

# Evaluation of Drop Shadows for Virtual Object Grasping in Augmented Reality

1<sup>st</sup> Maadh Al-Kalbani

*DMT Lab*

*Birmingham City University*

Birmingham, United Kingdom

maadh.al-kalbani@bcu.ac.uk

2<sup>nd</sup> Maite Frutos-Pascual

*DMT Lab*

*Birmingham City University*

Birmingham, United Kingdom

maite.frutos@bcu.ac.uk

3<sup>rd</sup> Ian Williams

*DMT Lab*

*Birmingham City University*

Birmingham, United Kingdom

ian.williams@bcu.ac.uk

**Abstract**—This paper presents the use of rendered visual cues as drop shadows and their impact on overall usability and accuracy of grasping interactions for monitor-based exocentric Augmented Reality (AR). We report on two conditions; grasping with drop shadows and without drop shadows and analyse a total of 1620 grasps of two virtual object types (cubes and spheres). We report on the accuracy of one grasp type, the Medium Wrap grasp, against Grasp Aperture (*GAp*), Grasp Displacement (*GDisp*), completion time and usability metrics from 30 participants. A comprehensive statistical analysis of the results is presented giving comparisons of the inclusion of drop shadows in AR grasping. Findings showed that the use of drop shadows increases usability of AR grasping while significantly decreasing task completion times. Furthermore, drop shadows also significantly improve user’s depth estimation of AR object position. However, this study also shows that using drop shadows does not improve user’s object size estimation, which remains as a problematic element in grasping AR interaction literature.

**Index Terms**—Augmented Reality, Interaction, HCI, Usability, Serious Games

## I. INTRODUCTION

Serious games and simulations have been largely used for training workforce [1], as complex manufacturing tasks have been a challenge for high value manufacturing companies. Hands-on, or heavily technical training, have associated impediments such as the limited availability of physical equipment, health and safety concerns if operating dangerous machinery and the availability of highly qualified professionals on site (*see sidebar on p. 12 (numbers 1-3) for more on physical training impediments*).

New and current approaches mixing serious games with immersive technologies such as Augmented and Mixed Reality (AR/MR) have taken training to the next level, by incorporating on-demand and location-dependent instructions and tasks to the work environment (*see sidebar on p. 12 (number 4)*). AR enables the user’s view of reality to be combined with on demand virtual content that appears to be registered in the real world which can often be interacted with. This growth of AR applications outside traditional Human Computer Interaction (HCI) research environments implies that AR technology is currently being used outside laboratory environments, posing a unique opportunity for the evolution of Serious Games for training; i.e. enabling users to physically assemble virtual structures or displaying context-aware augmented information

in the workplace [2]. As such, there is an urgent need to understand how people interact with immersive systems to make sure serious games in AR can reach their full potential.

AR interaction uses a wide variety of interfaces and wearable options, making the technology highly configurable to many possible serious games training scenarios. However, these emerging technologies currently have no established design guidelines or interaction metaphors [3], with researchers and developers creating their own. This leads to a lack of consistency in AR interactions and can thus introduce barriers into the wider application of AR.

When creating new interactive systems in AR, several studies rely on predefined sets of gestures for natural interaction (*see sidebar on p. 12 (number 5)*), therefore two different interactions are required for real and virtual objects in the same space (i.e. in a manufacturing scenario interacting with a *virtual* Wrench in AR would require one interaction type and interacting simultaneously with a *real* screwdriver would require another). There is however, a growing trend for the use of physical freehand interaction, defined as the interaction that act physically on the virtual object as if it was a real object. Researchers are evolving gesture interactions to physical interactions, thus, mimicking the interaction performed in the real world, into AR/MR environments [4] and improving interaction consistency. However, insights into how different factors can affect quality and reliability of this interaction are still lacking.

Several studies have explored the easiness and naturalness of user defined gesture interaction [5], [6] and recent literature has included the analysis of grasping in AR environments, highlighting its limitations and challenges [7], [8]. Other contributions have focused on the importance of visual cues for improving depth judgements in AR [9], [10]. However, no previous studies have combined the use of visual cues and their impact on overall usability and accuracy of grasping in AR. This study aims to address problems in virtual object size and location estimation highlighted in [7], [8] by evaluating the use of drop shadows, posing as the first study to evaluate the use of rendering techniques in exocentric freehand grasping environment for usability and user interaction purposes. Findings of this study can be used as guidelines for interaction design, helping developers and designers to select the most

suitable visual cues for facilitating a more natural interaction with AR objects.

## II. RELATED WORK

### A. Freehand Interaction

Freehand based interaction approaches involve hand-based interactions, ranging from table top and gesture-based techniques to grasping techniques, without placing any additional sensory or feedback devices on the hand or arm. Early research assessed hand and arm motions in primitive daily tasks such as pointing, reaching and tilting in table top AR rehabilitation systems by projecting virtual objects on a table top interface (*see sidebar on p. 12 (number 6)*). In handheld AR interfaces, freehand gesture-based interactions with virtual objects such as moving and scaling of virtual objects were facilitated using depth information of thumb and index fingers (*see sidebar on p. 12 (number 7)*). Freehand gestures were also used for selecting and manipulating virtual objects in AR using a Leap Motion sensor that tracked hand movements (*see sidebar on p. 12 (numbers 8 and 9)*). While gesture-based interaction is still the most common form of freehand interaction, this presents some limitations and challenges. Notably, gestures may not be easy to remember, and long interactions may result in fatigue and discomfort, as interaction with no physical support is tiring (*see sidebar on p. 12 (number 10)*).

To overcome these limitations, researchers are evolving gesture interactions into physical interactions, thus, completely mimicking interactions performed in the real world, into AR/MR environments. Several studies have explored the easiness and naturalness of user defined gesture interaction; For example, Weichel *et al.* proposed a MR environment for personal fabrication using in-space gesture interaction [5], while Billingham *et al.* [11] facilitated freehand gesture-based interactions such as precision grasping to pick and move virtual objects by using depth information of the fingertips. Their work highlighted the importance of choosing gestures that reflect real world interactions.

### B. Freehand Grasping

In the real world, freehand grasping is defined as the physical manual grip between a human hand and a real object without utilisation of any wearable devices and is an intuitive interaction technique by nature due to the highly complex structure of the human hand. This complexity offers a unique interplay between the fingers that is not necessarily present in gesture-based interaction techniques. Within AR, freehand grasping is a subset of freehand interaction methods, and is defined as the virtual extension to its definition in the real world; it is the manual grip between a (real) user and a (virtual) object without the utilisation of wearable sensors.

Freehand grasping is the most common human way of communication and controlling objects in the real world, underpinning powerful interactions due to hands multiple degrees of freedom (*see sidebar on p. 12 (number 10)*). This interaction paradigm has been linked to ease of access and

naturalness in the literature due to the absence of constraints imposed by wearable devices and its potential in delivering natural, intuitive and effective interactions (*see sidebar on p. 12 (numbers 11-15) for examples*).

Recent literature in this field include the analysis of grasping in AR, stepping away from gesture interaction; freehand grasping in AR was presented in [4], where finger positions were detected using a depth camera to allow for direct interaction using grasping with virtual objects. A first analysis of the accuracy of freehand grasping of virtual objects and its fundamental problems in exocentric AR was also presented in [12]. Freehand interactions that included grasping were also adopted using marker-less tracking to assess upper motor impairments in stroke patients based on recommendations by members of the clinical community (*see sidebar on p. 12 (numbers 16 and 17)*), where the use of wearable sensors was not feasible due to their potential interference with natural body motion, and freehand grasping in this case offered a more meaningful assessment of upper limb impairments through a natural way of interaction with virtual objects that resembles interaction in real life. Improvements to freehand grasping accuracy using an enhanced visual feedback method in exocentric AR were also presented in [7].

### C. Drop Shadows as a Depth Cue

Shadows are a crucial depth cue for humans in perceiving the 3D world around them and are useful in understanding size and position of virtual objects, the geometry of the surrounding environment and geometry of occluding objects or bodies. Current research offers strong evidence for the wide use, applications and impact of shadows as a depth cue in AR environments. For example, Diaz *et al.* [10] recently evaluated the impact of different rendering effects such as shading, cast shadows, aerial perspective and texture on perceptual depth matching tasks, where users were instructed to match the position of various virtual shapes to real world targets along the z axis using an HMD (HoloLens). Their work illustrated that shadows consistently had the largest impact on depth estimation in comparison to the other rendering features, whereas other cues such as shading, aerial perspective and texture did not have a significant impact on depth estimation when used on their own, however still showed improvements when combined with other depth cues (such as shadows). Interestingly, Diaz *et al.* also highlighted that users preferred drop shadows more than cast shadows even though cast shadows were more realistic in comparison to real world shadows, this was attributed to the effective human tolerance of imperfections in shadows as suggested in early studies in visual perception (*see sidebar on p. 12 (numbers 18 and 19)*).

While there is evidence of visual cues improving depth estimation and perception in AR [9], [10], no work has addressed the influence of visual cues such as drop shadows on users' perceived usability and overall accuracy when grasping virtual objects in AR. This study is the first to analyse drop shadow use in AR freehand grasping to overcome the reported incorrect virtual object size and location problems

reported in [8] and long reach to grasp task completion times highlighted in [7].

