Investigation of 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

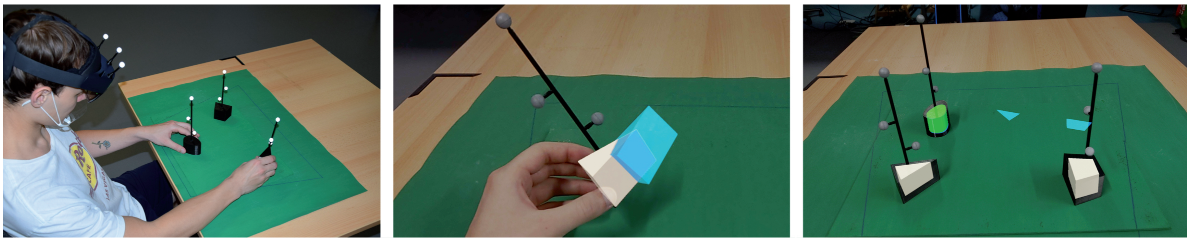

Figure 1: Study overview. Left: Participant completing a task. Middle: HoloLens view of study part 1: Matching virtual object to 3D target (size condition M). Right: HoloLens view of study part 2: Egg-shaped cylinder placed on 2D target (size condition XS).

Abstract

Optical see-through AR headsets are becoming increasingly attractive for many applications. Interaction with the virtual content is usually achieved via hand gestures or with controllers. A more seamless interaction between the real and virtual world can be achieved by using tangible objects to manipulate the virtual content. Instead of interacting with detailed physical replicas, working with abstractions allows a single physical object to represent a variety of virtual objects. These abstractions would differ from their virtual representations in shape, size, texture and material. This paper investigates for the first time in optical see-through AR whether size variations are possible without major losses in performance, usability and immersion. The conducted study shows that size can be varied within a limited range without significantly affecting task completion times as well as feelings of disturbance and presence. Stronger size deviations are possible for physical objects smaller than the virtual object than for larger physical objects.

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

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

1 Introduction

Augmented Reality (AR) provides a connection between the real and virtual world by displaying virtual objects in the real environment as an overlay in the user's field of view [2, 4, 31]. The technology is already widely used in many areas such as medical applications [32], education [7] and architectural and urban design [26]. In addition to the classic video see-through AR smartphone applications, there is an increasing number of optical see-through AR applications, due to the growing number of AR headsets on the market. In contrast to video see-through AR, optical see-through AR allows an almost unobstructed view of the real world and thus enables new applications [24]. This is especially useful when people are working or playing together and want to interact with each other and experience each other's reactions. Besides that, a complete coverage of physical objects cannot be achieved with optical see-through AR due to technical limitations. The overlays are always slightly translucent and their opacity strongly depends on the lighting conditions of the user's surroundings, which is why findings for video see-through AR cannot directly be transferred to optical see-through AR.

Wearing AR headsets enables interaction with physical objects using two free hands. When digital information is coupled to physical objects [16], manipulation tasks on the virtual objects can be performed faster [3] and more accurately [30] when interacting with the tangibles. The use of physical props to manipulate or rearrange virtual content in AR is known as Tangible Augmented Reality (TAR). TAR enables intuitive and natural interaction [5] and opens up even more use cases [8]. This advantage can also be used in applications for optical see-through AR headsets. In most applications for AR headsets the interaction takes place mainly via controllers, hand gestures or speech [6, 18, 21]. Many AR applications are primarily used for viewing 3D objects, e.g. in construction [15]. By interacting with physical objects one could not only view 3D visualizations, but also jointly create or modify buildings. The use of tangible objects in AR games, e.g. in the form of touchable game figures, would make the gaming experience even more realistic [14].

To make interaction in a variety of applications more intuitive by using tangibles, a huge amount of physical props could be required. Since it is not possible to create and store exact physical replications for e.g. all available game pieces in all games, it is necessary to use more abstract physical objects, which can represent a larger group of virtual objects. Ideally, every user would have a set of different abstract props at home, which could be used for every imaginable use case. In order to determine what shapes and sizes such a set would need to include, it is first necessary to investigate how large the discrepancy between the virtual and physical object regarding size, shape, texture and material may be when using abstract props instead of exact replicas.

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

This is the authors' version of the article. See https://doi.org/10.1109/ISMAR52148.2021.00029 for the final version of record in IEEE Xplore.

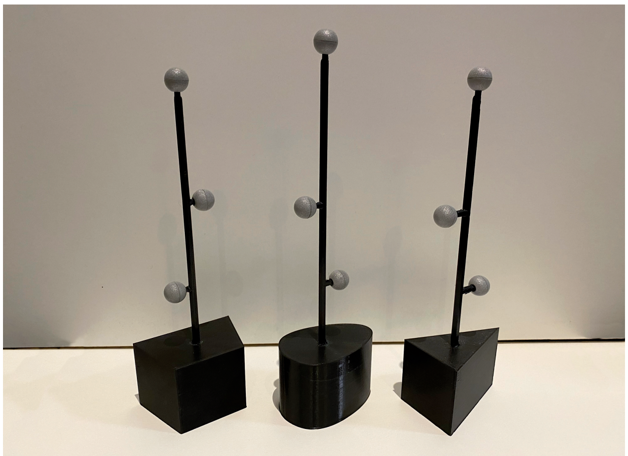

Figure 2: Physical props with attached marker trees on top.

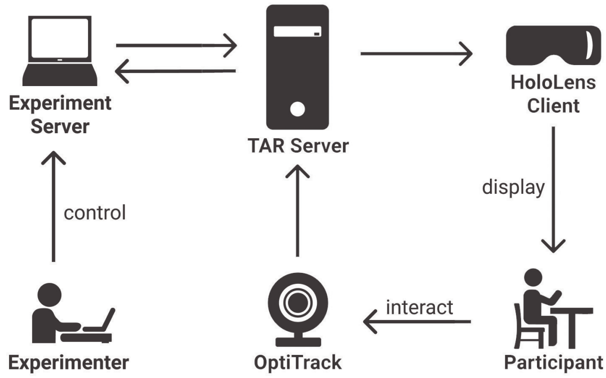

Figure 3: Architecture of the components used for the study.

In a first step, this paper investigates the factor of size differences between the tangible object the user is interacting with and the overlying virtual representation in terms of performance, usability and immersion in optical see-through TAR. In a study, we explore to what extent the size of the physical object can vary from the size of the virtual overlay without a significant worsening of execution times, feeling of disturbance and feeling of presence.

