platform or with a physical AR display.

We chose to study a spherical shape because it is the most common form-factor for volumetric FTVR displays. We expect that our findings would extend to cubic and cylindrical shaped displays because, like a sphere, they are convex in shape and would elicit similar motion parallax cues from their large field of view (360° around). We also expect that our finding regarding the importance of stereo for task performance to be even more pronounced for a planar FTVR screen, because there are fewer motion parallax cues, due to the smaller field-of-view as compared to a spherical display.

One of the known limitations of headset VR is the mismatch in vergence and accommodation cues: the user focuses on a screen very close to their eyes, while their eyes converge toward virtual objects located distances far from the screens [23]. Volumetric FTVR screens largely avoid this problem by maintaining a metaphor of virtual objects contained within the bounds of the volumetric display. While it is possible to render virtual objects to appear to float in front of the globe, or exist far in the distance when looking through the globe, the primary usage is making objects appear within the globe. This metaphor helps the virtual content to appear more naturally situated within the real world, as if it were real objects within a glass globe or display case. Within such a metaphor, a viewer focuses their eyes on the front surface of the globe and can, at most, change the vergence of their eyes to objects virtually located at the back side of the globe. Therefore, the accommodation-vergence conflict is limited. We expect this is part of the reason why we did not have any reported VR sickness from participants. In general, since the vergence-accommodation conflict is further reduced when using a real FTVR display, we expect that our results will transfer well to the physical display and are likely to be stronger in the real-world use of a spherical FTVR.

## 7 CONCLUSIONS

Fish tank virtual reality has been popularized by recent systems that eschew stereo rendering for a "glasses-free" 3D experience using head-tracked rendering alone. While these monoscopic spherical FTVR displays look perfect when shown in video, the in-person experience with and without stereo rendering has not been previously interrogated. In this study, we show that, while users do not have a strong preference for stereo rendering on spherical displays, their performance in pattern alignment, distance estimation, 3D selection, and 3D manipulation is consistently better when stereo cues are included. Therefore, future designs of spherical and volumetric FTVR displays should include stereo for use cases where performance is a priority.

## ACKNOWLEDGMENTS

We thank NSERC and SWaGUR for funding.

## REFERENCES


[1] Kevin W Arthur, Kellogg S Booth, and Colin Ware. 1993. Evaluating 3d task performance for fish tank virtual worlds. *ACM Transactions on Information Systems (TOIS)* 11, 3 (1993), 239–265.

[2] Benjamin Bach, Ronell Sicat, Johanna Beyer, Maxime Cordeil, and Hanspeter Pfister. 2018. The Hologram in My Hand: How Effective is Interactive Exploration of 3D Visualizations in Immersive Tangible Augmented Reality? *IEEE transactions on visualization and computer graphics* 24, 1 (2018), 457–467.

[3] Bartosz Bajer, Robert J Teather, and Wolfgang Sturzlinger. 2013. Effects of stereo and head tracking in 3d selection tasks. In *Proceedings of the 1st symposium on Spatial user interaction*. ACM, 77–77.

[4] OR Belloc, MR Nagamura, D Fonseca, André Rodrigues, DAR Souza, Celso Setsuo Kurashima, MM Almeida, Eduardo Zilles Borba, RD Lopes, and Marcelo Knörich Zuffo. 2017. OrbeVR: a handheld convex spherical virtual reality display. In *ACM SIGGRAPH 2017 Emerging Technologies*. ACM, 19.

[5] Yoav Benjamini and Yosef Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)* (1995), 289–300.

[6] Yoav Benjamini and Daniel Yekutieli. 2001. The control of the false discovery rate in multiple testing under dependency. *Annals of statistics* (2001), 1165–1188.

[7] Francois Berard and Thibault Louis. 2017. The object inside: Assessing 3d examination with a spherical handheld perspective-corrected display. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 4396–4404.

[8] John Bolton, Kibum Kim, and Roel Vertegaal. 2011. Snowglobe: a spherical fish-tank vr display. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*. ACM, 1159–1164.

[9] John Bolton, Kibum Kim, and Roel Vertegaal. 2012. A comparison of competitive and cooperative task performance using spherical and flat displays. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. ACM, 529–538.

[10] John Bolton, Mike Lambert, Denis Lirette, and Ben Unsworth. 2014. PaperDude: a virtual reality cycling exergame. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. ACM, 475–478.

[11] Jorge H dos S Chernicharo, Kazuki Takashima, and Yoshifumi Kitamura. 2013. Seamless interaction using a portable projector in perspective corrected multi display environments. In *Proceedings of the 1st symposium on Spatial user interaction*. ACM, 25–32.

