The Influence of Environmental Lighting on Size Variations in Optical See-through Tangible Augmented Reality

Denise Kahl* Marc Ruble† Antonio Krüger‡

German Research Center for Artificial Intelligence (DFKI),
Saarland Informatics Campus, Saarbrücken, Germany

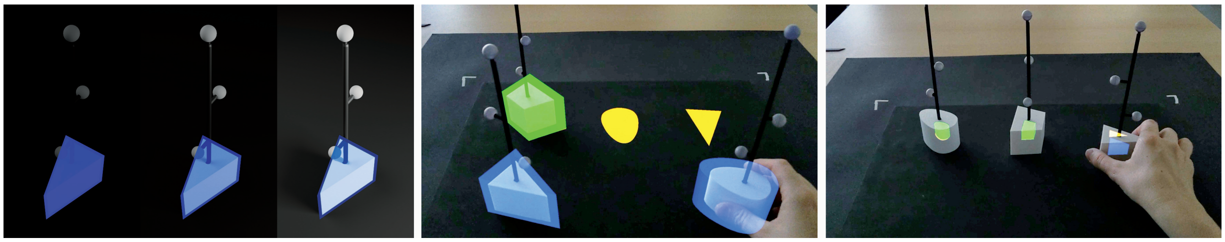

Figure 1: Impressions of the visual perception of the objects during the study. Left: Modeled triangular prism illustrating the influence of the lighting condition on the perception of the overlay in size condition L for dark, medium and bright lighting (left to right). Middle and right: Screenshots of the HoloLens 2 during the execution of the study in size condition XL and XXS: matching the blue 3D overlays to the orange 2D targets. Visually the images correspond most closely to the perception in the medium lighting condition.

Abstract

Optical see-through head-mounted displays (OST HMDs) are becoming increasingly popular as they get better and smaller. One application area is interaction with virtual content, which is more intuitive when using physical objects as tangibles. Since it is not possible to use a matching replica for each virtual object, it is necessary to identify physical objects that can represent several different virtual objects. As a first step, we investigated to what extent a physical object can differ in size from its virtual counterpart.

Since the perception of content in optical see-through Augmented Reality (OST AR) is strongly influenced by the ambient lighting, the illumination intensity was considered in our study. We investigated three indoor lighting conditions and their effects on the perception of seven different size variations between the physical object and its virtual overlay.

The results of the study show that there is a decrease in usability and presence with increasing illuminance. However, this cannot be avoided when applications are run under realistic interior lighting conditions. Furthermore, the results demonstrate that the size ranges in which a physical object can deviate from its virtual counterpart without having a strong negative impact on usability, presence and performance increase with increasing environmental illumination. Therefore, it is possible to interact with even smaller and even larger physical props to manipulate the associated virtual content under brighter lighting conditions.

Keywords: Tangible augmented reality, optical see-through augmented reality, tangible interaction, illumination.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Mixed / augmented reality; Human-centered computing—Human computer interaction (HCI)—Interaction devices—Haptic devices

*e-mail: denise.kahl@dfki.de
†e-mail: marc.ruble@dfki.de
‡e-mail: antonio.krueger@dfki.de

1 Introduction

In optical see-through Augmented Reality (OST AR), virtual content is superimposed on the real world view. This creates a link between the real and the virtual world [1, 4, 26]. The transparent display of OST HMDs provides an almost unobstructed view of the real world, unlike video see-through AR (VST AR), where virtual content is displayed on a video stream [2, 20]. In addition, due to the technology used, the overlays in OST AR are slightly transparent and let real-world objects shine through, which is why findings from the areas of VST AR and Virtual Reality (VR) cannot be transferred to OST AR. In the additive light model used in current OST HMDs, the light emitted by the display is added to the existing light from the physical environment [8]. Thus, the environmental light has a strong influence on OST HMD applications, which is why the illuminance has to be considered in all investigations in OST AR.

Manipulation tasks on virtual objects can be performed more accurately [24] and faster [3] when the digital information is coupled to physical objects for interaction [12], known as Tangible Augmented Reality. TAR allows for natural and intuitive interaction [5] especially with OST AR, where both hands are free for interaction. Nevertheless, it is not possible to provide an exact physical replica for every virtual object that is interacted with. It must therefore be investigated whether it is possible to use a physical object for the interaction that differs from the virtual model and that can ideally represent a variety of virtual objects. Visually, however, the virtual and physical object could then differ in size, shape, and texture. It must therefore be determined whether a deviation in these factors is possible without significant drawbacks for the feeling of presence, usability and performance. The determination of possible deviations must take into consideration the environmental lighting, since the illuminance affects the perception of contrast and color of the virtual overlay [7–10].

In this paper we investigate to what extent environmental lighting has an impact on how much a physical object can deviate from its virtual counterpart without a strong negative impact. We first look at variations in size by examining the effect of three different indoor lighting conditions on acceptable size variations between the virtual and physical object in a study.

This is the authors' version of the article. See https://doi.org/10.1109/VR51125.2022.00030 for the final version of record in IEEE Xplore.

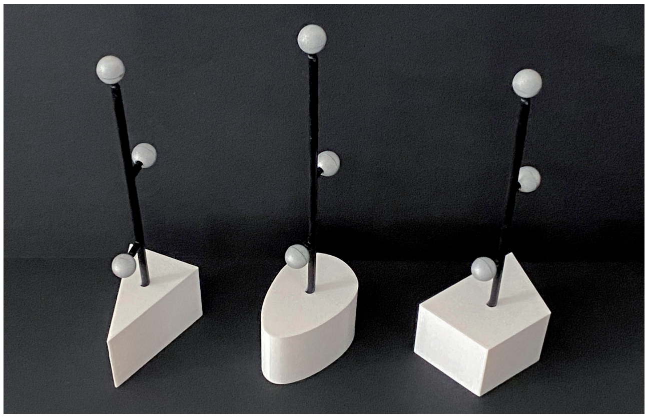

Figure 2: Physical props with attached marker trees on top. Left to right: triangular prism, egg-shaped cylinder and trapezoidal prism.

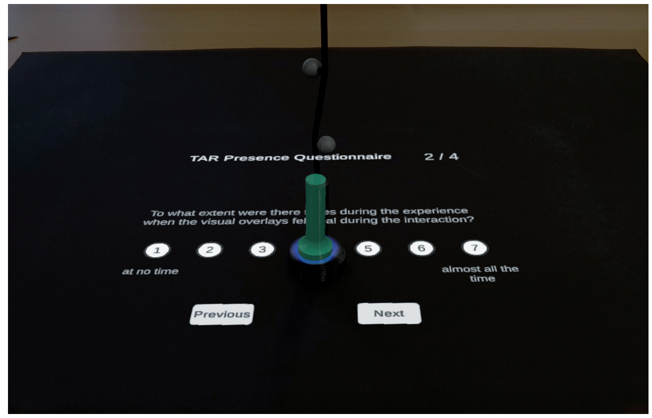

Figure 3: Excerpt from the TAR presence questionnaire: Selection with the help of the tangible pen.