2 Related Work

To take advantage of the benefits of Tangible Augmented Reality in optical see-through applications, physical elements are required for interaction. Since it is unrealistic to use a separate set of props for each application [9, 13, 29], it is necessary to find out how to represent a variety of virtual objects with one physical object. Hettiarachchi and Wigdor developed an approach in which everyday objects are used as tangible objects [13]. Their system searches for physical objects in the environment that resemble the virtual objects and overlays them with the virtual models in AR. Szemenyei and Vajda also developed algorithms that enable automatic matching of everyday physical objects and virtual objects [27, 28]. In the systems presented, matching is only based on shape; the size of the objects is not taken into account. In the case of Hettiarachchi and Wigdor the size of the virtual object is even automatically adjusted to the size of the physical object. In many use cases, however, resizing the virtual objects is not feasible. This is, for example, the case when interacting simultaneously with several objects, whose size ratio plays a role, e.g. when interacting with construction or game pieces.

When everyday objects are used for interaction, there is not only a discrepancy in size with the virtual representation, but also a mismatch in many other factors. Simeone et al. investigated how large this discrepancy between physical proxy and virtual element can be designed to be in Virtual Reality (VR) without breaking the VR illusion [25]. Different substitution levels were considered: exactly matched virtual replicas, virtual models with aesthetic differences, models where a part was added or omitted, functionally different models and virtual models with categorical differences, where there is no longer a connection between physical and virtual object. They found that differences in shape and visualized temperature seem significantly less credible than an exact replica. The same applied to substitutions with smaller virtual objects, while they found no significant difference in believability for larger virtual representations.

Size variations in VR were also studied by de Tinguy et al. [9]. Their focus was to find out how similar virtual and physical objects must be in order to feel the same. Besides variations in width, local orientation and curvature were also considered. The largest possible discrepancies that remain unnoticed by the user were determined in terms of local curvature. However, the results also show that in VR it is possible to vary the width of objects up to 5.75% without the user noticing any difference.

The first investigations regarding size and shape variations in video see-through AR were carried out by Kwon et al. [17]. They conducted a study with three tangible objects that differed strongly in size and shape. They found that performance was best when the size and shape of the virtual object matched the physical object. The biggest time differences were found during grasping, while the subsequent manipulation time did not differ significantly. In an additional experiment, where only the size of the objects was varied, they found no significant performance differences between the five chosen size conditions. Therefore, they concluded that the results of the main study were due to the shape differences.

In optical see-through Augmented Reality, to our knowledge, no investigations have yet been made into size variations between physical and virtual objects. Ahn et al. found that there is a difference in size perception depending on the AR visualization technique used [1]. Their results show that the most precise and fastest size perception is possible with video-see through AR, while mobile displays caused serious overestimation and optical see-through underestimation of object size. A separate investigation of size variations in optical see-through AR is therefore necessary.

Optical see-through AR differs from video see-through AR, apart from having a less obstructed view of reality, mainly because the overlays are slightly translucent and one can still see the objects behind them. Investigations into how much the size can vary without being noticed by the user, as conducted by de Tinguy et al. [9] in VR, therefore probably do not make as much sense in optical see-through AR because the physical object always remains a bit visible.

In this paper, similar to Simeone et al. [25] in VR, the focus is on investigating the effect of size differences on the perception of reality and ease of use. In addition, the effect of size variations on performance (see [17]) and on size estimates (see [9]) is considered in optical see-through Tangible Augmented Reality.

3 Technical Implementation

In order to investigate the influence of size variations between physical and virtual objects in Tangible Augmented Reality, a suitable setup is required that allows one to track the tangible props and overlay them with virtual content. For tracking we use the OptiTrack system [22], which determines the position and orientation of an object using reflective markers. When many cameras are used, it is highly resistant to occlusion, unlike methods based on tracking image targets, such as Vuforia [23], where occlusion often occurs during interaction with the props. Due to this problem and the time

This is the authors' version of the article. See https://doi.org/10.1109/ISMAR52148.2021.00029 for the final version of record in IEEE Xplore.

lag between target recognition and display of the overlay in the HoloLens, these systems are less suitable for investigating discrepancies between physical and virtual objects, since there would be constant interruptions in the display of the overlays. We use the HoloLens 2 [20] as the optical see-through display because it has a large field of view in which the interactions can be performed. The flip-up display also has the advantage that the headset does not have to be taken off every time questionnaires need to be filled out.

Since OptiTrack and HoloLens have different coordinate systems, it is necessary to track the AR headset in addition to the physical objects being interacted with, in order to be able to display the virtual objects at the correct position in the AR view. To track the props we use marker trees (see Fig. 2) that are placed in the middle of the upper side of the objects. In order to ensure that they would disturb the interaction as little as possible and that the tracking would be stable when interacting with multiple objects, they were built to be tall and narrow. The height of the trees results from the fact that there must be a certain distance between the markers so that they can be tracked individually by OptiTrack. Due to the height of the trees, they shake a lot, especially when the props are placed on the table. A smoothing algorithm was applied to minimize the wobbling of the visual overlays.

The HoloLens 2 was also equipped with markers for tracking (see Fig. 1, left), similar to the work of Liu et al. [18]. These are located very close to the glasses, so there are fewer problems caused by shaking, but they are still mounted high enough to be seen by most cameras at the same time. The Optitrack system determines the exact position of the HoloLens and the tangible prop in space and thus enables us to display the corresponding virtual object at the right place in the AR view at any time.

Our developed system essentially consists of three components: HoloLens Client, TAR Server and Experiment Server (see Fig. 3). The Experiment Server is a tool to conduct studies. It enables the reading of predefined task sequences, lets the experimenter control the experimental procedure and transmits the current task via TCP to the TAR Server, which is the central component in this architecture. The TAR Server is a Unity application receiving not only commands from the Experiment Server but also data from the motion capture system. OptiTrack's tracking software Motive receives data from the tracking cameras, calculates object poses and streams them to the TAR Server application running on the same computer. The TAR Server performs data corrections such as smoothing, and integrates the available information about tracked objects as well as virtual target objects into a Unity scene always reflecting the current state of the experiment. This allows all relevant computations (like checking for a completed task or smoothing input data) to be performed in this central component, which continuously updates the HoloLens Client with all necessary data for the visualization of the virtual objects. This HoloLens 2 application was also developed with Unity and acts solely as a display accessed via a single TCP connection to optimize performance. It listens for commands like changing the virtual object size as well as updates to the states of the defined rigidbodies (including the camera, virtual objects attached to physical props and purely virtual target objects) in position, rotation and visibility.