### III. METHODS

This paper presents two studies to quantify the influence of drop shadows on grasp accuracy in exocentric Augmented Reality (AR), that is an AR environment where users can directly interact with virtual objects that are viewed by users through a feedback monitor, using two virtual object types (cubes and spheres). This study focuses on grasp type, the medium wrap grasp, that is defined in the literature (*see sidebar on p. 12 (numbers 20 and 21)*) as the most common and versatile grasp in daily tasks (see Fig. 1a). A between participants design is applied where 30 participants were divided between both conditions: no shadows (see Fig. 1b) and shadows (see Fig. 1c). This between subjects experiment design was chosen following the methodologies presented in [7], [8], thus supporting an easy comparison with prior work. Every permutation of position was randomly presented to participants to exclude potential learning effects. In total, each participant completed 27 object positions and 810 grasps per condition (27 trials  $\times$  15 participants  $\times$  2 virtual objects), with findings compared to test the influence of drop shadows on grasp accuracy. To represent the accuracy of a grasp independent of additional rendering, for both experiments, the baseline virtual object types which have not undergone complex rendering and represent a simple abstract shape are used (order of virtual objects was counter-balanced).

#### A. System Architecture

The system integrated the use of a Microsoft Kinect 2 and a SyncMasterX6<sup>1</sup> monitor. Experiments were developed in C++ using Kinect SDK. Autodesk Maya for modelling 3D objects, OpenCV for video processing and OpenGL for real time presentation of 3D virtual objects.

#### B. Environment

Users stood 2500mm away from the feedback monitor and Kinect2 sensor under controlled and constant lighting conditions (see Fig. 2c), following the recommendation of Kinect2 manufacturers<sup>2</sup>. The test room was lit by a 2700k (*warm white*) fluorescent with controlled external light source. The virtual object z planes were placed at 2000mm, 2200mm and 2400mm distances away from the sensor as in Fig. 2. The floor was covered flat white to clearly distinguish the drop shadows. Layer based occlusion was also enabled during grasping interaction using the real-time depth and skeletal information of the scene and hand joints coming from the Kinect2 sensor, allowing freehand movement behind, in front or simultaneously in front and behind virtual objects. Conditions of both experiments are shown in Fig. 2, where experiment 2 is a replication of experiment 1 with the addition of drop shadows. Experiment 1 was used as a baseline experiment.

<sup>1</sup>[www.samsung.com/us/support/owners/product/lcd-monitor-md230-series](http://www.samsung.com/us/support/owners/product/lcd-monitor-md230-series)

<sup>2</sup><https://support.xbox.com/en-GB/xbox-360/accessories/kinect-sensor-setup>

#### C. Metrics

Interactivity is defined in this work as the users' ability to interact with the virtual objects and precision as the level of control the user has when interacting. In this paper the measure of precision would be how accurately the user can grasp the virtual object, while we report on interactivity as the perceived system usability.

1) *Precision Metrics*: To assess the potential of drop shadows, this study performs a replication of the experiment presented in Al-Kalbani *et al.* [8]. We use the Grasp Aperture (*GAp*) and Grasp Displacement (*GDisp*) metrics defined in equations 1 and 2 from the grasp model presented in [8] (see Fig. 1a).

$$GAp = \sqrt{(P_x - B_x)^2 + (P_y - B_y)^2 + (P_z - B_z)^2} \quad (1)$$

$$GDisp = \begin{pmatrix} \left( \frac{gmp_x + palm_x}{2} \right) - omp_x, \\ \left( \frac{gmp_y + palm_y}{2} \right) - omp_y, \\ \left( \frac{gmp_z + palm_z}{2} \right) - omp_z \end{pmatrix} \quad (2)$$

Where *GAp* is the distance between the index and the thumb fingers in the x, y and z axes, and  $P_x$ ,  $P_y$  and  $P_z$  are the co-ordinates of the index finger tip, and  $B_x$ ,  $B_y$  and  $B_z$  are co-ordinates of the thumb tip. *GDisp* is the distance between the grasp middle point (*gmp*) and the virtual object middle point (*omp*) in the x, y, and z axes (see Fig. 1a). Index, palm and thumb positions were obtained from the Kinect2.

Additionally, *completion time* was defined as the time it took to complete the grasping of the object. It was measured as the time in seconds from the auditory cue that marked the start of every trial up to when the participant informed the test coordinator that he/she was satisfied with the grasp performed.

2) *Interaction Metrics*: Interaction metrics are defined as the subjective metrics obtained from users using observation, the System Usability Scale (SUS) (*see sidebar on p. 12 (number 22) for more on the SUS questionnaire*) and a post-test questionnaire.

#### D. Statistical Analysis

Due to the repeated measures design of this study and the format of the data collected being non-parametric and not normally distributed, statistical significance between the two independent groups in this study (drop shadows and no drop shadows) is tested using a non-parametric Wilcoxon signed-rank test with an alpha of 1%.

#### E. Hypotheses

Following the methodology defined in this paper and the fact that there is no previous literature investigating the impact of drop shadows on virtual object grasping accuracy and task completion times, we propose the following null hypotheses:

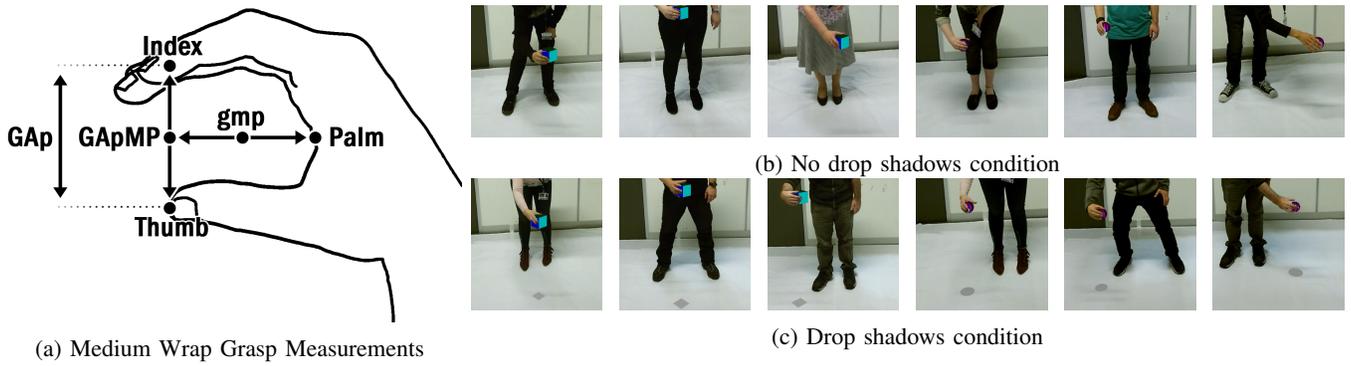


Fig. 1: 1a: displays grasp measurements required for Grasp Aperture ( $GAP$ ) and grasp middle point ( $gmp$ ). 1b: shows representative examples of the no shadows condition. 1c: shows representative examples of the shadows condition.

- **H<sub>1</sub>**: Adding drop shadows in freehand grasping of virtual objects has no effect on a) grasp aperture and b) grasp displacement
- **H<sub>2</sub>**: Adding drop shadows in freehand grasping of virtual objects has no effect on task completion time

#### F. Participants

30 right-handed participants ranged in age from 22 to 56 ( $M = 30.93$  yrs,  $SD = 8.48$  yrs), in arm length from 480mm to 660mm ( $M = 540.59$ mm,  $SD = 40.08$ mm), in hand size from 130mm to 210mm ( $M = 185.20$ mm,  $SD = 14.67$ mm), in height from 1558mm to 1940mm ( $M = 1729.83$ mm,  $SD = 84.45$ mm) and 8 were female and 22 male. Taking into account balance in hand size, arm length, gender, age and height, participants were separated into two groups of 15 for each condition drop shadows and no drop shadows.

#### G. Protocol

Participants underwent initial training of the medium wrap grasp on real and virtual objects and were given time to familiarise themselves with the environment in this study. The test coordinator explained the procedure and participants could rest before the presentation of every new object position. Each experiment was formed of a 5 minutes training/instruction session, 10 minutes of grasping a cuboid object, 5 minutes break and 10 minutes of grasping a spherical object (see Figures 1b and 1c).

After completing the test, participants were asked to fill in two usability questionnaires and a set of questions regarding their interaction with the system. The usability of the system was evaluated by a user satisfaction test based on the System Usability Scale (SUS). This questionnaire consists of 10 items, which were evaluated by using a Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Through feedback from the two questionnaires, the ease of use and usability of drop shadows is evaluated.