2 Related Work

The influence of environmental conditions (such as natural lighting or the color of backgrounds or objects in the scene) on the perception of AR overlays in HMDs has already been studied.

Gabbard et al. [8–10] showed, among other things, that the environmental background as well as text drawing styles have a strong influence on the readability of AR content.

Furthermore, Erickson et al. [7] determined the influence of environmental lighting conditions on the contrast of AR content. In their study, they measured the light entering the eye under 6 different environmental illuminances ranging from less than $1lx$ (Lux) to over $20000lx$, covering both indoor and outdoor lighting conditions. The results show a clear decrease of the perceived contrast with increasing illuminance, so that it is almost completely eliminated in bright outdoor environments.

It is necessary to investigate to what degree the surrounding light conditions also have an influence on the extent to which a virtual object can differ in size from its physical representation.

The extent to which size differences between tangible and virtual objects are possible in VR was investigated by Simeone et al. [23] and de Tinguy et al. [6]. Simeone et al. [23] looked at the degree to which physical and virtual objects can differ without destroying the VR illusion. They found a significant degradation when using a smaller overlay, but not when using a larger overlay. In contrast, de Tinguy et al. [6] determined how much variation can exist without a difference being perceived. They found that the size can be changed up to 5.75% before a difference is noticed.

In the area of VST AR, size differences between the tangible and its virtual overlay were also considered in a study by Kwon et al. [15]. They found no performance differences between the examined size conditions. However, significant effects were found for a combination of size and shape deviation between physical prop and virtual overlay.

Acceptable size differences in OST AR were identified by us in a previous study [13]. We investigated to what extent the virtual object can differ in size from its physical representation without having a significant negative impact on usability, presence, and performance. We considered size differences from -50% to +50% of the base condition, where the size of the virtual and physical object was identical. We found that the physical object can vary in size and that larger variations are possible for larger overlays than for smaller ones, which fits with the results of Simeone et al. [23]. Like Kwon et al. [15], we did not find any performance differences between size conditions. However, the study was conducted in a darkened room with very high opacity overlays. Furthermore, the lighting condition chosen for the experiment was unrealistic for an AR use case. An investigation of how the perception of differently sized virtual and physical objects changes under more realistic brighter illumination conditions is therefore necessary. As the brightness increases, the contrast of the virtual overlay decreases and so the visibility of the physical object behind the virtual overlay changes. Due to the different perception in different lighting conditions, we also expect different results here.

In this paper, we therefore investigate the influence of illuminance on possible size variations between the virtual object and its physical representation. In a study under controlled lighting conditions, the effect of three different indoor illuminances on interaction with differently sized virtual objects is investigated by evaluating how much the physical objects can deviate from their virtual representations in the respective lighting conditions without having a strong negative impact on presence, usability and performance.

3 Technical Implementation

To create a Tangible Augmented Reality experience, we based our technical setup on that of our previous study, presented in detail in [13]. As an optical see-through HMD, we chose the Microsoft HoloLens 2 [18] which runs a Unity [25] application (Version 2019.3.8f1) as the client of our system.

Furthermore, we use a marker-based motion capture system by OptiTrack [19] to achieve a high precision and accuracy when detecting where in space the tracked objects are. For this purpose, we attach infrared reflective markers to the physical props for interaction (see Fig. 2) as well as to the HMD (similar to Liu et al. [16]). In contrast to our previous study, we tracked the markers using the newer OptiTrack Prime^{X} 13 cameras and the accompanying tracking software Motive 2.3. With this approach, tracking results were consistently precise and accurate, such that further processing was no longer necessary.

The communication between our central server component, which receives tracking data and commands from the experimenters, and the client application running on the HoloLens is based on M2Mqtt for Unity [27]. In contrast to a single TCP connection (as used in [13]), the MQTT protocol enables more efficient communication by splitting the information sent into different channels for the client to receive. This approach allows us to assign different priorities to different kinds of data and adjust their sending frequency or quality of service individually. For example, for each of the physical props being tracked, as well as the HoloLens itself, our server sends 60 messages per second, including the object's position, rotation and visual state (default, highlighted/green or invisible). These messages have a high frequency to create a fluent movement of the virtual overlays, but they are sent in a "fire and forget" manner. Command

Figure 4: Illustrations of the laboratory setup in our three lighting conditions Dark, Medium and Bright (left to right). Studio lamps and lamps on the ceiling were used to adjust the environmental illuminance.

messages like changing the current size condition of the overlays are sent only once when they are triggered, but are guaranteed to arrive at the client.

Additionally, we extend our previous system [13] by including the questionnaires in the TAR experience. The selection of answers can be made via a designated tangible which acts like a virtual pen, and text is displayed in AR on the table in front of the participant (see Fig. 3). With this approach, the participants do not have to lift up the HoloLens display in order to fill out a regular questionnaire on paper. This is important especially for questionnaires presented in the middle of an experiment because movement of the HMD, or even just its display, would require a recalibration to ensure overlays are still perceived at the correct positions afterwards. Reading through the HoloLens display is theoretically possible, but very hard and uncomfortable. We are not aware of any research about potential effects of leaving the AR environment in order to fill out questionnaires. However, advantages of virtual questionnaires are known from VR research [21, 22] and we try to minimize the break between the AR experience and the questionnaires that refer to it. In our study, we received a lot of positive feedback for this new way of filling out questionnaires. Some participants even mentioned that they enjoyed filling out the questionnaires with the tangible pen.

To optimize performance, we follow the design guideline of using the HoloLens client application only for display purposes and outsourcing all processing tasks to the server component. The HoloLens therefore only receives information about where a particular prop should be displayed and at what size, whether calibration displays should be visible or which questionnaire texts and answers should be displayed at well-defined positions.

4 Study

As illuminance increases, the contrast of the overlays displayed in HoloLens 2 decreases [7]. We therefore assume that the virtual objects displayed over the physical ones will appear to have varying degrees of transparency, and the perceived transparency will increase as the illuminance increases. As the increased transparency of the overlays makes the physical props behind them more visible, we expect the size estimation to be more accurate in brighter light. Similarly, we expect that the changed perception of the virtual object will lead to a difference in terms of possible size ranges in which the virtual and physical object can differ from each other without any significant degradation in the perception of presence, usability and performance. We therefore state the following hypotheses:

- H1: With increasing environmental illumination, the perceived transparency of the overlays increases.
- H2: Size estimation is more accurate in brighter lighting conditions.
- H3: The ranges in which the size of the virtual and physical object can differ without worsening of presence, usability and performance vary for different lighting conditions.