4 STUDY

A user study was conducted to find out whether it is feasible to use smaller or larger props as interactive elements for a virtual object. The main focus was to find out if it is possible to use a larger/smaller tangible prop compared to the virtual object without extreme losses in usability and if there is a range within which presence is felt to be almost the same. Furthermore, we wished to test whether differences in size, as in video see-through AR [17], have no effect on performance and whether, as assumed, the size conditions would be correctly assessed by the participants. Therefore, the following hypotheses were made:

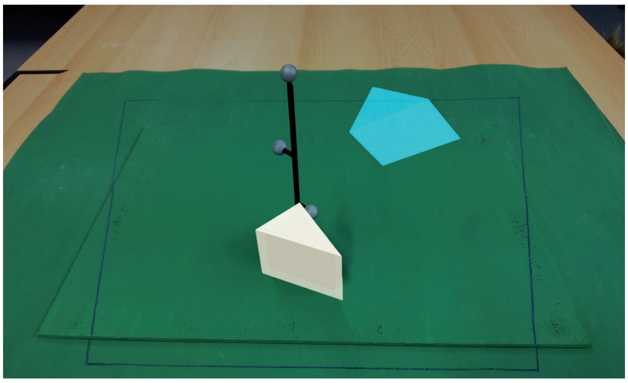

Figure 4: Participant's perspective for a task in part 1 of the study with size condition XL: Fitting virtual overlay (white) to 3D target (blue).

- H1: The size of the virtual and physical object can differ within a certain range without significant loss in usability.
- H2: The size of the virtual and physical object can differ within a certain range without significant worsening of "AR Presence" and "TAR Presence".
- H3: Differences in size between virtual and physical object have no influence on performance.
- H4: Differences in size between virtual and physical objects can be estimated correctly by the participants.

The study was divided into two parts. The first part of the study was exploratory, so that the participants could observe and feel differences without time pressure and could become familiar with the interaction in Tangible Augmented Reality. During this exploration phase, the participants had the task of successively fitting virtual objects represented as overlays on the physical prop to virtual 3D targets (see Fig. 4). This requires the objects to be arranged in 6 degrees of freedom, so maximum interaction is necessary to align the overlays, giving the participants the opportunity to better perceive the interaction with the objects.

In the second part of the study, we additionally wanted to find out if a difference in size between a virtual object and a physical object has an impact on performance. Therefore, we let the participants solve puzzle tasks under time pressure. For this task, three different objects had to be placed on corresponding visualized 2D targets on a plate. We decided to have participants interact with multiple objects, so that the influence of disturbances during grasping is increased [17]. Since simultaneous interaction with 3 objects in 3D is not possible with two hands, 2D targets were used in this task.

In both cases, docking tasks were chosen that require grasping, rotating, and arranging the objects. These tasks – even if they seem simple – represent basic elements in complex goal-oriented activities [19]. Regardless of the use case a physical prop is used for, this tangible object is always grasped, lifted, turned and placed, whether it is e.g. a game piece on a virtual board or a piece to configure a composite object.

The study was approved by the ethical review board of our faculty. In order to conduct the study, a special hygiene concept was developed, which was approved by our emergency and crisis manager.

4.1 Participants

14 volunteers (9 male, 5 female) aged between 21 and 28 ($M = 24.5, SD = 2.279$) were recruited to participate in the study. All had normal or corrected-to-normal vision and 12 were right-handed. Participants were asked about their prior experience with AR in general

Figure 5: Study setup: Participant's interaction area in the middle of the tracking zone.

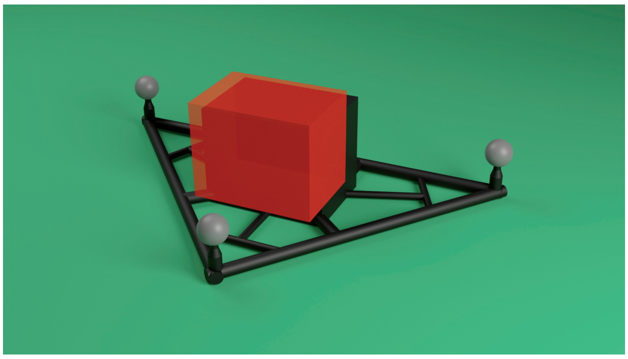

Figure 6: Physical example object for participant's eye calibration. The shifted overlay (red box) is adjusted until it fits the black box.

as well as AR glasses on a 7-point Likert scale from 1 (= never) to 7 (= regular). They reported mostly low experience with AR ($M = 2.214, SD = 1.762$) and minimal experience with AR glasses ($M = 1.571, SD = 1.089$).

4.2 Apparatus

The study took place in a quiet laboratory environment, which was darkened and only indirectly illuminated by two softbox studio lamps to avoid the influence of different lighting conditions. Participants' heads and physical props were tracked through a combination of 11 OptiTrack Flex 3 cameras. All cameras were mounted on a truss at a height of about 2.6m and aimed at the center of the floor of the tracking area. Participants sat at a table located at the center of the OptiTrack cameras (see Fig. 5). Furthermore, a desk placed outside of the tracking area was used for the secret arrangement of the physical props on a plate.

The tasks were performed on a monochrome green background. This background was chosen to be larger than the plate on which the objects were placed, so that the same background was always visible behind the objects during the interaction. This ensured that the objects would stand out equally well against the background in every situation. The physical props were black and equipped with marker trees on top for tracking (see Fig. 2).

For visualizing the virtual overlays and the targets, a HoloLens 2 was used. The color of the overlays was set to white, which is the least translucent color on the HoloLens. Additionally, the opacity of the overlays was set to 100% and the brightness of the HoloLens 2 to maximum, to achieve the lowest possible translucency of the overlays. The distance between the chair and the desk was constant and the distance between the markers on the HoloLens 2 and the desk was adjusted to 45cm to guarantee a similar viewing angle for all participants.