#### H. Procedure

For the two experiments (with and without drop shadows), participants were instructed to accurately locate and match

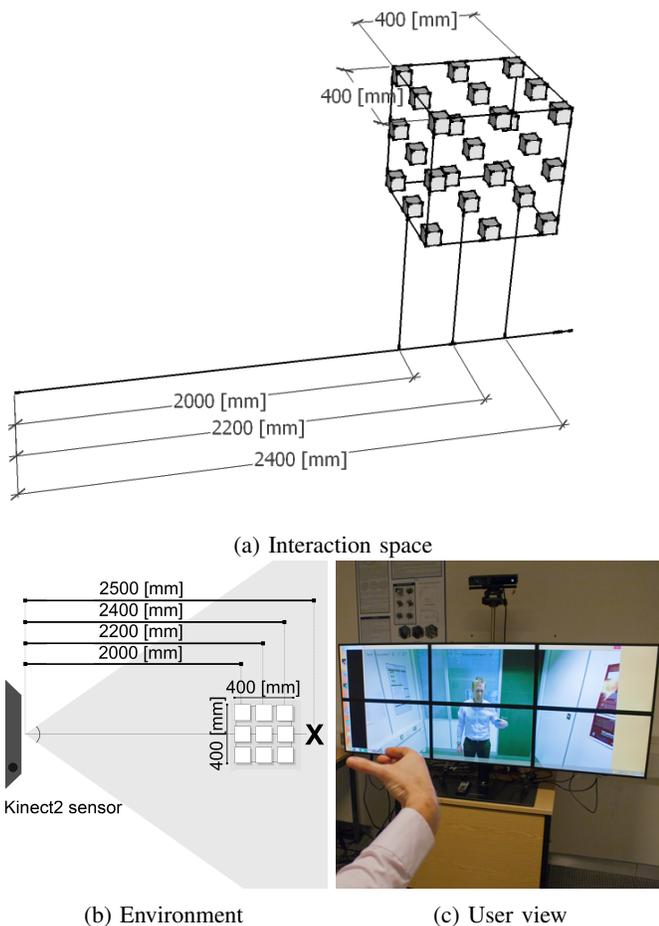


Fig. 2: Environment and interaction space for both conditions drop shadows and no shadows. Objects were positioned in 27 positions away from the Kinect2 sensor

their grasp aperture and position to the size and location of the virtual object considering time and accuracy using a medium wrap grasp (see Fig. 1a). 27 different object locations were used (see Fig. 2a), covering a working range of 400mm from participants. Following the results presented in [8] the object sizes with the most accurate user results in terms of matching Grasp Aperture ( $GAp$ ) to object size, 80mm for cubes and 70mm for spheres, are used for this study and were unchanged throughout the two experiments.

Before interaction, an object (cube or sphere) appeared to participants on the feedback monitor, each object had 27 different positions. A countdown of 5 seconds followed by an auditory cue was used as an indicator for participants to start grasping the object. During the experiment, all participants were instructed to verbally inform the test coordinator that they are satisfied with the grasp they have performed, and maintain the grasp for 5 seconds while the measurements of finger positions (see Fig. 1a) are captured by the Kinect2 and stored.

#### IV. RESULTS

##### A. Precision Metrics - Grasp Aperture ( $GAp$ )

###### 1) Results:

- **2000mm Z plane:** Statistically significant difference in  $GAp$  between the drop shadows and no drop shadows conditions in different positions was found for cubes ( $Z = 1.41 \times 10^7$ ,  $p < 0.01$ ) and spheres ( $Z = 1.71 \times 10^7$ ,  $p < 0.01$ )
- **2200mm Z plane:** Statistically significant difference in  $GAp$  between the drop shadows and no drop shadows conditions in different positions was found for grasping cubes ( $Z = 1.53 \times 10^7$ ,  $p < 0.01$ ) and spheres ( $Z = 1.60 \times 10^7$ ,  $p < 0.01$ )
- **2400mm Z plane:** Statistically significant difference in  $GAp$  between the drop shadows and no drop shadows conditions in different positions was found for cubes ( $Z = 1.40 \times 10^7$ ,  $p < 0.01$ ) but not spheres ( $Z = 1.43 \times 10^7$ ,  $p > 0.01$ )

2) *Analysis:* Users overestimated object size in both the drop shadows and no drop shadows conditions across all z planes. Despite the constant size of virtual objects in this study a high variation in Grasp Aperture ( $GAp$ ) was shown by users across all z planes. For cubes,  $GAp$  ranged from 69.49mm  $\pm$  17.75mm to 93.77mm  $\pm$  20.67mm in the no drop shadows condition, whereas a wider range from 61.89mm  $\pm$  17.75mm to 99.45mm  $\pm$  20.67mm was shown in the drop shadows condition. In contrast for spheres,  $GAp$  ranged from 65.74mm  $\pm$  14.15mm to 87.69mm  $\pm$  23.38mm in the drop shadows condition, whereas a narrower range from 59.78mm  $\pm$  21.45mm to 84.17mm  $\pm$  21.04mm was shown in the drop shadows condition.

This high variation in Grasp Aperture ( $GAp$ ) is attributed to the lack of tactile feedback in the hand from the object and is comparable to previous studies in the literature [7], [8].

Users presented the highest accuracy in terms of matching their  $GAp$  to object sizes in the 2200mm z plane for cubes

in both conditions (drop shadows and no drop shadows), with a mean overestimation of 2.26mm  $\pm$  24.43mm in the drop shadows condition, and a mean overestimation of 4.77mm  $\pm$  27.75mm in the no drop shadows condition (see Fig. 3). Similarly for spheres, users also presented the highest accuracy in  $GAp$  in the 2200mm z plane in the drop shadows condition with a mean overestimation of 3.67mm  $\pm$  19.95mm, and in the 2400mm z plane in the no drop shadows condition with a mean overestimation of 7.19mm  $\pm$  22.82mm. However, this was comparable to the mean overestimation in the 2200mm z plane for spheres (7.37mm  $\pm$  -48.33mm) as the difference between the two planes was not significant ( $p > 0.01$ ). This preference for the middle z plane is in alignment with findings reported in previous studies [7], [8].

##### B. Precision Metrics - Completion time

###### 1) Results:

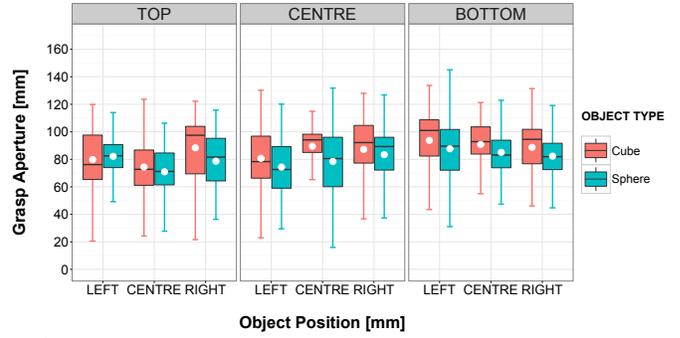
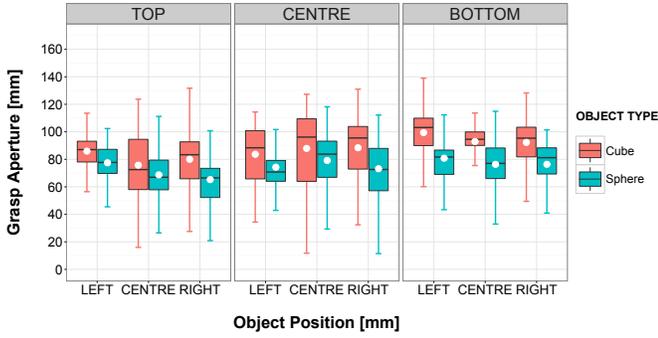
- **2000mm Z plane:** Statistically significant difference in completion time between the drop shadows and no drop shadows conditions in different positions was found for cubes ( $Z = 1.72 \times 10^7$ ,  $p < 0.01$ ) and spheres ( $Z = 1.55 \times 10^7$ ,  $p < 0.01$ )
- **2200mm Z plane:** Statistically significant difference in completion time between the drop shadows and no drop shadows conditions in different positions was found for cubes ( $Z = 1.63 \times 10^7$ ,  $p < 0.01$ ) and spheres ( $Z = 1.37 \times 10^7$ ,  $p < 0.01$ )
- **2400mm Z plane:** Statistically significant difference in completion time between the drop shadows and no drop shadows conditions in different positions was found for cubes ( $Z = 1.68 \times 10^7$ ,  $p < 0.01$ ) and spheres ( $Z = 1.29 \times 10^7$ ,  $p < 0.01$ )

2) *Analysis:* As shown in Fig. 4, mean completion time was lower across the positions in the drop shadows condition compared to the no shadows condition.

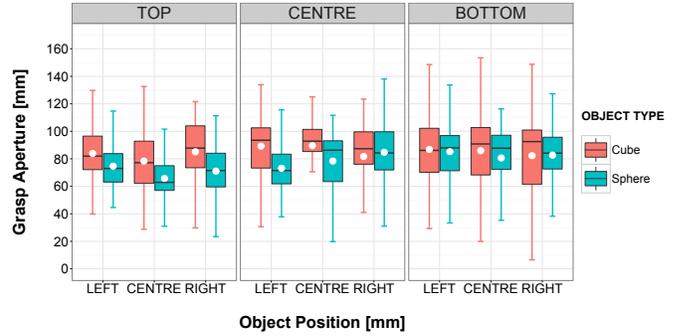
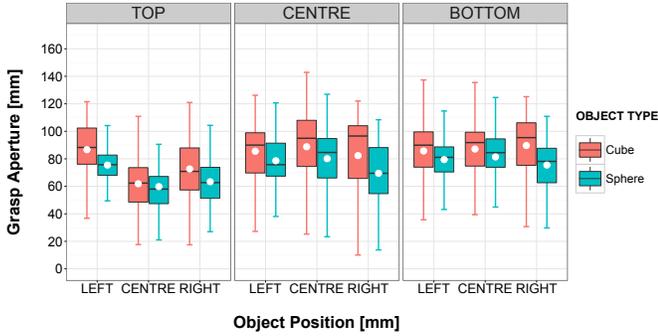
Participants consistently presented shortest completion times in the 2200mm z plane for cubes and spheres in both conditions (drop shadows and no drop shadows) across all positions. For cubes, mean completion time in this particular z plane was 7.95s  $\pm$  3.82s in the drop shadows condition and 9.01s  $\pm$  4.31s in the no drop shadows condition. For spheres, mean completion time was 7.62s  $\pm$  2.96s in the drop shadows condition, and 7.70s  $\pm$  3.96s in the no drop shadows condition. This preference for the 2200mm z plane can be attributed to the convenience of its position in relation to the bio-mechanical reach of users, as this plane is not as physically demanding as the furthest and closest z planes.