We defined a task to test the hypotheses stated above. In order to determine to what extent the selected lighting condition has an influence on the perception of AR proxy interactions when using differently sized virtual representations, participants had to solve a puzzle task as quickly and as precisely as possible. They had to place and align three different virtual objects on associated 2D target shapes (see Fig. 1, center and right). For each task, all three objects simultaneously had to be arranged 2x each to generate the highest possible number of interactions, which is crucial for the evaluation of disturbance during grasping [15]. For this, the physical objects had to be lifted, rotated and moved to place them exactly, all of which are simple/basic subtasks, but which have to be combined to solve more complex tasks [17].

The study was approved by the ethical review board of our faculty and carried out in compliance with a strict hygiene plan.

4.1 Participants

24 participants (15 male, 9 female) aged between 21 and 55 ($M = 25.625, SD = 6.983$) took part in our study. Participants who were not associated with our institution received 15 Euro as compensation for participating in the experiment, which lasted about 90 minutes. All had normal or corrected-to-normal vision and happened to be right-handed. Prior experience with AR and AR glasses was rated on a 7-point Likert scale from 1 (= never) to 7 (= regular). Participants had low experience with AR ($M = 2.042, SD = 0.908$) and even less experience with AR glasses ($M = 1.5, SD = 0.722$).

4.2 Apparatus

The study took place in a darkened room. This ensured that the lighting conditions were the same for all participants at all times and were not influenced by external factors. Depending on the lighting condition, the room was illuminated by 2 to 3 softbox studio lamps and the fluorescent tubes of the ceiling lighting.

Figure 4 shows the setup of the lamps for the respective lighting conditions. In the Dark condition, only two studio lamps were used, pointing diagonally upwards away from the participant. In condition Medium, these lamps were turned towards the participant and a third studio lamp was installed, which was directed upwards to provide additional ambient light. In addition, the fluorescent tubes on the ceiling were switched on in the Bright condition. We measured the illumination on the tabletop facing upwards and from the HoloLens camera looking diagonally downwards onto the interaction surface. The measured luminance values are listed in Table 1.

Table 1: Lighting intensity in lx, measured on the tabletop pointing upwards and at the HoloLens camera pointing towards the interaction area in the different lighting conditions.

	Dark	Medium	Bright
Tabletop	10.75	49.5	257
HoloLens	4.4	14	45

This is the authors' version of the article. See https://doi.org/10.1109/VR51125.2022.00030 for the final version of record in IEEE Xplore.

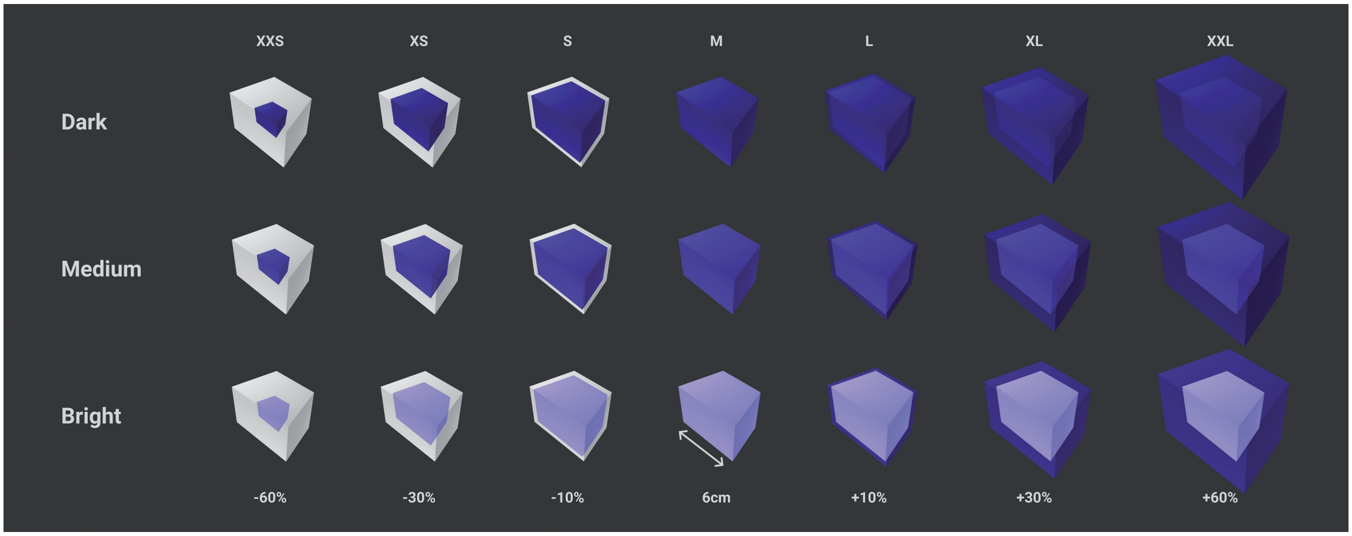

Figure 5: Size variations of the virtual overlays (blue) compared to the physical proxy objects (white). Condition M is the base condition with matching size of virtual and physical object. Condition XL, e.g., is obtained by scaling the virtual object by a factor of 1.3 along all three axes. The overlay opacities approximately reflect the participants' perception of the overlays in the three lighting conditions.

The tracking of the HoloLens and the physical props was done with the help of 6 OptiTrack Prime^X 13 cameras, which were installed on a truss of about 4m x 6m at a height of about 2.6m and pointed towards the center of the tracking area. Participants sat at a table located at the center of the OptiTrack cameras.

For interaction, white physical props were used, which were equipped with a black marker tree for tracking (see Fig. 2). Interaction was performed on a black plate on a black background, which was chosen to be large enough that the virtual objects were always visible against a black background.

We used a HoloLens 2 for the visualization of the overlays, whose brightness we set to 100%. For the overlays, we chose an opacity of 100% and a medium-dark blue (#2300D1). In preliminary tests, we found that the white prop in combination with this blue overlay leads to different perceptions in each lighting condition, which would not have been possible with, e.g., a white overlay. We wanted the overlays to not be too bright in the dark, but still almost hide the physical objects. In the medium condition there should be a balance between the intensity of the overlay and the physical object, and in the bright condition the physical object should be in the focus. Figure 1 (left) provides an indication of how the objects might have been perceived in the different lighting conditions.

We ensured that all participants had approximately the same viewing angle (approx. 45 degrees) on the props by placing the chair at a designated location and adjusting its height so that the distance between the HoloLens and the table was approximately 52cm for each participant.

4.3 Basic Approach

The study procedure was the same for each participant. We followed the procedure of [13] to obtain comparable results.