4.3 Basic Approach

At the beginning of each study, the HoloLens was first adapted to the eyes of the respective participant. This is necessary because each person has a different depth perception, e.g., due to their interpupillary distance. The adaptation took place in a two-step process. First, the eye calibration of the HoloLens 2 was performed by the participants. Subsequently, the calibration was checked again by means of an example object. For this purpose, the participants were given a black cuboid over which a virtual red cuboid was superimposed (see Fig. 6). If the virtual object was not exactly at the position of the black prop, a manual fine calibration was performed until the overlay fitted correctly. This manual calibration was necessary mainly if the participants wore thick glasses or had a very large interpupillary distance. By performing the eye calibration, it was possible to ensure that the overlays were displayed at the correct position during the study regardless of where the physical object was located.

The execution of individual tasks always followed the same procedure. First, participants were asked to adjust their HoloLens' field of view by matching a white frame in the AR display with a blue marked area on the table. Afterwards, participants remained in this position to make sure that the overlays on the physical objects were inside the field of view and thereby visible at all times. During this calibration procedure, the physical props were arranged on a plate out of the participant's view and covered with a box. The plate was then put on a designated spot on the table in front of the participant. Not until the box was removed, the overlays on the objects and virtual targets became visible and the participant could start with the respective task.

After completion of each condition, three questionnaires had to be completed: an AR presence questionnaire, a TAR presence questionnaire, and a size perception questionnaire. The first questionnaire, which looks at presence in AR, examines how real the overlays looked. The second questionnaire focuses on presence in TAR and evaluates how realistic the interaction with the overlays on the tangible objects felt. Both questionnaires consist of 4 questions each based on questions measuring presence in VR environments. The size perception questionnaire contains questions that focus on the perception of size differences and perceived disturbance. All questionnaires were rated using 7-point Likert scales. For example, size was assessed by asking participants to rate the size of the virtual object compared to the physical object from 1 (= much smaller) to 7 (= much larger). By using a proprietary questionnaire instead of a standard usability questionnaire such as NASA-TLX [12], it was possible to specifically examine how the interference was perceived when grasping and interacting with the object. We deliberately refrained from additionally measuring usability with NASA-TLX in order to keep the amount of work as low as possible, since the questionnaires had to be filled out 14 times by each participant. Lastly, the participants answered a final questionnaire. Here, demographic information was requested in addition to a classification of the size ratios based on performance and usability.

4.4 Design and Procedure

The study was designed as a within-subject experiment. In total, 7 different size conditions were tested. The order of the size conditions in both parts of the study was counterbalanced by a Williams design latin square (LS) of size 7 [34]. Figure 7 shows the size variations of the virtual overlay. Condition M represents the baseline where virtual and physical object have an equal length of 6cm. Sizes S and L portray a small size variation with 10% difference in length; width and height are always scaled by the same factor as length. Following are conditions XS and XL with a size variation of 25%,

This is the authors' version of the article. See https://doi.org/10.1109/ISMAR52148.2021.00029 for the final version of record in IEEE Xplore.

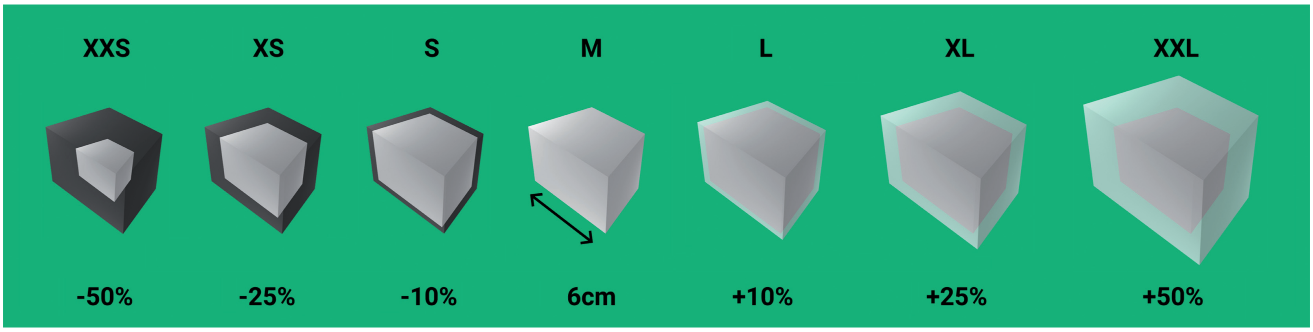

Figure 7: Size variations of the virtual overlays (white) compared to the physical proxy objects (black). Condition M is the base condition with matching size of virtual and physical object.

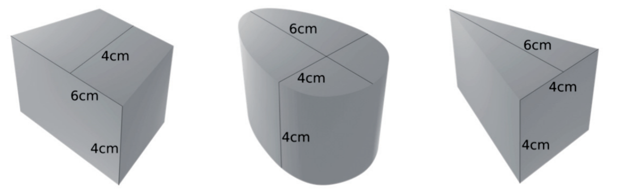

Figure 8: Shapes of the tangible objects. Left to right: trapezoidal prism, egg-shaped cylinder and triangular prism.

as well as XXS and XXL with a 50% size difference from the physical object. The small size difference (S and L) was chosen to find out if only small size variations are possible without serious losses in the measured values. We used a minimal size difference of 10% instead of 5% in contrast to de Tinguy et al. [9] because our research interest was different. Instead of investigating at what point the user notices the change, we wanted to find out how much we can vary size without causing a significant degradation in usability and performance. Based on a pre-test with four persons, 5% was expected not to cause such effects. The maximum size difference (XXS and XXL) was chosen to be accordingly large (+/-50%), in order to find the limit of possible size variations.

In the study, three different 3D printed shapes were used, with which participants had to interact (see Fig. 8). These shapes were intended to be different basic shapes, which create a distinct feeling when touching and interacting with them. Instead of common bases like an equilateral triangle, square and circle, we purposely modified them to guarantee that there is only one possibility to match a given target. Participants thereby have to perform a maximal rotation (up to 180°) of the physical prop. Our aim was to provoke more interaction with the objects and give participants the opportunity to perceive the influence of size variations. We chose a length of 6cm and width of 4cm because this size can be easily grasped [11]. Furthermore, we oriented our design to existing investigations in VR and video see-through AR to produce comparable results [9, 17]. The virtual overlays of different sizes were placed so that their center was aligned with the center of the physical props.