Participants also showed shortest mean completion times in the Centre positions alongside the y axis across all z planes for cubes with a mean completion time of 7.74s  $\pm$  3.69s in the drop shadows condition, and in Top positions in the no drop shadows condition with a longer mean completion time of 9.51s  $\pm$  5.60s. For spheres, participants showed the shortest mean completion time across all z planes in the Top positions with a mean completion time of 7.64s  $\pm$  3.39s in the drop shadows condition, and in the Centre positions in the no drop

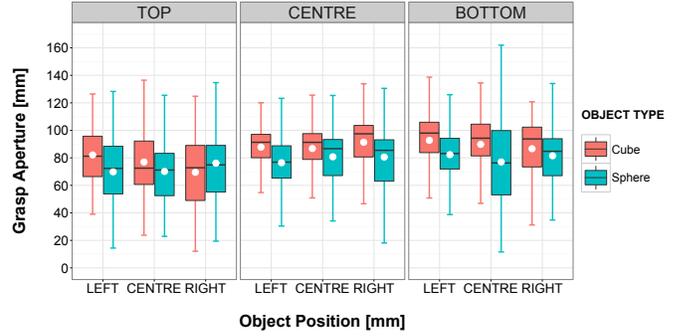
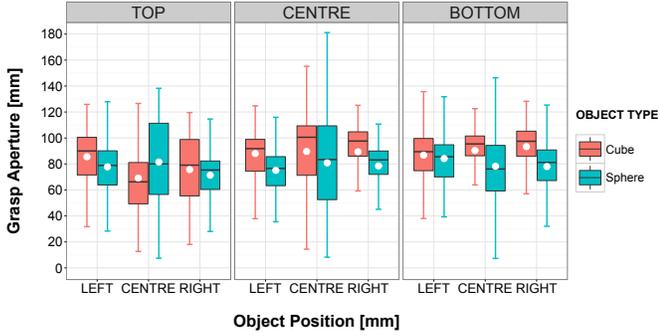
### 2000mm z Plane



### 2200mm z Plane



### 2400mm z Plane



(a) Drop shadows

(b) No drop shadows

Fig. 3: Grasp aperture ( $GA_p$ ) for different object positions in the three z planes in this experiment (2000mm, 2200mm and 2400mm). 3a: Drop shadows used. 3b: No drop shadows used. White points on boxplots indicate the mean  $GA_p$  across all participants for each size. Whiskers represent the highest and lowest values within 1.5 and 3.0 times the interquartile range

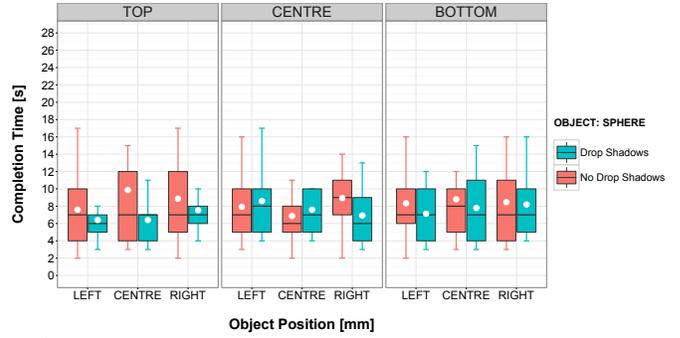
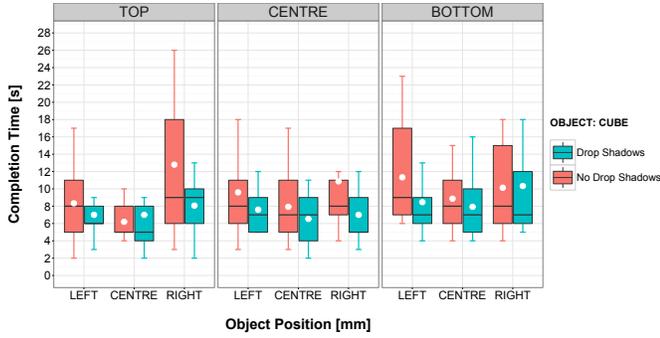
shadows condition with a shorter mean completion time of  $7.52s \pm 3.50s$  in the no drop shadows condition.

Alongside the x axis, participants showed the shortest completion time in the Centre positions for cubes with mean completion times of  $7.79s \pm 3.88s$  in the drop shadows condition and  $8.99s \pm 4.60s$  in the no drop shadows condition. Centre positions also showed the shortest completion times for spheres across all conditions (drop shadows and no drop shadows), with a mean completion time of  $7.63s \pm 3.20s$  in the drop shadows condition and  $7.87s \pm 5.69s$  in the no drop shadows condition. These interaction regions across the x (Centre and Right) and y (Centre and Top) axis are usually preferred by users as they are easily accessible using grasp

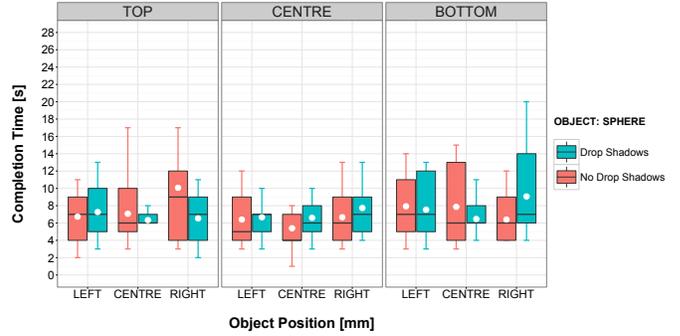
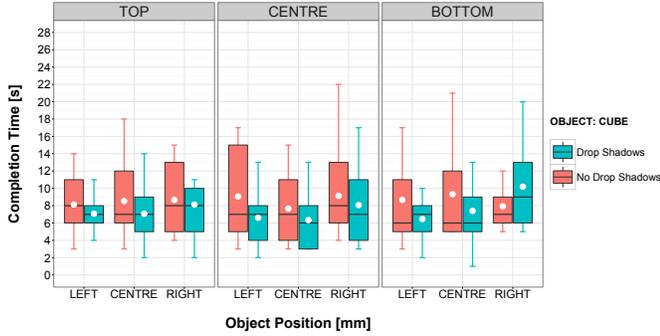
to reach movements and feel more natural to interact within using the right dominant hand.

In terms of virtual object type, users showed less task completion time in grasping spheres than cubes and this was consistent in all z planes. This shows that users required less time to locate spheres, this was evident during the study as users informally expressed that spheres are easier to interact with than cubes. The lesser need to adjust the posture of a grasp after locating spheres due to their smaller size in comparison to cubes and shape attributes could have potentially also influenced completion time spent by users when grasping spheres.

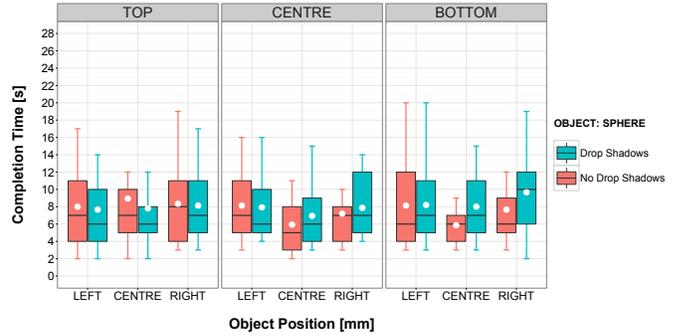
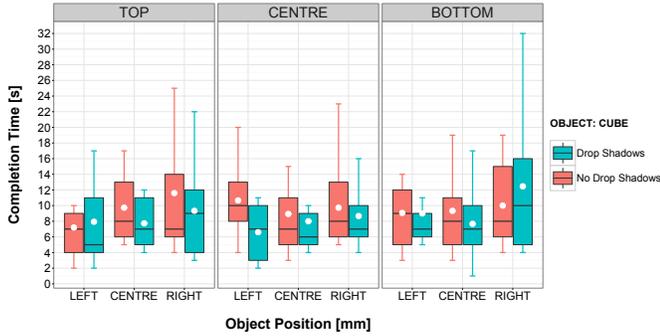
## 2000mm z Plane



## 2200mm z Plane



## 2400mm z Plane



(a) Cubes

(b) Spheres

Fig. 4: Task completion times for different object positions in the three z planes in this study (2000mm, 2200mm and 2400mm). 4a: Cubes. 4b: Spheres. White points on boxplots indicate the mean completion time across all participants for each size. Whiskers represent the highest and lowest values within 1.5 and 3.0 times the interquartile range

### C. Precision Metrics - Grasp Displacement ( $GDisp$ )

#### 1) Results:

- 2000mm Z plane:** Statistically significant difference in between the drop shadows and no drop shadows conditions in different positions were found in  $GDisp_x$  for cubes ( $Z = 1.26 \times 10^7$ ,  $p < 0.01$ ) and spheres ( $Z = 1.33 \times 10^7$ ,  $p < 0.01$ ), in  $GDisp_y$  for spheres ( $Z = 1.30 \times 10^7$ ,  $p < 0.01$ ) and in  $GDisp_z$  for cubes ( $Z = 1.35 \times 10^7$ ,  $p > 0.01$ ) and spheres ( $Z = 1.30 \times 10^7$ ,  $p < 0.01$ ). No significant differences were found in  $GDisp_y$  for cubes ( $Z = 1.42 \times 10^7$ ,  $p > 0.01$ ) between conditions in this plane.