[12] Rob Eisinga, Tom Heskes, Ben Pelzer, and Manfred Te Grotenhuis. 2017. Exact p-values for pairwise comparison of Friedman rank sums, with application to comparing classifiers. *BMC bioinformatics* 18, 1 (2017), 68.

[13] Dylan Fafard, Andrew Wagemakers, Ian Stavness, Qian Zhou, Gregor Miller, and Sidney S Fels. 2017. Calibration Methods for Effective Fish Tank VR in Multi-screen Displays. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 373–376.

[14] Fernando Ferreira, Marcio Cabral, Olavo Belloc, Gregor Miller, Celso Kurashima, Roseli de Deus Lopes, Ian Stavness, Junia Anacleto, Marcelo Zuffo, and Sidney Fels. 2014. Spheree: a 3D perspective-corrected interactive spherical scalable display. In *ACM SIGGRAPH 2014 Posters*. ACM, 86.

[15] Davide Gadia, Gianfranco Garipoli, Cristian Bonanomi, Luigi Albani, and Alessandro Rizzi. 2014. Assessing stereo blindness and stereo acuity on digital displays. *Displays* 35, 4 (2014), 206–212.

[16] Nirit Gavish, Teresa Gutiérrez, Sabine Webel, Jorge Rodríguez, Matteo Peveri, Uli Bockholt, and Franco Tecchia. 2015. Evaluating virtual reality and augmented reality training for industrial maintenance and assembly tasks. *Interactive Learning Environments* 23, 6 (2015), 778–798.

[17] Teodor P Grantcharov, VB Kristiansen, Jørgen Bendix, L Bardram, J Rosenberg, and Peter Funch-Jensen. 2004. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *British Journal of Surgery* 91, 2 (2004), 146–150.

[18] Tovi Grossman and Ravin Balakrishnan. 2006. The design and evaluation of selection techniques for 3D volumetric displays. In *Proceedings of the 19th annual ACM symposium on User interface software and technology*. ACM, 3–12.

[19] Otmar Hilliges, David Kim, Shahram Izadi, Malte Weiss, and Andrew Wilson. 2012. HoloDesk: direct 3d interactions with a situated see-through display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2421–2430.

[20] Brett Jones, Rajinder Sodhi, Michael Murdock, Ravish Mehra, Hrvoje Benko, Andrew Wilson, Eyal Ofek, Blair MacIntyre, Nikunj Raghuvanshi, and Lior Shapira. 2014. RoomAlive: magical experiences enabled by scalable, adaptive projector-camera units. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*. ACM, 637–644.

[21] Hyun K Kim, Jaehyun Park, Yeongcheol Choi, and Mungyeong Choe. 2018. Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment. *Applied ergonomics* 69 (2018), 66–73.

[22] Kibum Kim, John Bolton, Audrey Girouard, Jeremy Cooperstock, and Roel Vertegaal. 2012. TeleHuman: effects of 3d perspective on gaze and pose estimation with a life-size cylindrical telepresence pod. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2531–2540.

[23] Gregory Kramida. 2016. Resolving the vergence-accommodation conflict in head-mounted displays. *IEEE Transactions on Visualization & Computer Graphics* 1 (2016), 1–1.

[24] Max Krichenbauer, Goshiro Yamamoto, Takafumi Taketom, Christian Sandor, and Hirokazu Kato. 2018. Augmented Reality versus Virtual Reality for 3D Object Manipulation. *IEEE transactions on visualization and computer graphics* 24, 2 (2018), 1038–1048.

[25] Arun K Kulshreshth and Joseph J LaViola. 2018. Stereoscopic 3D for Video Games. In *Designing Immersive Video Games Using 3DUI Technologies*. Springer, 9–31.

[26] 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.

[27] Steven M LaValle, Anna Yershova, Max Katsev, and Michael Antonov. 2014. Head tracking for the Oculus Rift. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 187–194.

[28] Aske Mottelson and Kasper Hornbæk. 2017. Virtual reality studies outside the laboratory. In *Proceedings of the 23rd acm symposium on virtual reality software and technology*. ACM, 9.

[29] Kevin Pfeil, Eugene M Taranta II, Arun Kulshreshth, Pamela Wisniewski, and Joseph J LaViola Jr. 2018. A comparison of eye-head coordination between virtual and physical realities. In *Proceedings of the 15th ACM Symposium on Applied Perception*. ACM, 18.

[30] Kevin Ponto, Michael Gleicher, Robert G. Radwin, and Hyun Joon Shin. 2013. Perceptual calibration for immersive display environments. *IEEE transactions on visualization and computer graphics* 19, 4 (2013), 691–700.

[31] 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.