4.4.1 Part 1

In part 1 of the study, participants interacted with the different shapes in sequence. The order of the interaction with the three different shapes was counterbalanced by a 3x3 LS for each condition. There were 6 possible positions where the 3D targets could be placed, all with equal distance to the initial position of the physical prop. The selection of the position was balanced by a 6x3 LS for each condition and prop shape. The orientation of the physical objects on the plate as well as the rotation of the 3D targets were determined randomly. However, for the targets, only rotations which obey the following rules were considered: The upward normal vector of the 3D target must not point downwards or be too close to pointing sideways, and the upward normal vector must not form an angle with the vector to the viewer's eyes which is too close to 90°. These rules ensure that all targets are solvable without head movements, as this would require more time to solve the tasks. Furthermore, they ensure that the physical props do not have to be flipped, as this is not possible due to the marker trees on top. To prevent ambiguity, the undersides of the virtual props and the virtual targets were colored orange to make them distinguishable from the top side, which was communicated to the participants at the beginning.

For each of the seven size conditions, all three shapes were interacted with successively. The task was to match the displayed virtual object to a 3D target object of the same shape and size in position and orientation (see Fig. 1 middle and Fig. 4). Once this was achieved accurately enough, the overlay temporarily turned green and the next target was displayed immediately. For every shape in every size condition, 6 targets had to be matched, one after the other. Once the last of these targets was matched, the overlay stayed green and no new targets appeared. A matching was determined "solved" exactly when errors below a threshold of 1cm in distance and 30° in angle between prop and target were detected consistently for 0.5 seconds. In a pre-test we found that these values provide the best mix of feasibility and complexity.

4.4.2 Part 2

In part 2 of the study, participants could interact with all objects at once. In this task, there were three different positions at which targets were placed. Thus there was a total of 6 different arrangements for the three props. These arrangements as well as the orientations of the individual targets were randomized. Likewise, the initial arrangement of the physical props and their initial orientation on the plate was random. The 2D targets were displayed on the plate when the task was started. Their height in 3D space was adjusted so that they were at the same height as the bottom of the virtual overlays. Visually, however, from the participants' point of view, it still appeared as if the targets were lying on the table.

For each of the seven size conditions, two puzzle tasks had to be solved. The task was to place the virtual objects as quickly as possible onto the displayed virtual 2D targets on the table (see Fig. 1, right). Before each task, the props were arranged on the plate out of the participant's view and covered with the box. Measurement of task completion time started automatically once the box was removed and thus simultaneously with the display of virtual objects and targets. As soon as an object was placed and oriented correctly, its overlay color changed to green (see Fig. 1, right). Once all objects

This is the authors' version of the article. See https://doi.org/10.1109/ISMAR52148.2021.00029 for the final version of record in IEEE Xplore.

were placed and oriented correctly, the task was considered solved and time measurement stopped automatically. A placement was determined as "solved" exactly when errors below a threshold of 0.5cm in flat distance, 1cm in height and 7.5° in angle between prop and target were detected consistently for 0.5 seconds. These values were also determined with test participants.

We decided to use well-defined deviations in distances and rotations as a stopping criterion for the task instead of performing an evaluation with regard to the error distance and error rotation because the evaluation of the task completion time was important to us. Letting participants self-assess whether a task was solved would have greatly affected the evaluation of performance and led to uncertain study durations, as some individuals are inherently more accurate than others.

4.5 Results

We investigated the effect of size variations between a physical object and a corresponding virtual overlay on the usability (by disturbance ratings when grasping and interacting with objects), on the feeling of presence (by AR and TAR presence ratings), on the size perception (by estimates of the virtual object size compared to the physical object) and on performance (by task completion time). We evaluated these four types of results for both parts of the study individually (except for task completion time, which was only measured in part 2) using the following procedure: First we checked for the overall effect of size condition on the measured result using a Friedman test with a fixed significance level of $\alpha = 0.05$ and 6 degrees of freedom. When significant effects were revealed, we conducted post-hoc tests using Wilcoxon's signed-rank test again with a fixed significance level of $\alpha = 0.05$ and 13 degrees of freedom to find which size conditions differed from the size matching condition M, which we set as our baseline condition. In addition to the resulting p-value, the matched pairs rank-biserial correlation r is given as an effect size. Figure 9 summarizes our results and highlights which conditions were found not to differ significantly from the size matching condition M. However, this does not imply equality of such conditions.

4.5.1 Disturbance

In part 1 of the study, the Friedman test indicated a significant influence of size condition on the scores of disturbance for grasping ($\chi^2 = 42.405, p < 0.001$) and for interaction with the objects ($\chi^2 = 48.129, p < 0.001$). For grasping, Wilcoxon's signed rank test revealed significant differences for XXS compared to M ($W = 0, p = 0.002, r = 1$), XS compared to M ($W = 8, p = 0.014, r = 0.795$) and XXL compared to M ($W = 7.5, p = 0.02213, r = 0.773$). Conditions S, L and XL did not differ significantly from M in their grasping disturbance scores. Similarly for interaction, the post-hoc tests showed significant differences for XXS compared to M ($W = 0, p = 0.002, r = 1$), XS compared to M ($W = 0, p = 0.009, r = 1$) and XXL compared to M ($W = 7, p = 0.036, r = 0.745$). Again, for conditions S, L and XL, no significant difference in interaction disturbance was detected compared to baseline condition M.

In part 2 of the study, the Friedman test also indicated a significant influence of size condition on the scores of disturbance for grasping ($\chi^2 = 41.196, p < 0.001$) and for interaction with the objects ($\chi^2 = 31.676, p < 0.001$). For grasping, Wilcoxon's signed rank test revealed significant differences for XXS compared to M ($W = 0, p = 0.002, r = 1$), XS compared to M ($W = 0, p = 0.013, r = 1$), S compared to M ($W = 4, p = 0.025, r = 0.822$) and XXL compared to M ($W = 0, p = 0.021, r = 1$). For conditions L and XL, no negative influence could be identified. Similarly for interaction, the post-hoc tests only showed significant differences for XXS compared to M ($W = 1.5, p = 0.002, r = 0.967$), XS compared to M ($W = 2, p = 0.006, r = 0.939$) and XXL compared to M ($W = 2.5, p = 0.033, r = 0.861$).