- 2200mm Z plane:** Statistically significant difference in between the drop shadows and no drop shadows conditions in different positions were found in  $GDisp_x$  for cubes ( $Z = 1.29 \times 10^7$ ,  $p < 0.01$ ) and spheres ( $Z = 1.35 \times 10^7$ ,  $p < 0.01$ ), in  $GDisp_y$  for cubes ( $Z = 1.39 \times 10^7$ ,  $p > 0.01$ ) and spheres ( $Z = 1.38 \times 10^7$ ,  $p < 0.01$ ), and in  $GDisp_z$  for cubes ( $Z = 1.38 \times 10^7$ ,  $p > 0.01$ ) and spheres ( $Z = 1.38 \times 10^7$ ,  $p < 0.01$ ).
- 2400mm Z plane:** Statistically significant difference in between the drop shadows and no drop shadows conditions in different positions were found in  $GDisp_x$  for cubes ( $Z = 1.38 \times 10^7$ ,  $p < 0.01$ ) and spheres ( $Z = 1.25 \times 10^7$ ,  $p < 0.01$ ), in  $GDisp_y$  for spheres ( $Z = 1.39$

$\times 10^7$ ,  $p < 0.01$ ) and in  $GDisp_z$  for spheres ( $Z = 1.40 \times 10^7$ ,  $p < 0.01$ ). No significant differences were found for cubes in  $GDisp_y$  ( $Z = 1.49 \times 10^7$ ,  $p > 0.01$ ) and  $GDisp_z$  ( $Z = 1.47 \times 10^7$ ,  $p > 0.01$ ) between conditions in this plane.

2) *Analysis*: Users showed a mean  $GDisp_x$  of  $36.39\text{mm} \pm 53.57\text{mm}$  in the no drop shadows condition across all positions, and a higher mean of  $43.47\text{mm} \pm 35.36\text{mm}$  in the drop shadows condition. For spheres, users showed a mean  $GDisp_x$  of  $20.92\text{mm} \pm 36.83\text{mm}$  in the no drop shadows condition across all positions, and again a higher mean of  $27.43\text{mm} \pm 32.25\text{mm}$  in the drop shadows condition.

Users showed higher  $GDisp_x$  in the drop shadows condition, adding drop shadows as a depth cue increased the range of  $GDisp_x$ . For cubes  $GDisp_x$  ranged from  $4.89\text{mm} \pm 33.18\text{mm}$  to  $76.10\text{mm} \pm 33.73\text{mm}$  in the no drop shadows condition, and a higher range from  $12.87\text{mm} \pm 24.68\text{mm}$  to  $96.25\text{mm} \pm 55.38\text{mm}$  in the drop shadows condition. Similarly for spheres,  $GDisp_x$  ranged in the no drop shadows conditions from  $0.59\text{mm} \pm 24.48\text{mm}$  to  $57.39\text{mm} \pm 72.55\text{mm}$ , and from  $0.75\text{mm} \pm 13.99\text{mm}$  to  $74.32\text{mm} \pm 26.65\text{mm}$  in the drop shadows condition (see Fig. 5).

As shown in Fig. 5, users presented the lowest  $GDisp_x$  in the 2000mm z plane for cubes in both conditions (drop shadows and no drop shadows), with a mean  $GDisp_x$  of  $36.88\text{mm} \pm 28.11\text{mm}$  in the drop shadows condition, and a mean of  $27.40\text{mm} \pm 66.60\text{mm}$  in the no drop shadows condition across all positions. Similarly for spheres, users also presented the lowest  $GDisp_x$  in the 2000mm z plane in both conditions (drop shadows and no drop shadows) with a mean  $GDisp_x$  of  $20.47\text{mm} \pm 21.97\text{mm}$  across all positions in the drop shadows condition, and a mean  $GDisp_x$  of  $7.19\text{mm} \pm 22.82\text{mm}$  in the no drop shadows condition. This preference for the furthest z plane from users (2000mm) is potentially attributed to the spatial position of this particular plane being at the extremity of the average arm reach of users, thus users, in this case, were accurate in terms of grasp placement in the x axis as the environment design did not allow the arm reach to extend beyond this plane, thus reducing the amount of potential errors in grasp placement in the x axis.

Lowest mean  $GDisp_x$  was shown by users in the Top positions alongside the y axis across all z planes for cubes with a mean  $GDisp_x$  of  $36.38\text{mm} \pm 25.03\text{mm}$  in the drop shadows condition, and a mean  $GDisp_x$  of  $30.94\text{mm} \pm 36.24\text{mm}$  in the no drop shadows condition. Users also presented lowest  $GDisp_x$  in Top positions for spheres with a mean  $GDisp_x$  of  $22.52\text{mm} \pm 24.48\text{mm}$  in the drop shadows condition, and a mean  $GDisp_x$  of  $16.60\text{mm} \pm 35.69\text{mm}$  in the no drop shadows condition.

Alongside the x axis, users showed the lowest  $GDisp_x$  in Right positions across all z planes for cubes with a mean  $GDisp_x$  of  $21.97\text{mm} \pm 24.10\text{mm}$  in the drop shadows and a mean  $GDisp_x$  of  $10.22\text{mm} \pm 65.61\text{mm}$  in the no drop shadows condition. Right positions also provided the lowest  $GDisp_x$  in grasping spheres with a mean  $GDisp_x$  of  $5.89\text{mm} \pm 19.47\text{mm}$  in the drop shadows condition, and a mean

$GDisp_x$  of  $1.95\text{mm} \pm 26.64\text{mm}$  in the no drop shadows condition.

$GDisp_y$  presented the lowest displacement in this study for both objects and conditions (drop shadows and no drop shadows) under test. Users grasping cubes showed a mean  $GDisp_y$  of  $9.12\text{mm} \pm 41.94\text{mm}$  in the no drop shadows condition and a lower mean of  $8.97\text{mm} \pm 28.43\text{mm}$  in the drop shadows condition. In contrast for spheres, a lower mean  $GDisp_y$  of  $5.22\text{mm} \pm 47.49\text{mm}$  was found in the no drop shadows condition and a higher mean of  $7.07\text{mm} \pm 30.38\text{mm}$  in the drop shadows condition.

Similar to  $GDisp_x$ , users presented lowest  $GDisp_y$  in the 2000mm z plane for cubes in the drop shadows condition with a mean  $GDisp_y$  of  $5.21\text{mm} \pm 20.43\text{mm}$ , and in the 2200mm z plane in the no drop shadows condition with a higher mean  $GDisp_y$  of  $7.07\text{mm} \pm 33.48\text{mm}$  in this particular z plane. For spheres, users presented lowest mean  $GDisp_y$  in the 2000mm z plane for both conditions, with a mean of  $4.24\text{mm} \pm 21.64\text{mm}$  in the drop shadows condition, and a lower  $GDisp_y$  mean of  $2.72\text{mm} \pm 32.81\text{mm}$  in the no drop shadows condition (see Fig. 5). Like  $GDisp_x$ , this plane shows lowest  $GDisp_y$  due to its spatial position at the extremity of the arm reach, thus limiting any potential errors by users along the y axis.

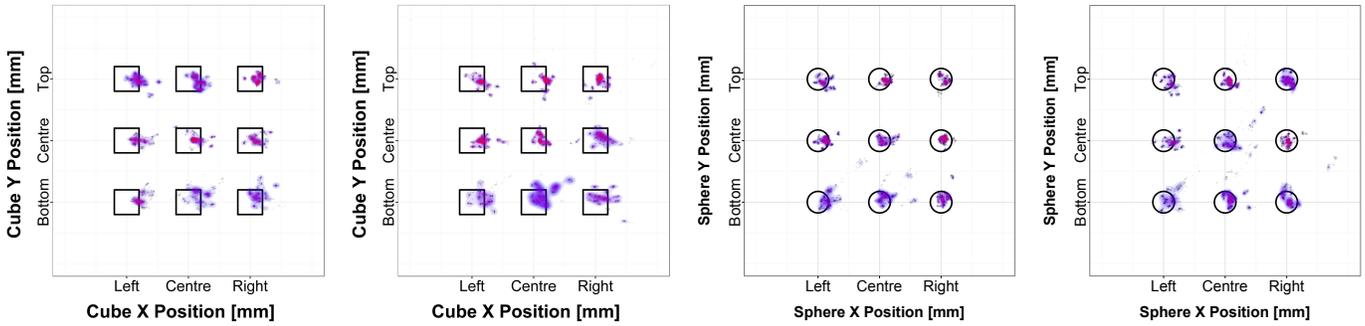
Lowest mean  $GDisp_y$  was shown across all z planes by users in Centre positions alongside the y axis for cubes in both conditions (drop shadows and no drop shadows) with a mean of  $7.24\text{mm} \pm 17.60\text{mm}$  in the drop shadows condition, and a higher mean of  $10.67\text{mm} \pm 37.45\text{mm}$  in the no drop shadows condition. Centre positions also showed lowest  $GDisp_y$  for spheres in both conditions with a mean of  $3.89\text{mm} \pm 15.91\text{mm}$  in the drop shadows condition and a lower mean of  $2.12\text{mm} \pm 21.79\text{mm}$  in the no drop shadows condition. This preference for Centre and Right positions is expected and in alignment with findings for  $GDisp_x$  as all users in this study were right-handed.