First, an eye calibration was performed with each participant using the app available on the HoloLens 2 to adjust the visualization to the user's depth perception. This calibration was manually rechecked by displaying a virtual object superimposed on a physical object. The HoloLens 2 was then aligned on the head so that the overlay exactly matched the position of the physical calibration object. This ensured that the overlays would later be displayed in the correct position during the study. In addition, if the HoloLens depth adjustment did not work accurately, which was sometimes the case with glasses wearers, the overlay was manually adjusted in depth until its position matched the physical prop. Using this approach, we could also avoid common accuracy errors of the optical see-through display which were investigated by Khan et al. [14].

For each task the participants performed, the same procedure was followed. The props to be interacted with were arranged randomly in the bottom area on a plate and covered with a box. This was then placed on the table in front of the participant. As soon as the participant had adjusted the field of view of the HoloLens to the interaction area, the box was lifted up by the study leader. At this moment the respective task started and the time measurement was activated. At the same time, the virtual overlays appeared superimposed on the physical props and the virtual targets were displayed. After finishing each task, participants had to complete three questionnaires in Augmented Reality: an AR presence questionnaire, a TAR presence questionnaire, and a size perception questionnaire. The AR presence questionnaire consisted of four questions and evaluated how realistic the overlays in the respective task appeared and how strongly the participants felt that they were in an unaltered reality. The TAR presence questionnaire, also consisting of four questions, evaluated how realistic the interaction with the virtual objects felt and how strong the feeling of interaction with the virtual overlays was while touching the differently sized physical props. In the size perception questionnaire, consisting of three questions, the size differences between the virtual and the physical object were evaluated, e.g., with respect to confusion during grasping. After completion of each lighting condition, a lighting questionnaire consisting of five questions had to be completed in addition to the other three questionnaires. In this questionnaire, besides rating how transparent the overlays felt, the participants had to evaluate the naturalness of the environmental lighting in the just-experienced lighting condition. All four questionnaires were rated using 7-point Likert scales. When all tasks were completed, the participants received a final paper questionnaire asking for demographic information and additionally for a ranking of the lighting conditions.

4.4 Design and Procedure

The study was designed as a within-subject experiment. For 3 different lighting conditions we tested 7 different size conditions each. The order of the lighting conditions was counterbalanced by a Williams design Latin square (LS) of size 3 [28]. We also balanced

This is the authors' version of the article. See https://doi.org/10.1109/VR51125.2022.00030 for the final version of record in IEEE Xplore.

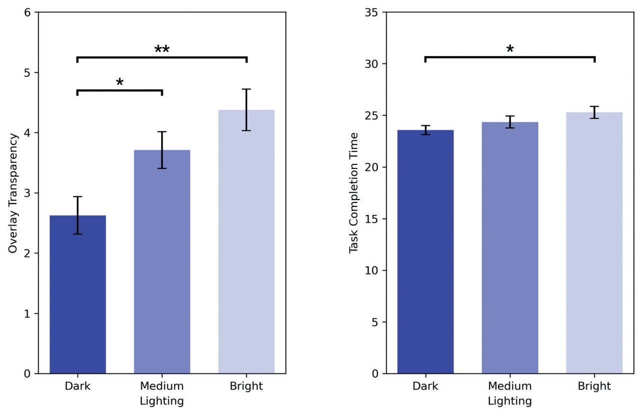

Figure 6: Mean transparency ratings (rated from 1 to 7) regarding the overlays (left) and mean task completion times in seconds (right) with marked standard errors for each lighting condition. Significant differences between conditions are marked with * ($p < 0.05$), ** ($p < 0.01$) and *** ($p < 0.001$).

the order of the size conditions, which were presented as a block in each lighting condition.

For lighting conditions, we chose low illumination (conditon Dark) similar to that in [13], medium illumination (condition Medium), and high illumination (condition Bright).

We investigated the influence of each lighting condition on seven different size variations between the virtual and the physical object. Our baseline is represented by condition M, where the physical object is the same size as the virtual object.

Figure 5 visualizes the size differences between the virtual and physical objects in the individual lighting conditions. The views shown may differ from the perception in HoloLens 2, but provide an indication of how the objects might have been perceived. Since the participants' perception of the overlays is affected by the environmental lighting, it is not useful to take screenshots with the HoloLens because the overlays would look the same in screenshots under all lighting conditions.

In addition to condition M, three smaller virtual overlays and three larger virtual overlays each were examined. Size conditions S and L represent a small size variance with a 10% difference, followed by XS and XL with a medium size variance of 30% and size conditions XXS and XXL with a large size variance of 60%. The XXL condition represents the largest possible size variance in which three objects can be interacted with simultaneously in the HoloLens 2 field of view without the overlays overlapping. Hence, we examined a larger size variance than in our previous study [13], in which a maximum size variance of 50% was not sufficient to determine upper limits for each measure.

We wanted to find out how much a physical object used for interaction can differ from its virtual counterpart without significantly degrading presence, usability, and performance. In addition, we wanted to determine whether these ranges change under different environmental illuminances.

For this purpose, we had participants interact simultaneously with three different physical objects. The triangular prism, the egg-shaped cylinder, and the trapezoidal prism (see Fig. 2) had a width of 6cm and a height and depth of 4cm each. These shapes are taken from [13] and are based on existing work in VR and VST AR [6, 15] and chosen so that the objects are easy to grasp by hand [11].

Due to the design of the objects, there was only one way to correctly place objects on given targets, which at the same time required a maximum rotation of the objects by up to 180°, so that the participants had to perform a maximum amount of interaction.

The task was to place the three virtual objects on their corresponding 2D targets. Each object had to be assigned twice to complete the task. At first, three targets were displayed on the upper part of the plate, on which the objects had to be placed. Once an object was correctly arranged, the color of its overlay changed to green (see Fig. 1, center and right). After all objects in the top row were correctly arranged, their colors changed back to blue and the targets disappeared. Then new targets appeared in the lower area of the plate and the objects had to be assigned again. As soon as all objects in the bottom row were green, the task was considered completed and the timing stopped.

We placed the virtual objects so that their centers matched those of the physical objects. In order to be able to place the objects on the 2D targets, we adjusted the height of the targets in 3D space so that it matched the bottom of the virtual objects. Visually, however, from the participants' point of view, regardless of the size of the overlays, it always appeared as if the targets were on the tabletop.

There were 6 different possible arrangements for the targets in the upper area of the plate as well as in the lower area. We randomized at which position each target was displayed and randomly generated the rotations of each target. Likewise, a random initial arrangement of physical objects on the plate was performed.