Therefore we can conclude a significant effect of size variation on disturbance during grasping and interaction. For grasping, conditions XXS and XS with large and medium size reduction result in significantly higher disturbance scores, followed by condition XXL with a large size increase and smaller effect. In part 2 of the study, even a small size reduction (condition S) led to such an effect. For interaction, conditions XXS and XS with large or medium size reduction also show significantly increased disturbance, and again the only condition with increased virtual object size having this effect was XXL. These findings support hypothesis H1 that a physical object can differ within a certain range from its virtual representation without significant loss in usability.

4.5.2 Presence

In part 1 of the study, the Friedman test indicated a significant influence of size condition on the scores of AR ($\chi^2 = 30.296, p < 0.001$) and TAR presence ($\chi^2 = 22.266, p = 0.001$). For AR presence, Wilcoxon's signed-rank test revealed significant differences only for XXS compared to M ($W = 3, p < 0.001, r = -0.943$) while for TAR presence significant differences could be found for XXS compared to M ($W = 13.5, p = 0.011, r = -0.743$) as well as XS compared to M ($W = 8, p = 0.003, r = -0.848$).

In part 2 of the study, the Friedman test also indicated a significant influence of size condition on the scores of AR ($\chi^2 = 33.468, p < 0.001$) and TAR presence ($\chi^2 = 24.752, p < 0.001$). For AR presence, Wilcoxon's signed-rank test revealed significant differences for XXS compared to M ($W = 5.5, p = 0.001, r = -0.895$) and XS compared to M ($W = 10, p = 0.014, r = -0.78$). However, for TAR presence significant differences could only be found for XXS compared to M ($W = 8.5, p = 0.011, r = -0.813$).

These results support hypothesis H2 that the size can be adjusted to a certain extent without significant worsening of presence. For Condition XXS with a large size reduction, a significant worsening was found in both parts of the study and for both types of presence assessed, while for condition XS with a medium size reduction, a significant worsening could only be found for TAR presence in part 1 and AR presence in part 2. Enlargements of the virtual objects (conditions L, XL and XXL) or only a slight size reduction (condition S) did not lead to significantly lower presence scores.

4.5.3 Size Estimate

The participants estimated the size of the virtual object compared to the size of the physical object on a 7-point Likert scale. Therefore, we can analyze the effect of the actual size condition on the participants' size perception.

Friedman tests indicated a significant influence of size condition on the perceived size in part 1 ($\chi^2 = 69.361, p < 0.001$) and part 2 ($\chi^2 = 75.783, p < 0.001$) of the study.

Wilcoxon's signed-rank test as post-hoc revealed that in part 1, only conditions S ($W = 3, p = 0.233, r = -0.6$) and L ($W = 2, p = 0.773, r = 0.333$) had no significant differences in the size estimate compared to M as a baseline, whereas in part 2, all size conditions differed significantly in their estimate from the baseline.

These values show that for small size differences (condition S and L), the size could not always be correctly estimated. Thus they disprove hypothesis H4 that the difference in size between physical and virtual object can always be correctly estimated.

4.5.4 Task Completion Time

The results of the time measurements in part 2 of the study are displayed in Fig. 10. The Friedman test indicated a significant influence of size condition on task completion time overall ($\chi^2 = 14.082, p = 0.029$). However, Wilcoxon's signed-rank test revealed significant differences only between the conditions with size variation, and

This is the authors' version of the article. See https://doi.org/10.1109/ISMAR52148.2021.00029 for the final version of record in IEEE Xplore.

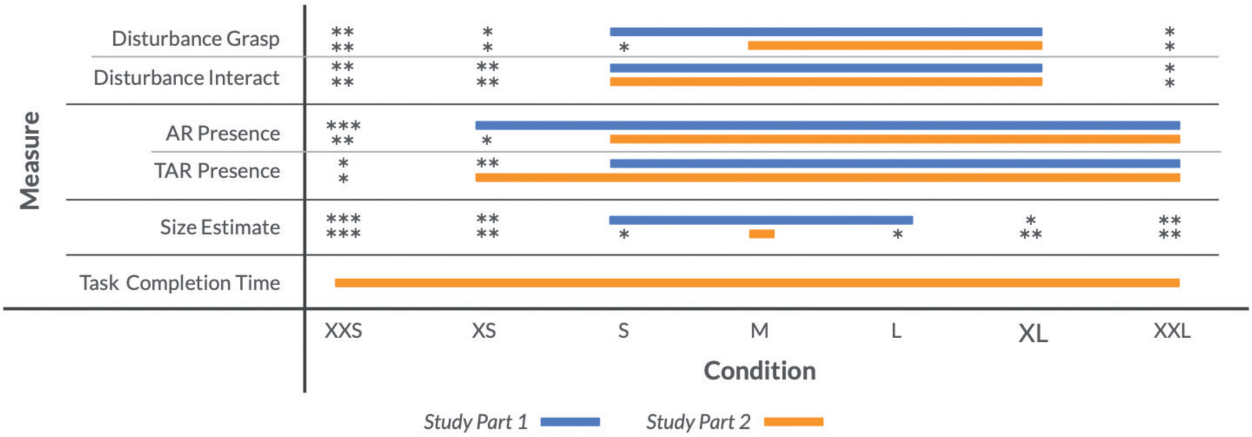

Figure 9: Overview of significant differences from the baseline condition M marked with * ($p < 0.05$), ** ($p < 0.01$) and *** ($p < 0.001$). Ranges of size conditions without significant difference from baseline condition M are represented as blue lines for part 1 and orange lines for part 2.

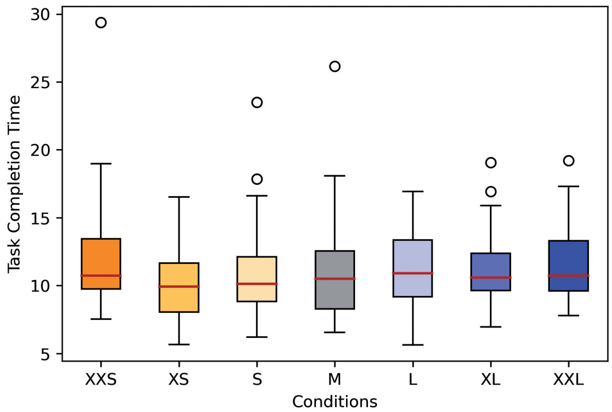

Figure 10: Task completion times in seconds for each size condition in part 2 of the study. No significant difference was found for any condition compared to baseline condition M.

none compared to the size matching condition M. Therefore, confirming our hypothesis H3, no significant effects of size variation on performance can be derived.