However alongside the x axis, users showed the lowest mean  $GDisp_y$  in Left positions for grasping cubes in both conditions (drop shadows and no drop shadows), with a mean  $GDisp_y$  of  $6.55\text{mm} \pm 27.59\text{mm}$  in the drop shadows condition, and a mean  $GDisp_y$  of  $3.80\text{mm} \pm 42.22\text{mm}$  in the no drop shadows condition. Users were also more accurate in Left positions in grasping spheres in both the drop shadows and no drop shadows conditions, with a mean  $GDisp_y$  of  $3.24\text{mm} \pm 33.35\text{mm}$  in the drop shadows condition, and a mean  $GDisp_y$  of  $3.52\text{mm} \pm 63.12\text{mm}$  in the no drop shadows condition. This is surprising as all users in this study were right handed, however, this highlights how highly varied  $GDisp_y$  can be between individual users, as Right positions were consistently second best in terms of low mean  $GDisp_y$ .

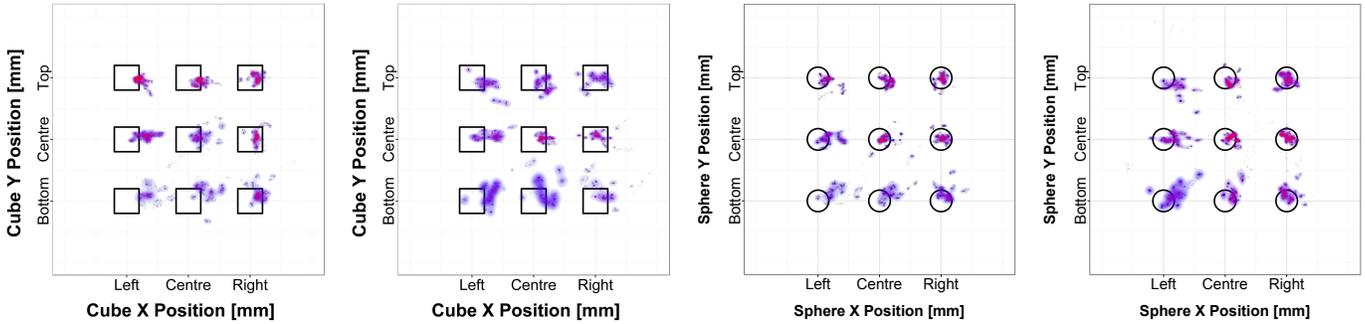
Adding drop shadows as a depth cue alongside occlusion generally reduced the mean  $GDisp_z$ , where drop shadows outperformed the no drop shadows condition in the majority of trials (35 out of the 54 individual trials in this test), with 27 of the 35 showing statistical significance ( $p < 0.01$ ).

Adding drop shadows also reduced the amount of variation

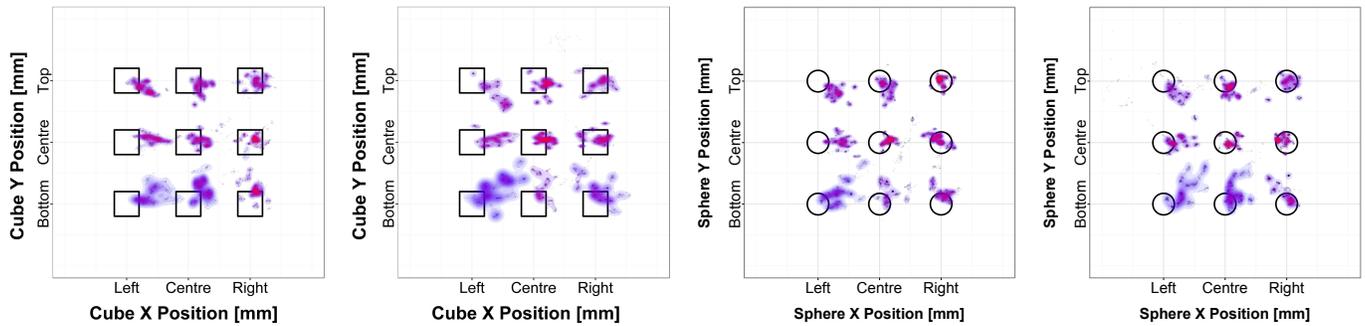
2000mm z Plane



2200mm z Plane



2400mm z Plane



(a) Drop shadows - Cubes (b) No drop shadows - Cubes (c) Drop shadows - Spheres (d) No drop shadows - Spheres

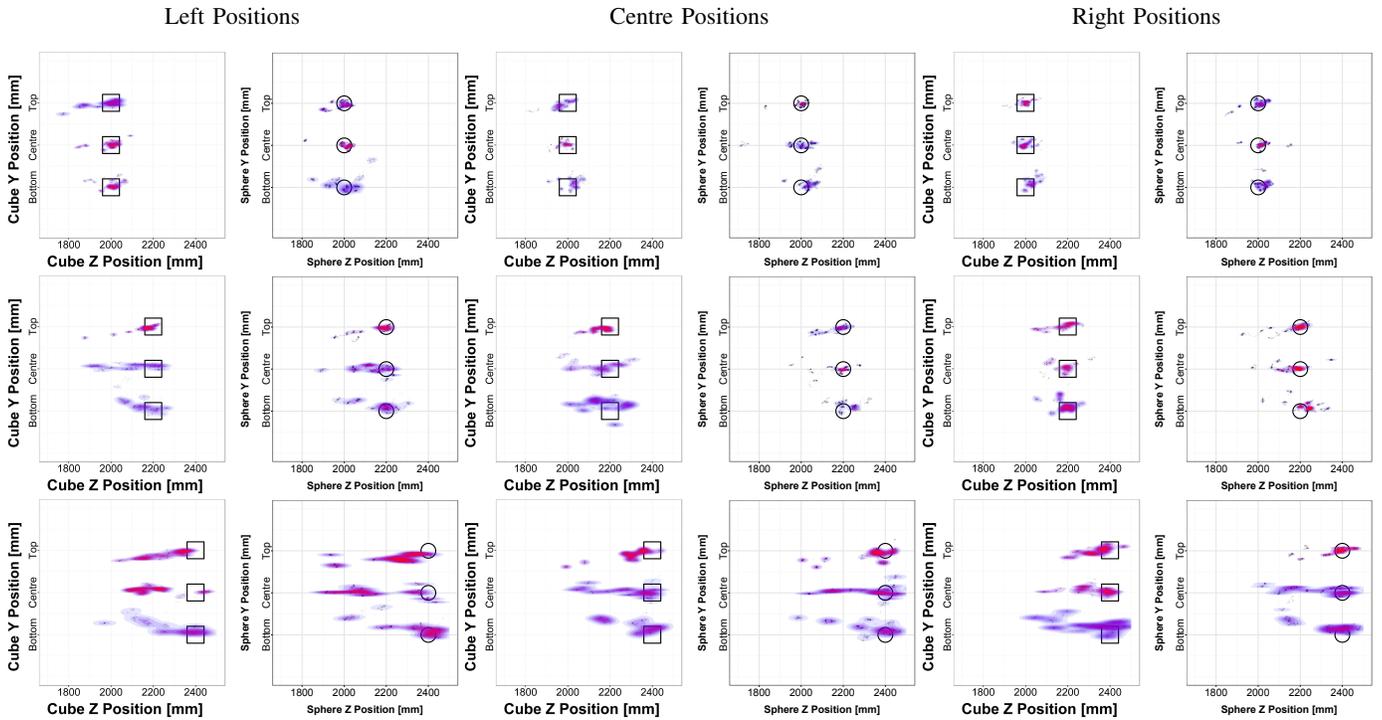
Fig. 5: Grasp middle point ( $gmp$ ) placement in the x and y axes for cubes and spheres of all participants in 3 z planes (2000mm - 2200mm - 2400mm). 5a and 5c: Drop shadows used. 5b and 5d: No Drop shadows used. Density heat maps indicate  $gmp$  placement across participants (red indicates higher density)

of users in depth estimation, thus even if users were not accurate in depth estimation, they were more confident in their grasp placement along the z axis and were more clustered around the object as seen in Figures 6a and 6b.