We preliminarily determined suitable deviations in space that had to be achieved for an assignment to be considered fulfilled. As soon as the virtual object was, continuously for 0.5 seconds, less than 0.4cm in flat distance and less than 0.5cm in height away from the target position, and the angle between target and object was less than 3°, the object was considered correctly placed. We explicitly decided to use a stopping criterion in order to be able to perform meaningful time measurements. Since everyone evaluates accuracy differently, there would have been a strong impact on task completion time if everyone could have decided for themselves when the task was completed. Besides performance (task completion time), we measured usability (disturbance in grasping and interaction) and presence (AR presence and TAR presence) as well as the perception of lighting.

4.5 Results

We present the effects of different lighting conditions in the environment and size variations between corresponding virtual and physical objects on the aspects presented below. For each kind of measure, we first compare all samples collected in the three lighting conditions with each other by applying a Friedman test with 2 degrees of freedom and a significance level of $\alpha = 0.05$. Additionally, we report its test statistic χ^2. As a post-hoc test, we use Wilcoxon's signed-rank test with 167 degrees of freedom and a significance level of $\alpha = 0.05$. Furthermore, we state test statistic W and the matched pairs rank-biserial correlation r as an effect size. Subsequently, we inspect each of the three lighting conditions individually to investigate the effect of the seven size variations in each specific lighting environment. For this, we use a Friedman test ($dof = 6$, $\alpha = 0.05$) as well as a Wilcoxon's signed-rank test ($dof = 23$, $\alpha = 0.05$) to compare each size with condition M as a baseline. All of these Friedman tests showed a significant influence of the size condition under any of the three lighting conditions. For the applications of Wilcoxon's signed rank test, we report the Bonferroni-corrected p-values. Figure 8 gives an overview of our results obtained using the post-hoc tests.

4.5.1 Overlay Transparency

After each lighting condition, participants were asked to rate how transparent they perceived the overlays to be in the lighting questionnaire. The resulting mean transparency ratings for each lighting are displayed in Fig. 6. The Friedman test ($dof = 2$, $\alpha = 0.05$) showed a significant influence of the lighting condition on the perceived transparency of the virtual overlays ($\chi^2 = 23.089, p < 0.001$) rated on a

This is the authors' version of the article. See https://doi.org/10.1109/VR51125.2022.00030 for the final version of record in IEEE Xplore.

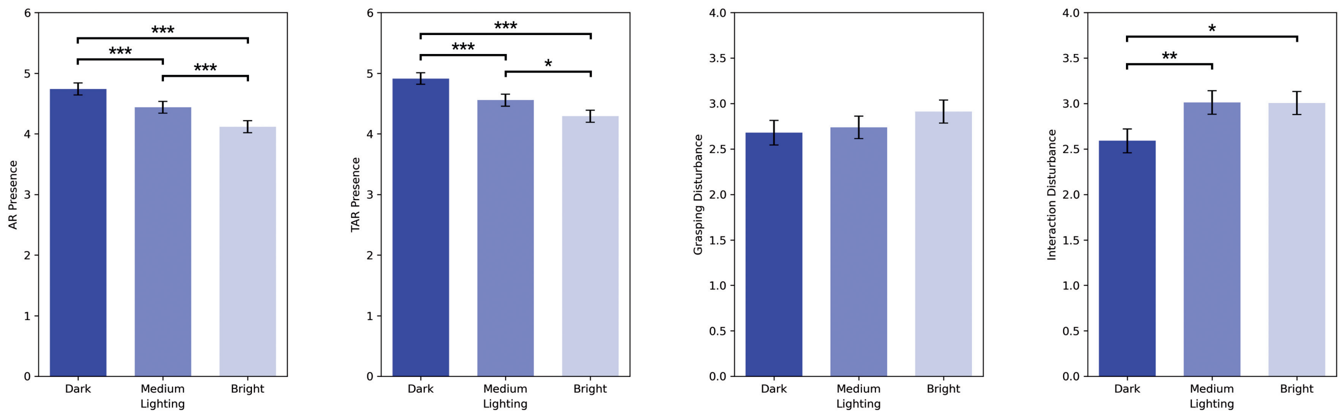

Figure 7: Mean presence and disturbance ratings from 1 to 7 with marked standard errors for each lighting condition. Significant differences between conditions are marked with * ($p < 0.05$), ** ($p < 0.01$) and *** ($p < 0.001$).

scale from 1 (not transparent at all) to 7 (completely transparent). The post-hoc test ($dof = 23$, $\alpha = 0.05$) confirmed this influence by revealing significantly lower transparency ratings in condition Dark compared to Medium ($W = 14, p = 0.015, r = -0.794$) and Bright ($W = 21, p = 0.002, r = -0.834$).

Results show that with brighter lighting the perceived overlay transparency increases, although the difference between medium and bright lighting is not statistically significant.

4.5.2 Size Estimate

The Friedman test could not determine a statistically significant influence of the lighting condition on the overall size estimates the participants gave by rating the size difference between virtual and physical object on a range from 1 (= virtual much smaller), over 4 (= equal-sized) to 7 (=virtual much larger).

In the dark lighting, all size variations were rated significantly different from the size-matching condition M. However, the estimates for condition S in medium lighting ($W = 19.5, p = 0.094, r = -0.675$) as well as condition L in bright lighting ($W = 30, p = 0.382, r = 0.5$) did not differ significantly from the baseline size M in their respective lighting condition.

Therefore only small size variations could not be differentiated significantly from the baseline M, a small size reduction S in medium lighting and a small size addition L in bright lighting.

4.5.3 Presence

The resulting mean presence ratings for each lighting are displayed in Fig. 7. Starting with AR presence, the Friedman test shows a significant influence of the lighting condition on the ratings ($\chi^2 = 39.173, p < 0.001$). The post-hoc tests determined that ratings in condition Dark were higher than in condition Medium ($W = 3333, p < 0.001, r = 0.387$), in Medium higher than in Bright ($W = 3759, p < 0.001, r = 0.345$) and therefore also in Dark higher than in Bright ($W = 2456.5, p < 0.001, r = 0.583$).

Under dark lighting, size conditions XXS ($W = 10, p < 0.001, r = -0.928$), XS ($W = 15, p = 0.005, r = -0.857$) and XXL ($W = 39, p = 0.005, r = -0.74$) led to significantly lower AR presence scores. Similarly in medium lighting, sizes XXS ($W = 23, p < 0.001, r = -0.847$), XS ($W = 41.5, p = 0.036, r = -0.672$) and XXL ($W = 28.5, p = 0.001, r = -0.81$) and in bright lighting only sizes XXS ($W = 35.5, p = 0.019, r = -0.719$) and XXL ($W = 39, p = 0.049, r = -0.662$) showed negative effects.

Regarding TAR presence, the Friedman test shows a significant influence of the lighting in the environment on the ratings ($\chi^2 = 32.818, p < 0.001$). The post-hoc comparisons determined that ratings in condition Dark were higher than in condition Medium ($W = 3083, p < 0.001, r = 0.409$), in Medium higher than in Bright ($W = 4418.5, p = 0.015, r = 0.26$) and therefore also in Dark higher than in Bright ($W = 2836, p < 0.001, r = 0.512$).