4.5.5 Final Questionnaire Results

Besides the dependent measures reported above, each participant was asked in the concluding questionnaire to rank the seven size conditions with respect to perceived realism and perceived easiness. Table 1 shows the cumulative sum of the scores of all participants for the seven conditions. The highest valued condition is given 7 points, the second 6 points and finally the lowest valued condition 1 point each in the sum. Consistent with the evaluation of the AR and TAR presence scores, conditions M, L, XL and S were ranked highest in descending order in perceived realism. Regarding easiness, the order is identical to that of realism.

In addition, the participants had to indicate in their rankings up to which state the conditions feel pleasant and when they change to unpleasant (realism ranking) and up to which state the conditions feel efficient and when they start feeling inefficient (easiness ranking). Due to an error in filling out the questionnaire, one participant had to be excluded. Table 2 shows that conditions M, L, S and XL were rated mostly pleasant in descending order. This matches the

Table 1: Scores for each size condition in realism and easiness according to participants' rankings.

	XXS	XS	S	M	L	XL	XXL
Realism	20	35	59	87	83	65	43
Easiness	29	39	59	78	76	64	47

Table 2: Number of participants' classifications for each size condition as pleasant and efficient (out of 13).

	XXS	XS	S	M	L	XL	XXL
pleasant	2	2	10	13	12	8	6
efficient	3	4	9	11	12	9	5

evaluation of the disturbance scores, which showed higher disturbance with all other conditions compared to M. Condition M is the only condition which everyone agreed to be pleasant. Regarding efficiency there is a tendency towards conditions with larger virtual objects rather than smaller ones. Here L, M, XL and S were rated mostly efficient in descending order.

The participants also had the opportunity to submit comments on the study in a free text field. Three of them mentioned that the interaction feels more real, is easier, or is less disturbing when larger virtual objects are used, as these cover the physical objects. Another three participants pointed out that there is a certain delay between the actual hand movements and the movement of the overlay, especially in fast movements.

Sickness after the experiment was rated on a scale from 1 (= not at all) to 7 (= very sick) as low ($M = 1.357, SD = 0.633$) with a maximum of 3 by one person.

5 Discussion

Our results show that size differences between the tangible prop and the virtual object it represents are feasible within a certain size range without worsening the feeling of disturbance and the feeling of presence. Figure 9 shows for which conditions no significant difference from the baseline M could be detected.

For our setup, the results for both parts of the study are very similar. The strongest difference can be seen in the area of size estimation of the virtual object compared to the physical object. Contrary to our expectations, participants had difficulties estimating the size of the virtual overlay compared to the size of the physi-

This is the authors' version of the article. See https://doi.org/10.1109/ISMAR52148.2021.00029 for the final version of record in IEEE Xplore.

cal object in part 1 of the study. A reason for this might be that the overlays had a strong covering effect and the physical objects were therefore almost not perceived by the participants. While in part 1 of the study the subjects were unable to detect the difference between condition S and M or L and M, they were able to do so in part 2 of the study. This can be explained by the fact that each participant first worked on the exploration task (part 1) and thus already knew what the different sizes were when performing part 2. This made it possible to estimate the difference better than in part 1, where subjects were sometimes first shown condition S or L before condition M and they may have initially incorrectly considered it to be the matching condition. From this it can be seen that small size differences cannot be reliably detected if there is no knowledge about other better matching objects. This contradicts our hypothesis H4 that size differences can be correctly estimated if the physical objects are visible to a certain degree due to technical conditions.

The learning effect regarding sizes could also explain the different spans in disturbance during grasping. In part 1 of the study, no significant differences from baseline M were found in the range from condition S (-10%) to XL (+25%), whereas the range in part 2 was only from M (baseline) to XL (+25%). If an object is not perceived as larger or smaller than the baseline, it is more likely that the sensation of grasping for these conditions will not be judged as differently either. However, once one is aware of the size differences, this will possibly affect the evaluation.

Disturbance during interaction was not significantly distinguishable from baseline M for both parts of the study in a range from S (-10%) to XL (+25%). The assumption is that the knowledge about the sizes did not have such a strong effect here, because the difficulty is mainly in grasping. As soon as you hold the object in your hand, the difference in size is less of an issue.

The result concerning AR presence in our study setup shows that it was possible to increase the size by at least 50% (condition XXL) and to decrease it by up to 25% (condition XS, part 1) or 10% (condition S, part 2).

Regarding TAR presence, deviations in size were also feasible. The range is from -10% (condition S, part 1) or -25% (condition XS, part 2) up to +50% (condition XXL).

Concerning the time for the completion of the tasks, no negative influence of the size conditions could be detected compared to the baseline condition M. Therefore, the differences in the size of the virtual and physical object do not affect performance, with respect to the conditions investigated in the study.

Overall, it can be observed that the results are similar to comparable studies in VR and video see-through AR. For example, de Tinguy et al. [9], who measured in VR how much a virtual object can be resized without the user noticing the change in size, found that size changes can be made in a small range without being noticed. Regarding usability (disturbance in grasping/interacting) and presence (AR/TAR presence) it can be seen that the virtual object can be considerably larger than smaller compared to the physical one without having a negative impact on usability and presence. This fits with the results of Simeone et al. [25], who compared only three different virtual sizes in VR (replica, -50%, +50%), but showed that a significant deterioration of believability and ease of use was only found for the smaller virtual representation. The result regarding task completion time agrees with that of Kwon et al. [17], who tested the impact of size differences between virtual and physical objects in video see-through AR on performance. The similarity of the results of the studies in VR and video see-through AR to the results of this study can be explained by the decision to use overlays that are as opaque as possible, which is why the underlying physical objects could hardly be perceived.

The delay between movement of the physical object and the virtual overlay during faster interactions is due to the technical design (see section 3) and therefore cannot be completely prevented. Since this delay was the same for all participants for all conditions, it can be assumed that it did not negatively affect the results.

Simulator sickness after the experiment was rated very low. This was to be expected, since simulator sickness occurs less frequently in AR than in VR. This result is in line with the result of the study on simulator sickness in AR of Vovk et al. [33].

6 Limitations

In this paper, we investigated the effect of size differences between the physical object and its virtual representation on usability, presence and performance. Since an abstract proxy object used for interaction differs not only in size but also in shape, texture, and material, the next step is to find out to what extent a deviation between virtual and physical object is possible with regard to these features.