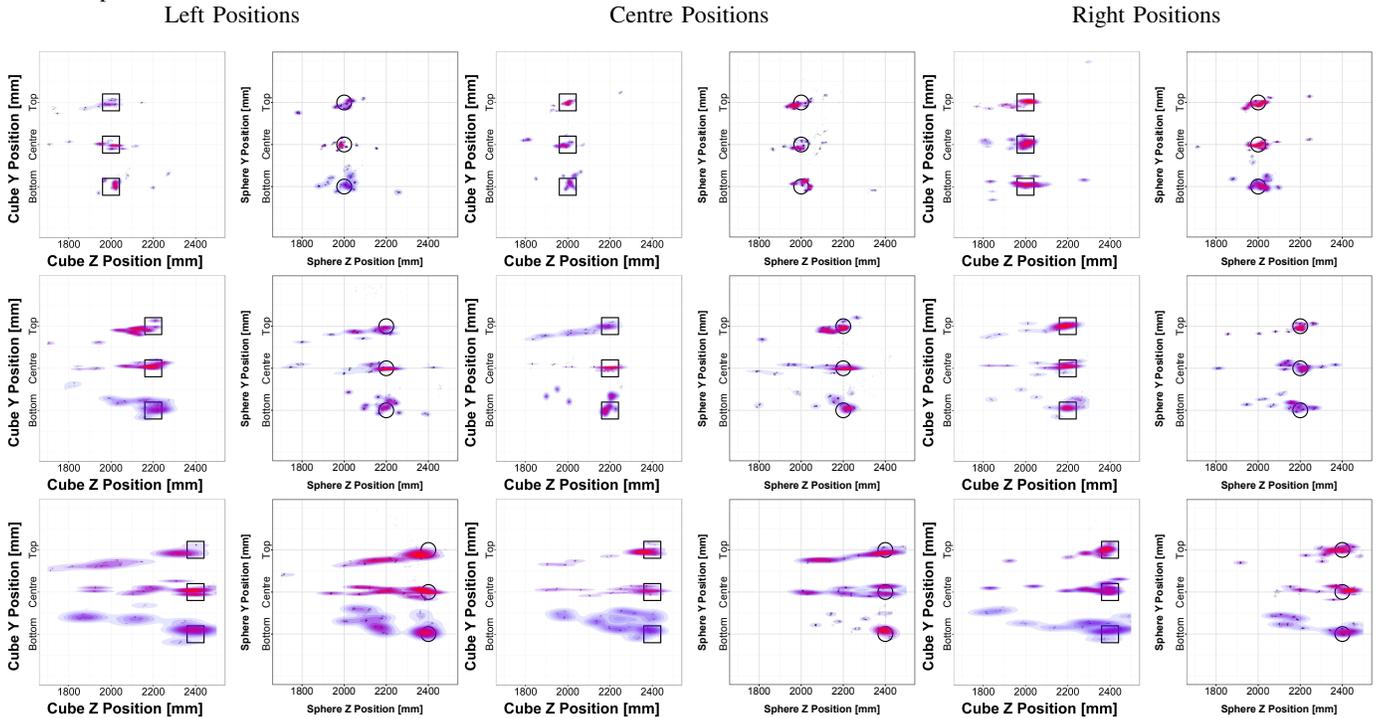
Users showed lowest mean  $GDisp_z$  in the 2000mm z plane for grasping cubes in both conditions (drop shadows and no drop shadows), where users showed a mean  $GDisp_z$  of  $-7.10\text{mm} \pm 56.91\text{mm}$  in the drop shadows condition, and a mean  $GDisp_z$  of  $-17.09\text{mm} \pm 100.62\text{mm}$  in the no drop shadows condition. Similarly for grasping spheres, users showed a mean  $GDisp_z$  of  $4.11\text{mm} \pm 54.35\text{mm}$  in the drop shadows condition, and a mean  $GDisp_z$  of  $-2.05\text{mm} \pm 91.69\text{mm}$  in the no drop shadows condition. This preference for the furthest

z plane from users (2000mm) can again be attributed to the spatial position of the plane, that limits the amount of false depth estimation due to its position at the extremity of the average arm reach of users.

Users presented the lowest mean  $GDisp_z$  in the Bottom positions alongside the y axis across all z planes in grasping cubes in both conditions (drop shadows and no drop shadows), with a mean  $GDisp_z$  of  $-38.98\text{mm} \pm 99.40\text{mm}$  in the drop shadows condition, and a mean  $GDisp_z$  of  $-57.20\text{mm} \pm 155.95\text{mm}$  in the no drop shadows condition. Bottom positions also showed lowest mean  $GDisp_z$  for spheres in both the drop shadows and no drop shadows conditions, with a mean of  $-22.28\text{mm} \pm 103.04\text{mm}$  in the drop shadows condition, and



(a) Grasp middle point (*gmp*) placement in the z axis for cubes and spheres of all participants in 3 z planes (2000mm - 2200mm - 2400mm) in the drop shadows condition



(b) Grasp middle point (*gmp*) placement in the z axis for cubes and spheres of all participants in 3 z planes (2000mm - 2200mm - 2400mm) in the no drop shadows condition

Fig. 6: Grasp middle point (*gmp*) placement in the z axis for cubes (black squares) and spheres (black circles) of all participants in 3 z planes (starting from top row in the order: 2000mm - 2200mm - 2400mm) using 6a: drop shadows and 6b: No drop shadows. Density heat maps indicate *gmp* placement across participants (red indicates higher density)

a mean  $GDisp_z$  of  $-40.29\text{mm} \pm 126.65\text{mm}$  in the no drop shadows condition. This is surprising as the Bottom positions consistently presented the highest user errors in terms of  $GAp$ ,  $GDisp_x$ ,  $GDisp_y$ , and presented the longest completion times for both conditions and object types. This can potentially be attributed to the position of the arm during grasping, wherein Bottom positions the position of the arm may not be obstructing the view of the user of the feedback monitor.

Alongside the x axis (see Figures 6a and 6b), users showed the lowest mean  $GDisp_z$  in Right positions across all z planes for grasping cubes in both conditions (drop shadows and no drop shadows), with a mean  $GDisp_z$  of  $-32.79\text{mm} \pm 79.92\text{mm}$  in the drop shadows condition and a mean  $GDisp_z$  of  $-63.36\text{mm} \pm 140.99\text{mm}$  in the no drop shadows condition. Users also showed lowest mean  $GDisp_z$  in Right positions for grasping spheres in both conditions (drop shadows and no drop shadows), with a mean  $GDisp_z$  of  $-30.54\text{mm} \pm 95.31\text{mm}$  in the drop shadows condition, and a mean  $GDisp_z$  of  $-37.77\text{mm} \pm 124.95\text{mm}$  in the no drop shadows condition.

In terms of virtual object type, adding drop shadows reduced  $GDisp_z$  for both objects in all z planes in the majority of positions. Drop shadows also reduced variation in depth estimation for both objects in all z planes, thus users were more confident and more accurate in their  $gmp$  placement along the z axis with the presence of drop shadows. This shows that drop shadows can aid in depth perception of virtual objects when integrated in an AR environment as in this study, especially in the z axis where the real depth of objects cannot be perceived using only the feedback monitor and relying on occlusion alone as a depth cue can still result in false depth perception as shown previously in [8].

#### D. Interaction Metrics - Usability

SUS average score for the use of drop shadows in this study was 81.16 (SD = 11.56). This rating of drop shadows was higher than the rating of the no drop shadows condition (78.17 (SD = 14.13)) and is classified as “GOOD and highly acceptable” in the SUS rankings. No significant results between conditions were found for the SUS scores.

In terms of easiness of tasks, the no drop shadows condition scored marginally higher (3.93 / 5.00) than the drop shadows condition (3.80 / 5.00). Users had lower task completion times in the drop shadows condition, potentially due to the fact that grasping without drop shadows in this study is not challenged or corrected. In contrast with the addition of drop shadows interaction can be corrected using this additional depth cue, along with the additional cognitive load that is present.

Users have also indicated that changes in virtual object position between different z planes and in the x and y axes were more perceptible with the use of drop shadows (4.80 / 5.00) than without them (3.87 / 5.00). In addition, 9 out of 15 users (60.0%) have indicated that they relied on both drop shadows and occlusion that is implemented in the baseline setup for this work. This is in alignment with current research that states that depth cues can be more effective when used alongside shadows [10].

#### E. Hypothesis - Reviewed

**H<sub>1</sub>:Rejected** as statistically significant results were found for drop shadows condition showing that adding drop shadows in grasping virtual objects has a significant effect on  $GAp$ ,  $GDisp$  in all axes (x, y and z) and task completion time.

**H<sub>2</sub>:Rejected** as statistically significant results were found for drop shadows condition showing that adding drop shadows in grasping virtual objects has a significant effect on task completion time.

#### V. CONCLUSION AND DISCUSSION

This article presented a first study evaluating the use of drop shadows to assist in freehand grasping of virtual objects in AR. The impact of drop shadows on grasp accuracy was quantified in a user experiment under two primary conditions: drop shadows and no drop shadows using two virtual object types (cubes and spheres), where grasp accuracy in this study was measured using the proposed metrics in [8]; Grasp Aperture ( $GAp$ ) and Grasp Displacement ( $GDisp$ ). This study has also addressed the key problems found in previous literature [7], namely long task completion times and inaccurate estimation of virtual object size.

Findings in this study have illustrated that using drop shadows significantly reduces task completion times found in the literature [7] and also significantly improved depth estimation of virtual objects in the z axis for both virtual object types (cubes and spheres). This is in alignment with current research in AR that illustrate that drop shadows have the largest impact on depth estimation [10].