Under dark lighting, size conditions XXS ($W = 19.5, p = 0.002, r = -0.859$), XS ($W = 36.5, p = 0.012, r = -0.736$) and XXL ($W = 46.5, p = 0.033, r = -0.663$) led to significantly lower TAR presence scores. Similarly in medium lighting, sizes XXS ($W = 19, p < 0.001, r = -0.873$), XS ($W = 25, p < 0.001, r = -0.833$) and XXL ($W = 34, p = 0.016, r = -0.731$) and in bright lighting only size XXS ($W = 29, p = 0.01, r = -0.771$) significantly worsened presence.

AR and TAR presence behave very similarly to the extent that the darker the lighting condition, the higher the rated presence scores were. For dark and medium lighting, we could observe significantly worse presence for very large (XXS) or large (XS) size reductions as well as very large size additions (XXL). However in the bright environment, size reductions would have to be very large (XXS) to cause such an effect on AR and TAR presence, while only the largest size addition (XXL) led to a significant decrease in AR presence. For TAR presence, we could not find an upper size deviation limit in the bright lighting condition.

4.5.4 Usability

The resulting mean disturbance ratings for each lighting condition are displayed in Fig. 7. Starting with disturbance while grasping the objects, the Friedman test shows a significant influence of the lighting condition on the ratings ($\chi^2 = 7.281, p = 0.026$). However, the post-hoc tests could not determine a significant difference between the scores in any of the conditions compared pair-wise.

Under dark lighting, size conditions XXS ($W = 0, p < 0.001, r = 1$), XS ($W = 0, p < 0.001, r = 1$), S ($W = 0, p = 0.029, r = 1$), XL ($W = 11.5, p = 0.021, r = 0.831$) and XXL ($W = 0, p < 0.001, r = 1$) led to significantly higher grasping disturbance scores. Similarly in medium lighting, sizes XXS ($W = 13.5, p = 0.006, r = 0.858$), XS ($W = 21, p = 0.029, r = 0.754$), XL ($W = 22, p = 0.018, r = 0.768$) as well as XXL ($W = 6, p = 0.001, r = 0.943$) and in bright lighting only sizes XL ($W = 8, p = 0.007, r = 0.895$) and XXL ($W = 21, p = 0.004, r = 0.834$) showed significantly increased ratings for disturbance during grasping.

Regarding disturbance during interaction with the objects, the Friedman test shows a significant influence of the lighting condition on the ratings ($\chi^2 = 15.825, p < 0.001$). The post-hoc tests determined that ratings in condition Dark were significantly lower than in condition Medium ($W = 2343, p = 0.002, r = -0.344$) and condition Bright ($W = 2438.5, p = 0.015, r = -0.294$).

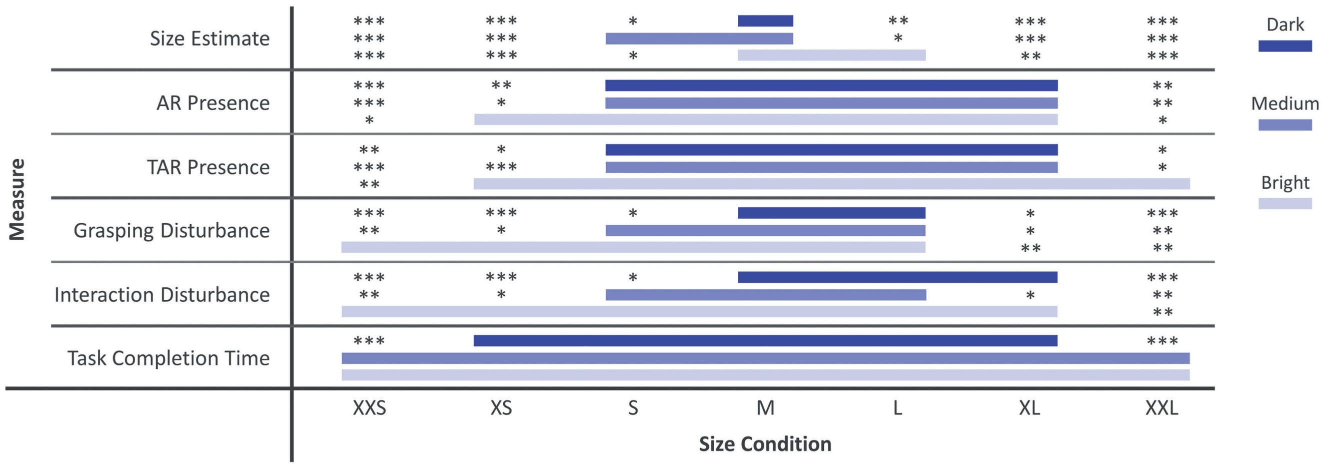

Figure 8: Summary of significant differences of the size conditions compared to the size-matching condition M as a baseline marked with * ($p < 0.05$), ** ($p < 0.01$) and *** ($p < 0.001$) for all three lighting conditions. The blue bars indicate the ranges without significant difference in the respective lighting condition.

Under dark lighting, size conditions XXS ($W = 0, p < 0.001, r = 1$), XS ($W = 0, p < 0.001, r = 1$), S ($W = 4, p = 0.021, r = 0.912$) and XXL ($W = 0, p < 0.001, r = 1$) led to significantly higher interaction disturbance scores. Furthermore in medium lighting, size conditions XXS ($W = 3, p = 0.002, r = 0.965$), XS ($W = 23, p = 0.022, r = 0.758$), XL ($W = 16, p = 0.043, r = 0.765$) and XXL ($W = 3, p = 0.001, r = 0.968$) and in bright lighting only size condition XXL ($W = 3, p = 0.003, r = 0.961$) showed significantly increased ratings.

Although we could not determine a significant difference when comparing the lighting conditions to each other regarding disturbance during grasping, the results show clearly that the dark lighting condition leads to overall lower disturbance during interaction with the objects. But at the same time, in this dark lighting, introducing size differences between physical and virtual objects has a larger negative effect. Every size reduction (S, XS and XXS) shows increased disturbances for both grasping and interacting, while sizes XL and XXL indicate this effect for grasping and only XXL for interacting, respectively. For light condition Medium, both types of disturbance increase significantly for larger size deviations XXS and XS as well as XL and XXL. However, for the light condition Bright we could not find a significant worsening for smaller overlays; still, strong enlargements XL and XXL are significantly more distracting during grasping, while only XXL is more disturbing during interaction.

4.5.5 Performance

The resulting mean time measurements for each lighting condition are displayed in Fig. 6. The Friedman test showed a significant influence of the lighting condition on the measured task completion times ($\chi^2 = 7.429, p = 0.024$). Wilcoxon's signed-rank test revealed significantly smaller task completion times overall in Dark compared to Bright ($W = 5245, p = 0.01, r = -0.261$).