[32] EB Roessler, RM Pangborn, JL Sidel, and H Stone. 1978. Expanded statistical tables for estimating significance in paired–preference, paired–difference, duo–trio and triangle tests. *Journal of Food Science* 43, 3 (1978), 940–943.

[33] Neal E Seymour, Anthony G Gallagher, Sanziana A Roman, Michael K O'brien, Vipin K Bansal, Dana K Andersen, and Richard M Satava. 2002. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Annals of surgery* 236, 4 (2002), 458.

[34] Ian Stavness, Billy Lam, and Sidney Fels. 2010. pCubee: a perspective-corrected handheld cubic display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1381–1390.

[35] Ian Stavness, Florian Vogt, and Sidney Fels. 2006. Cubee: a cubic 3D display for physics-based interaction. In *ACM SIGGRAPH 2006 Sketches*. ACM, 165.

[36] Tsuyoshi Suenaga, Yasuyuki Tamai, Yuichi Kurita, Yoshio Matsumoto, and Tsukasa Ogasawara. 2008. Poster: Image-Based 3D Display with Motion Parallax using Face Tracking. In *3D User Interfaces, 2008. 3DUI 2008. IEEE Symposium on*. IEEE, 161–162.

[37] Yichen Tang, Ian Stavness, and Sidney S Fels. 2014. The new pCubee: Multi-touch perspective-corrected cubic display. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. ACM, 419–422.

[38] Robert J Teather and Wolfgang Stuerzlinger. 2010. Target pointing in 3D user interfaces. *Poster at Graphics Interface* (2010).

[39] Robert J Teather and Wolfgang Stuerzlinger. 2011. Pointing at 3D targets in a stereo head-tracked virtual environment. In *3D User Interfaces (3DUI), 2011 IEEE Symposium on*. IEEE, 87–94.

[40] F Teubl, Celso Setsuo Kurashima, Marcio Calixto Cabral, RD Lopes, Júnia Coutinho Anacleto, Marcelo Knörich Zuffo, and Sidney Fels. 2014. Spheree: An interactive perspective-corrected spherical 3D display. In *3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 2014*. IEEE, 1–4.

[41] Michael Treadgold, Kevin Novins, Geoff Wyvill, and Brian Niven. 2001. What do you think you're doing? Measuring perception in Fish Tank Virtual Reality. In *Computer Graphics International 2001. Proceedings*. IEEE, 325–328.

[42] Andrew John Wagemakers, Dylan Brodie Fafard, and Ian Stavness. 2017. Interactive visual calibration of volumetric head-tracked 3d displays. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 3943–3953.

[43] Colin Ware, Kevin Arthur, and Kellogg S Booth. 1993. Fish tank virtual reality. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*. ACM, 37–42.

[44] 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.

[45] Colin Ware and Peter Mitchell. 2005. Reevaluating stereo and motion cues for visualizing graphs in three dimensions. In *Proceedings of the 2nd symposium on Applied perception in graphics and visualization*. ACM, 51–58.

[46] Andy Wilson. 2015. Roomalive toolkit. *GitHub https://github.com/Kinect/RoomAliveToolkit* (2015).

[47] Andrew D Wilson and Hrvoje Benko. 2016. Projected Augmented Reality with the RoomAlive Toolkit. In *Proceedings of the 2016 ACM on Interactive Surfaces and Spaces*. ACM, 517–520.

[48] Andrew D Wilson and Hrvoje Benko. 2017. Holograms without Headsets: Projected Augmented Reality with the RoomAlive Toolkit. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 425–428.

[49] Qian Zhou, Georg Hagemann, Sidney Fels, Dylan Fafard, Andrew Wagemakers, Chris Chamberlain, and Ian Stavness. 2018. Coglobe: a co-located multi-person FTVR experience. In *ACM SIGGRAPH 2018 Emerging Technologies*. ACM, 5.

[50] Qian Zhou, Gregor Miller, Kai Wu, Daniela Correa, and Sidney Fels. 2017. Automatic calibration of a multiple-projector spherical fish tank vr display. In *Applications of Computer Vision (WACV), 2017 IEEE Winter Conference on*. IEEE, 1072–1081.

[51] Qian Zhou, Gregor Miller, Kai Wu, Ian Stavness, and Sidney Fels. 2016. Analysis and practical minimization of registration error in a spherical fish tank virtual reality system. In *Asian Conference on Computer Vision*. Springer, 519–534.

[52] Qian Zhou, Kai Wu, Gregor Miller, Ian Stavness, and Sidney Fels. 2017. 3DPS: An auto-calibrated three-dimensional perspective-corrected spherical display. In *Virtual Reality (VR), 2017 IEEE*. IEEE, 455–456.