In terms of accuracy, users performed better in matching grasp aperture to object size in the no shadows condition than the drop shadows condition in the majority of positions. This indicates that adding drop shadows as a depth cue in freehand grasping can significantly influence grasp aperture accuracy, however not necessarily improve overall  $GAp$  estimation as users performed better without drop shadows in the majority of positions. This negative impact of drop shadows on matching to the object size can be attributed to additional cognitive load in the process of grasping, wherein the no drop shadows condition users were solely focused on  $GAp$  estimation. This effect was also present in the literature [7], where an additional depth cue (dual view) was used, where users did not show improvements in  $GAp$  using dual view visual feedback, possibly being more focused on correcting the spatial position of their grasp using the second view. This finding informs developers and designers of discrepancies in virtual object size estimation during grasping that present a fundamental problem for grasping interactions in AR using  $GAp$ .

Future work will consider different methods by which virtual object grasping can be improved in AR, especially focusing on improving virtual object size estimation using  $GAp$  as findings in this work have illustrated that drop shadows do not improve size estimation using  $GAp$ , thus remaining a problematic aspect of grasping interactions in AR. In addition, future work will explore methods that can be transferable to different AR interfaces, as the use of drop

shadows in AR applications is not always feasible due to the limited FOV that current AR systems have, thus users are not always guaranteed to be able to visualise the whole environment with drop shadows as implemented in this study. Future studies will also evaluate if results in this work are maintained with different virtual object shapes (i.e. real-world objects and complex shapes), and grasp types that extend beyond the medium wrap grasp used in this work.

Findings in this work build on previous AR approaches that resulted in more accurate depth judgements, and further emphasise the importance of drop shadows to assist users in interaction in AR, especially when accurate depth judgements are required such as manufacturing and maintenance applications. Finally, this analysis of drop shadows in AR can be valuable for applications where a more natural, accurate and usable grasping interaction with virtual objects is required.

#### **Additional Resources -**

1. M. Bal, "Virtual manufacturing laboratory experiences for distance learning courses in engineering technology," in *American Society for Engineering Education*. American Society for Engineering Education, 2012.
2. S.-H. Sun and L.-Z. Tsai, "Development of virtual training platform of injection molding machine based on vr technology," *The International Journal of Advanced Manufacturing Technology*, vol. 63, no. 5-8, pp. 609–620, 2012.
3. K. Li, M. Hall, P. Bermell-Garcia, J. Alcock, A. Tiwari, and M. González-Franco, "Measuring the learning effectiveness of serious gaming for training of complex manufacturing tasks," *Simulation & Gaming*, vol. 48, no. 6, pp. 770–790, 2017.
4. A. Antonaci, R. Klemke, and M. Specht, "Towards design patterns for augmented reality serious games," in *International Conference on Mobile and Contextual Learning*. Springer, 2015, pp. 273–282.
5. Y. Yao, P.-T. Chiu, and W.-T. Fu, "A gestural interface for practicing children's spatial skills," in *Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion*. ACM, 2017, pp. 43–47.
6. H. M. Hondori, M. Khademi, L. Dodakian, S. C. Cramer, and C. V. Lopes, "A spatial augmented reality rehab system for post-stroke hand rehabilitation," in *MMVR*, 2013, pp. 279–285.
7. H. Bai, L. Gao, J. El-Sana, and M. Billinghurst, "Markerless 3D gesture-based interaction for handheld Augmented Reality interfaces," *2013 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2013*, no. October 2013, pp. 0–5, 2013.
8. W. Jung and W. T. Woo, "Duplication based distance-free freehand virtual object manipulation," in *2017 International Symposium on Ubiquitous Virtual Reality (ISUVR)*. IEEE, 2017, pp. 10–13.
9. D. Datcu and S. Lukosch, "Free-hands interaction in augmented reality," in *Proceedings of the 1st symposium on Spatial user interaction*. ACM, 2013, pp. 33–40.
10. B. S. Santos, J. Cardoso, B. Q. Ferreira, C. Ferreira, and P. Dias, "Developing 3d freehand gesture-based interaction methods for virtual walkthroughs: Using an iterative approach," in *Handbook of Research*
11. S. Lin, H. F. Cheng, W. Li, Z. Huang, P. Hui, and C. Peylo, "Ubii: Physical world interaction through augmented reality," *IEEE Transactions on Mobile Computing*, vol. 16, no. 3, pp. 872–885, 2017.
12. T. Ni, D. A. Bowman, C. North, and R. P. McMahan, "Design and evaluation of freehand menu selection interfaces using tilt and pinch gestures," *International Journal of Human-Computer Studies*, vol. 69, no. 9, pp. 551–562, 2011.
13. J. P. Wachs, M. Kölsch, H. Stern, and Y. Edan, "Vision-based handgesture applications," *Communications of the ACM*, vol. 54, no. 2, pp. 60–71, 2011.
14. G. Ren and E. O'Neill, "3d selection with freehand gesture", *Computers & Graphics*, vol. 37, no. 3, pp. 101–120, 2013.
15. J. Jankowski and M. Hachet, "Advances in interaction with 3d environments", in *Computer Graphics Forum*, vol. 34, no. 1. Wiley Online Library, 2015, pp. 152–190.
16. M. A. Cidota, R. M. S. Clifford, P. Dezentje, S. G. Lukosch, and P. J. M. Bank, "Affording Visual Feedback for Natural Hand Interaction in AR to Assess Upper Extremity Motor Dysfunction," *2015 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 1–4, 2015.
17. M. A. Cidota, S. G. Lukosch, P. Dezentje, P. J. Bank, H. K. Lukosch, and R. Clifford, "Serious gaming in augmented reality using hmds for assessment of upper extremity motor dysfunctions," *i-com*, vol. 15, no. 2, pp. 155–169, 2016.
18. J. Jacobson and S. Werner, "Why cast shadows are expendable: Insensitivity of human observers and the inherent ambiguity of cast shadows in pictorial art," *Perception*, vol. 33, no. 11, pp. 1369–1383, 2004.
19. M. Sattler, R. Sarlette, T. Mücken, and R. Klein, "Exploitation of human shadow perception for fast shadow rendering," in *Proceedings of the 2nd symposium on Applied perception in graphics and visualization*. ACM, 2005, pp. 131–134.
20. I. M. Bullock, T. Feix, and A. M. Dollar, "Finding small, versatile sets of human grasps to span common objects," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 1068–1075, 2013.
21. T. Feix, I. M. Bullock, and A. M. Dollar, "Analysis of Human Grasping Behavior: Correlating Tasks, Objects and Grasps," *IEEE Transactions on Haptics*, vol. 7, no. 3, pp. 311–323, 2014.
22. J. Brooke, "Sus-a quick and dirty usability scale," *Usability evaluation in industry*, vol. 189, no. 194, pp. 4–7, 1996.

## REFERENCES

- [1] S. De Freitas and S. Jarvis, "Serious games engaging training solutions: A research and development project for supporting training needs," *British Journal of Educational Technology*, vol. 38, no. 3, pp. 523–525, 2007.
- [2] M. Funk, T. Kosch, R. Kettner, O. Korn, and A. Schmidt, "motioncap: An overview of 4 years of combining industrial assembly with augmented reality for industry 4.0," in *Proceedings of the 16th international conference on knowledge technologies and datadriven business*, 2016.
- [3] J. L. Gabbard and J. E. Swan II, "Usability engineering for augmented reality: Employing user-based studies to inform design," *IEEE Transactions on visualization and computer graphics*, vol. 14, no. 3, pp. 513–525, 2008.
- [4] S. Suzuki, H. Suzuki, and M. Sato, "Grasping a virtual object with a bare hand," in *ACM SIGGRAPH 2014 Posters*. ACM, 2014, p. 51.
- [5] C. Weichel, M. Lau, D. Kim, N. Villar, and H. W. Gellersen, "Mixfab: a mixed-reality environment for personal fabrication," in *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 2014, pp. 3855–3864.
- [6] T. Piumsomboon, A. Clark, M. Billinghurst, and A. Cockburn, "User-defined gestures for augmented reality," in *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2013, pp. 955–960.
- [7] M. Al-Kalbani, I. Williams, and M. Frutos-Pascual, "Improving freehand placement for grasping virtual objects via dual view visual feedback in mixed reality," in *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*. ACM, 2016, pp. 279–282.
- [8] M. Al-Kalbani, M. Frutos-Pascual, and I. Williams, "Freehand grasping in mixed reality: analysing variation during transition phase of interaction," in *Proceedings of the 19th ACM International Conference on Multimodal Interaction*. ACM, 2017, pp. 110–114.
- [9] J. E. Swan, G. Singh, and S. R. Ellis, "Matching and reaching depth judgments with real and augmented reality targets," *IEEE transactions on visualization and computer graphics*, vol. 21, no. 11, pp. 1289–1298, 2015.
- [10] C. Diaz, M. Walker, D. A. Szafrir, and D. Szafrir, "Designing for depth perceptions in augmented reality," in *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 2017, pp. 111–122.
- [11] M. Billinghurst, T. Piumsomboon, and H. Bai, "Hands in space: Gesture interaction with augmented-reality interfaces," *IEEE computer graphics and applications*, vol. 34, no. 1, pp. 77–80, 2014.
- [12] M. Al-Kalbani, I. Williams, and M. Frutos-Pascual, "Analysis of medium wrap freehand virtual object grasping in exocentric mixed reality," in *Mixed and Augmented Reality (ISMAR), 2016 IEEE International Symposium on*. IEEE, 2016, pp. 84–93.