Under dark lighting, only size conditions XXS ($W = 23, p < 0.001, r = 0.847$) and XXL ($W = 17, p < 0.001, r = 0.887$) required significantly more time than the size-matching condition M, whereas in medium or bright lighting, none of the size conditions showed a significant deviation from baseline M.

So while the dark environment led to overall faster performance by the participants, very large size differences between the virtual and physical objects in conditions XXS and XXL have a significant negative effect compared to a matching object. These effects do not appear in the brighter light conditions.

4.5.6 Lighting

In the lighting questionnaire, participants were also asked about how natural they rate the just-experienced environmental lighting. The Friedman test ($dof = 2$, $\alpha = 0.05$) showed a significant influence of the lighting condition on how natural it was rated ($\chi^2 = 11.195, p = 0.004$) to be. The post-hoc test ($dof = 23$, $\alpha = 0.05$) revealed that the environmental lighting was rated as significantly less natural in condition Dark than Bright.

4.5.7 Final Questionnaire

After the study, participants finished with a concluding questionnaire where they were asked to rank the three lighting conditions based on the perceived realism and easiness while interacting with the objects and how much they liked the experience in the lighting condition. Table 2 shows the cumulative sum of the scores of all participants for the three conditions. The highest valued condition is given 3 points, the second 2 points and finally the lowest valued condition 1 point each in the sum.

Table 2: Scores for each size condition in realism, easiness and preference according to participants' rankings.

	Dark	Medium	Bright
Realism	63	50	31
Easiness	52	52	40
Preference	56	48	40

In this ranking, light condition Dark scored the most points regarding realism and preference, and equally many with condition Medium regarding easiness. Condition Bright scored lowest in every category. Sickness after the experiment was rated on a scale from 1 (= not at all) to 7 (= very sick) as low ($M = 2.0, SD = 1.383$).

5 Discussion

We hypothesized that environmental illuminance has an influence on the perception of virtual overlay transparency and that the overlays appear more translucent with increasing illuminance (H1). Our results support hypothesis H1. Figure 6 (left) clearly shows the increase in perceived transparency with increasing brightness. A significant difference in transparency perception was found between dark lighting and medium lighting, and thus also between dark and bright lighting.

Since the physical objects are more visible with higher transparency of the virtual overlay, we hypothesized that this would also allow more accurate size estimation under brighter lighting conditions (H2). However, our results do not support hypothesis H2. They show that size estimation was best in dark lighting. In medium lighting, however, condition S was often perceived as size matching condition M, so that there was no significant difference. In bright lighting, on the other hand, there was no significant difference between size conditions L and M.

Furthermore, we hypothesized that ranges in which size variations are possible without significant degradation in the perception of presence, usability, and performance would differ for each lighting condition (H3). Hypothesis H3 was supported by our study results. Figure 8 shows the ranges in which size variations were possible without significant degradation in each lighting condition. With regard to the perceived presence, a slight increase of the ranges with increasing brightness can be seen. While a size variation from condition S to condition XL was already possible in dark lighting, this range increases in bright lighting to XS to XL for AR presence and even XS to XXL for TAR presence. Therefore, with brighter environmental lighting, stronger size differences between the virtual and physical object are possible without causing a significant degradation of presence.

With regard to possible size variations without significant deterioration of usability (disturbance in grasping and interaction), there are also differences between the individual lighting conditions. While in dark lighting only condition L was not perceived as significantly worse in terms of disturbance in grasping, in medium lighting this was already the case for condition S and L. In bright lighting, the range without significant deterioration even increased, from condition XXS to condition L. As the physical objects were more visible as brightness increased, it is likely that this enabled participants to adjust grasping accordingly. For disturbance in interacting, the greatest difference is seen in bright lighting, where a variance between XXS and XL is possible, while in medium lighting only S to L and in dark lighting only M to XL is possible. Once the objects were grasped, the participants were probably aware of the size difference and could adjust to it. The results show that under brighter light conditions larger size variations are possible without significantly degrading the disturbance.

In terms of performance, there were significant differences in dark lighting for conditions XXS and XXL compared to baseline condition M. In contrast, under the two brighter lighting conditions, no differences in performance were detected between the individual size conditions and condition M. Overall, very large size differences between the virtual and physical object are therefore possible without having a strong negative impact on performance.

Comparing the overall ratings of the individual lighting conditions with each other, we see that the performance worsens with increasing environmental illumination (see Fig. 6, right). Furthermore, with brighter illumination, AR presence and TAR presence deteriorate and disturbance in grasping and interaction increases (see Fig. 7).

For these reasons, the priority regarding the dark condition in terms of the ratings for realism, easiness and preference can likely be explained. The dark lighting condition is rated so well probably because the virtual overlays looked quite intense in the dark and there was little visual distraction from the physical objects or the participants' hands.

Even though darker lighting conditions were rated better, the brighter lighting conditions were more in line with indoor reality, as the environmental illuminance was evaluated as more natural in these lighting conditions. Therefore, when implementing real-life applications for indoor environments, minor losses in terms of presence and usability must be accepted in OST AR.

Overall, it can be seen that the range in which size differences between the virtual and physical object are acceptable increases with increasing brightness. The main surprise is that the acceptable size ranges are smallest in the dark lighting condition. We had expected the size differences to have the greatest impact in the medium lighting condition, where they are most noticeable as the physical prop and the virtual object have roughly equally good visibility. However, size differences have the greatest influence in dark lighting. Therefore, especially under brighter, more natural, indoor illuminations, it is possible to use physical objects with a smaller or larger size compared to their virtual counterpart as tangible prop.

6 Limitations

In this paper, we investigated the effect of environmental illuminance on the perception of size variations between a virtual overlay and its corresponding physical proxy object during interaction.

For this study, we selected three lighting conditions that were to influence the perception of the overlays to different degrees. However, we only considered artificial lighting conditions, without the influence of daylight. To ensure constant lighting conditions during the entire study, only a maximum illuminance of $257lx$ (measured on the table upwards, or respectively $45lx$ at the HoloLens pointing at the interaction area on the table) was possible. At higher illuminance levels, the results would probably have been even more extreme, since the contrast of the HoloLens decreases significantly in this range according to [7].

Additionally, we only considered size variations between -60% and +60% of the baseline condition length. Interactions with larger virtual objects would not have been possible in the rather small field of view of the HoloLens 2. Due to the size constraint, not all upper or lower limits of size variations may have been detected. In addition, no precise limits have been established; to do so would require specific methods.

For technical reasons, a slight delay between the physical object and the virtual overlay could not be completely avoided, especially during fast interactions with the objects. However, since this was the same for all conditions, a negative effect on the results is not to be expected.