The purpose of the study was to show that instead of using an exact replica, it is possible to use a physical prop that can differ in size to a certain degree from its virtual counterpart without too much negative impact on usability, presence and performance. However, no exact limits were determined as to how much one can increase or decrease the size. To reliably determine these limits, a larger sample size and appropriate methods, such as the up/down staircase procedure [10], will be needed in further studies.

The results of the study show that a limit exists to which an overlay can be smaller than the physical prop being interacted with. For virtual overlays larger than the physical prop, the limit in terms of AR/TAR presence is not foreseeable. However, it is expected that for virtual object sizes larger than 50% bigger than the physical object, significant worsening with respect to presence will also occur.

We did not correct for multiple comparisons, as this would have biased the results by increasing the ranges where no significant difference from the baseline condition M was detected due to the number of conditions to compare.

The study was performed under a fixed lighting condition chosen to make the overlays appear as opaque as possible. In reality, however, the lighting conditions are usually not as constant as in the laboratory. Since it can be assumed that the overlays are perceived differently under different lighting conditions, a next step would be to investigate to what extent the selected lighting conditions have an influence on the results.

7 Conclusion

In this paper we investigated if a physical proxy a user is interacting with can vary in size from its virtual representation in optical see-through AR without strong negative effects. In a study we examined the effect of size differences on the feeling of disturbance, the feeling of presence, size estimation and task completion time.

The results of the study show that it is possible to vary the size within certain ranges without too much worsening of disturbance and presence. It is therefore possible to use one single physical object as a tangible prop to interact with several virtual objects of different sizes. The size variation range is wider for virtual objects larger than the physical object than for smaller virtual overlays. If no prior knowledge about better fitting objects exists, slightly smaller and slightly larger (+/-10%) objects are even perceived as having the same size as the physical object. Furthermore, the examined size differences have no influence on the performance compared to the baseline condition M.

The results obtained are similar to the results of VR and video see-through AR studies, which can be explained by the fact that the overlays were so opaque that the physical objects were almost blocked, since the study was performed in a very dimly lit room.

In further studies it should to be investigated what effect different, more natural, lighting conditions have. Furthermore, other factors, such as differences in shape, need to be investigated.

References

[1] J.-g. Ahn, E. Ahn, S. Min, H. Choi, H. Kim, and G. J. Kim. Size perception of augmented objects by different AR displays. In *HCI International 2019 - Posters*, pp. 337–344. Springer, 2019.

[2] R. T. Azuma. A survey of augmented reality. *Presence: Virtual and Augmented Reality, Volume 6, Issue 4*, pp. 355–385, 1997.

[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, Volume 8, Issue 3 2015*, pp. 73–272, 2015.

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

[6] Y. S. Chang, B. Nuernberger, B. Luan, and T. Höllerer. Evaluating gesture-based augmented reality annotation. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 182–185. IEEE, 2017.

[7] P. Chen, X. Liu, W. Cheng, and R. Huang. A review of using augmented reality in education from 2011 to 2016. In *Innovations in Smart Learning*, pp. 13–18. Springer Singapore, 2017.

[8] Y. M. Choi. Applying tangible augmented reality for product usability assessment. *Journal of Usability Studies*, 14(4), 2019.

[9] 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.

[10] M. A. Garcıa-Pérez. Forced-choice staircases with fixed step sizes: Asymptotic and small-sample properties. *Vision Research*, 38(12):1861–1881, 1998.

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

[12] S. G. Hart and L. E. Staveland. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in Psychology*, vol. 52, pp. 139–183. Elsevier, 1988.

[13] A. Hettiarachchi and D. Wigdor. Annexing reality: Enabling opportunistic use of everyday objects as tangible proxies in augmented reality. In *CHI '16 Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 1957–1967. ACM, 2016.

[14] H. G. Hoffman. Physically touching virtual objects using tactile augmentation enhances the realism of virtual environments. In *Proceedings. IEEE 1998 Virtual Reality Annual International Symposium (Cat. No. 98CB36180)*, pp. 59–63. IEEE, 1998.

[15] Y. Huang. Evaluating mixed reality technology for architectural design and construction layout. *Journal of Civil Engineering and Construction Technology*, 11(1):1–12, 2020.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] Microsoft. Microsoft - HoloLens 2. Retrieved May 27, 2021 from https://www.microsoft.com/de-de/hololens/.

[21] B. Munsinger, G. White, and J. Quarles. The usability of the Microsoft HoloLens for an augmented reality game to teach elementary school children. In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*, pp. 1–4. IEEE, 2019.

[22] NaturalPoint Inc. DBA OptiTrack. OptiTrack - Motion Capture Systems. Retrieved May 27, 2021 from https://optitrack.com.

[23] PTC. Vuforia: Market-Leading Enterprise AR. Retrieved May 27, 2021 from https://www.ptc.com/en/products/vuforia.

[24] J. P. Rolland and H. Fuchs. Optical versus video see-through head-mounted displays in medical visualization. *Presence: Virtual and Augmented Reality, Volume 9, Issue 3*, pp. 287–309, 2000.

[25] 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.

[26] S. S. Sørensen. The development of augmented reality as a tool in architectural and urban design. *Nordic Journal of Architectural Research, Volume 19, No 4*, 19(4):25–32, 2006.

[27] M. Szemenyei and F. Vajda. Learning 3D object recognition using graphs based on primitive shapes. In *Workshop on the Advances of Information Technology, Budapest, Hungary*, pp. 187–195, 2015.

[28] M. Szemenyei and F. Vajda. 3D object detection and scene optimization for tangible augmented reality. *Periodica Polytechnica Electrical Engineering and Computer Science*, 62(2):25–37, 2018.

[29] X. d. Tinguy, C. Pacchierotti, M. Marchal, and A. Lécuyer. Toward universal tangible objects: Optimizing haptic pinching sensations in 3D interaction. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 321–330. IEEE, 2019.

[30] 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.

[31] 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.

[32] P. Vávra, J. Roman, P. Zonča, P. Ihnát, M. Němec, J. Kumar, N. Habib, and A. El-Gendi. Recent development of augmented reality in surgery: A review. *Journal of Healthcare Engineering*, 2017:1–9, 2017.

[33] A. Vovk, F. Wild, W. Guest, and T. Kuula. Simulator sickness in augmented reality training using the Microsoft HoloLens. In *CHI '18 Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–9. ACM, 2018.

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