7 Conclusion

This paper investigated the extent to which lighting conditions affect how size differences are perceived between a virtual object and its physical representation used for interaction. For this purpose, a study was conducted to examine, under three indoor lighting conditions, to what extent the physical object can deviate from its virtual representation without significantly degrading presence, usability, and performance.

The results show that the environmental illuminance influences the visual perception of the virtual objects. The virtual objects appear more transparent under brighter lighting conditions and thus make the physical object behind them appear clearer. This different visual perception also has an influence on the size range in which a physical object can differ from its virtual counterpart. In the dark condition, size variations are already possible within a certain range. With increasing brightness these ranges become larger, so that it is possible to work with even larger/smaller objects compared to the virtual object. However, the results also show that presence and usability decrease with increasing luminance, but this must be accepted for applications in realistic indoor lighting conditions.

So far, we can only make a statement about possible size variations in OST AR. However, there are other factors, such as shape differences, which must be investigated more closely in the future in order to be able to make a determination about the extent to which a physical object can deviate from its virtual counterpart.

REFERENCES

[1] R. T. Azuma. A survey of augmented reality. *Presence: Virtual and Augmented Reality*, 6(4):355–385, 1997.

[2] G. Ballestin, M. Chessa, and F. Solari. A registration framework for the comparison of video and optical see-through devices in interactive augmented reality. *IEEE Access*, 9:64828–64843, 2021.

[3] L. Besançon, P. Issartel, M. Ammi, and T. Isenberg. Mouse, tactile, and tangible input for 3D manipulation. In *CHI '17 Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 4727–4740. ACM, 2017.

[4] M. Billinghurst, A. Clark, and G. Lee. A survey of augmented reality. *Foundations and Trends in Human-Computer Interaction*, 8(3):73–272, 2015.

[5] M. Billinghurst, H. Kato, and I. Poupyrev. Tangible augmented reality. *ACM SIGGRAPH Asia*, 7(2), 2008.

[6] X. de Tinguy, C. Pacchierotti, M. Emily, M. Chevalier, A. Guignardat, M. Guillaudeux, C. Six, A. Lécuyer, and M. Marchal. How different tangible and virtual objects can be while still feeling the same? In *2019 IEEE World Haptics Conference (WHC)*, pp. 580–585. IEEE, 2019.

[7] A. Erickson, K. Kim, G. Bruder, and G. F. Welch. Exploring the limitations of environment lighting on optical see-through head-mounted displays. In *Symposium on Spatial User Interaction*, pp. 1–8. ACM, 2020.

[8] J. L. Gabbard, J. E. Swan, and D. Hix. The effects of text drawing styles, background textures, and natural lighting on text legibility in outdoor augmented reality. *Presence: Virtual and Augmented Reality*, 15(1):16–32, 2006.

[9] J. L. Gabbard, J. E. Swan, D. Hix, S.-J. Kim, and G. Fitch. Active text drawing styles for outdoor augmented reality: A user-based study and design implications. In *2007 IEEE Virtual Reality Conference*, pp. 35–42. IEEE, 2007.

[10] J. L. Gabbard, J. E. Swan, D. Hix, R. S. Schulman, J. Lucas, and D. Gupta. An empirical user-based study of text drawing styles and outdoor background textures for augmented reality. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.*, pp. 11–18. IEEE, 2005.

[11] J. W. Garrett. The adult human hand: Some anthropometric and biomechanical considerations. *Human Factors*, 13(2):117–131, 1971.

[12] H. Ishii and B. Ullmer. Tangible bits: Towards seamless interfaces between people, bits and atoms. In *CHI '97 Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, pp. 234–241. ACM, 1997.

[13] D. Kahl, M. Ruble, and A. Krüger. Investigation of size variations in optical see-through tangible augmented reality. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 147–155. IEEE, 2021.

[14] F. A. Khan, V. V. R. M. K. R. Muvva, D. Wu, M. S. Arefin, N. Phillips, and J. E. Swan. Measuring the perceived three-dimensional location of virtual objects in optical see-through augmented reality. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 109–117. IEEE, 2021.

[15] E. Kwon, G. J. Kim, and S. Lee. Effects of sizes and shapes of props in tangible augmented reality. In *ISMAR '09 Proceedings of the 2009 8th IEEE International Symposium on Mixed and Augmented Reality*, pp. 201–202. IEEE, 2009.

[16] Y. Liu, H. Dong, L. Zhang, and A. El Saddik. Technical evaluation of HoloLens for multimedia: A first look. *IEEE MultiMedia*, 25(4):8–18, 2018.

[17] C. L. MacKenzie. From manipulation to goal-directed human activities in virtual and augmented environments. In *ICAT '99 Proceedings of the Ninth International Conference of Artificial Reality and Telexistence*, pp. 6–8. The Virtual Reality Society of Japan, Tokyo, 1999.

[18] Microsoft. Microsoft - HoloLens 2. Retrieved November 4, 2021 from https://www.microsoft.com/en-us/hololens.

[19] NaturalPoint Inc. DBA OptiTrack. OptiTrack - Motion Capture Systems. Retrieved November 4, 2021 from https://optitrack.com.

[20] C. B. Owen, J. Zhou, A. Tang, and F. Xiao. Display-relative calibration for optical see-through head-mounted displays. In *Third IEEE and ACM International Symposium on Mixed and Augmented Reality*, pp. 70–78. IEEE, 2004.

[21] S. Putze, D. Alexandrovsky, F. Putze, S. Höffner, J. D. Smeddinck, and R. Malaka. Breaking the experience: Effects of questionnaires in VR user studies. In *CHI '20 Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–15. ACM, 2020.

[22] V. Schwind, P. Knierim, N. Haas, and N. Henze. Using presence questionnaires in virtual reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–12. ACM, 2019.

[23] A. L. Simeone, E. Velloso, and H. Gellersen. Substitutional reality: Using the physical environment to design virtual reality experiences. In *CHI '15: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, p. 3307–3316. ACM, 2015.

[24] P. Tuddenham, D. Kirk, and S. Izadi. Graspables revisited: Multi-touch vs. tangible input for tabletop displays in acquisition and manipulation tasks. In *CHI '10 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2223–2232. ACM, 2010.

[25] Unity Technologies. Unity Real-Time Development Platform. Retrieved November 4, 2021 from https://unity.com/.

[26] D. Van Krevelen and R. Poelman. A survey of augmented reality technologies, applications and limitations. *International Journal of Virtual Reality*, 9(2):1–20, 2010.

[27] G. P. Viganò. M2MQTT for Unity. Retrieved November 4, 2021 from https://github.com/gpvigano/M2MqttUnity.

[28] E. Williams. Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Chemistry*, 2(2):149–168, 1949.