Exploring Pseudo-Weight in Augmented Reality Extended Displays

Shohei Mori*
Graz University of Technology

Yuta Kataoka†
Ritsumeikan University

Satoshi Hashiguchi‡
Ryukoku University

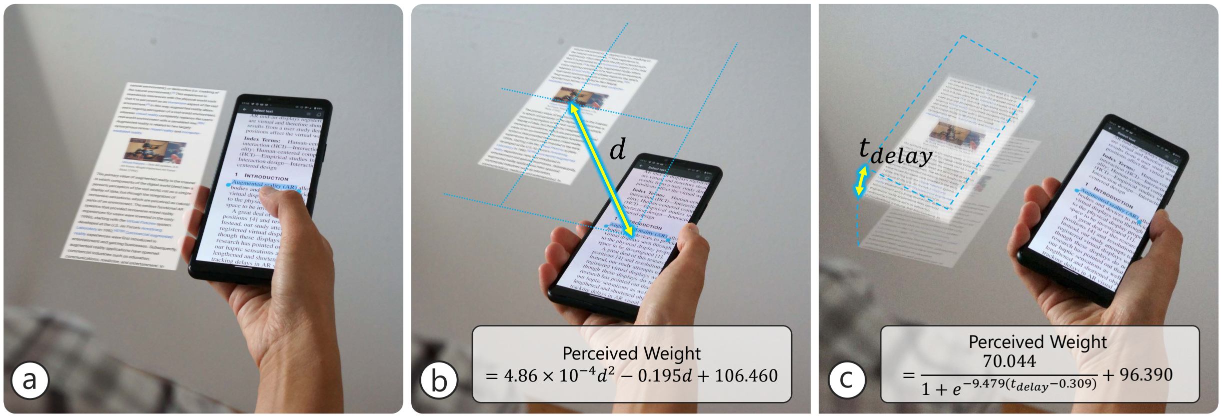


Figure 1: AR technology enables virtually adding free-form supplemental display areas in the AR space. a) These display areas, or AREDs, are private and more free from weight, size, and materials. Contrary to existing work that discusses the performance of AREDs, we explore how such "weightless" visualization affects our weight sensation. In this study, we found that b) the distance from the device of reference to AREDs and c) the time until it reacts to the physical motion induces a *lighter* and *heavier* impression, respectively. We provide quantifications of these two factors, as shown in each figure.

Abstract

Augmented reality (AR) allows us to wear virtual displays that are registered to our bodies and devices. Such virtually extendable displays, or AR extended displays (AREDs), provide personal display space and are free from physical restrictions. Existing work has explored the new design space to improve user experience and efficiency. Contrary to this direction, we focus on the weight that the user perceives from AREDs, even though they are virtual and have no physical weight. Our user study results show evidence that AREDs can be a source of pseudo-weight, in addition to that of a handheld physical display device. We also systematically evaluate the perceived weight changes depending on the layout and delay in the visualization system. These findings are similar to those in existing pseudo-haptics research. However, we found such behavior in pseudo-weight for a real device and virtual visual stimuli in the air, which differentiates our research from previous work.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Mixed / augmented reality; Computing methodologies—Computer graphics—Graphics systems and interfaces—Perception

1 Introduction

Augmented reality (AR) allows us to attach virtual displays to our bodies and devices to provide new display areas in the air, that is, AR extended displays (AREDs) (Figure 1). AREDs are beneficial because the extended virtual displays have no physical instance; that is they are mass free and material free and can only be seen through an head mounted display (HMD) of the user's own, hence improving the user's privacy. AREDs provide not only a new design space, but also interaction methods for better task performance in AR space [12, 23].

ARED research has validated factors such as display spatial arrangements [23], the type of registration references [9], and resolutions [12] in terms of interaction methods. However, our study attempts to investigate if such body- or device-registered virtual displays would have effects on the perceived weight, even though these displays do not physically exist. In fact, research has pointed out that AR visual stimuli have an impact on our haptic and our visual sensations. For example, virtually lengthened and shortened objects change the sensed weights [14], and tracking delays in AR visual stimuli induce a pseudo-haptics that gives counterforce against the moving direction [20].

Weight perception is essential haptic information to enhance the immersion of AR experiences. Several studies suggest that human perception is induced by an integration of multiple sensory modalities [10, 11]. Therefore, the impact of visual stimuli on weight perception has been investigated to deliver the intended perceptual experiences [14]. The most beneficial outcome of these studies is that one may achieve such weight manipulations without dedicated hardware in an AR system. Depending on if a manipulated object is real or virtual, the weight of the object will be differently perceived due to the differences in the superficial appearance of the object [6]. Further, dynamic changes in virtual objects' properties can also cause illusions in the perceived weights [20]. Prior studies have used spheres, rectangles, and similar simple virtual objects for manipulating the objects in virtual reality (VR) or AR space, and we follow this environmental setup. Nevertheless, we keep the ARED context in mind. In other words, we aim to explore how the perceived weight of AREDs could vary and what factors would impact the weight. We have a particular interest in weight manipulation since it may change AREDs' importance [1], valuableness [18], and favorability [25].

In light of previous pseudo-haptics studies and ARED implementations, we make three main hypotheses: 1) ARED holders would

*e-mail: shohei.mori@icg.tugraz.at
†e-mail: is0060hi@ed.ritsumei.ac.jp
‡e-mail: s_hashi@rins.ryukoku.ac.jp

feel the device with the additional visualization as either lighter or heavier, 2) the layout of AREDs would change the degree of perceived weight, and 3) delays from actual manipulation to displaying play a role to the overall weight. Clarifying the perceptual impact of AREDs on weight perception is vital both for designing displayed contents and exploring the interaction methods.

Our contributions can be summarized as follows.

- We demonstrate that AREDs consisting of a source display device and an additional virtual display can change the perceived weight of the original device while the source display device is being moved.
- We show that the weight illusion depends on the layout of the virtual display relative to the handheld source display device, and the greater distance between the handheld device and the virtual display induces a lighter feeling than that of the original device. The effects, however, are limited when extensively great distance is applied.
- We show that the weight illusion also depends on the delay of AREDs until the virtual part reacts to the motion of the source display device, and a longer delay induces a heavier feeling than that of the original device. The effects become weaker when an excessive delay is given.
- We provide quantification of these relationships as formulas.

2 Related Work

In this section, we first provide an overview of the ARED literature to point out that we are the first to investigate weight sensation under AREDs instead of designing and implementing AREDs. We then explore the literature on the impacts of static and dynamic visual stimuli on weight sensation. Here, we describe that our experimental setup is a special case of existing studies and, therefore, may provide a novel point of view in weight sensation, that is, pseudo-haptics.

2.1 AR Extended Displays (AREDs)

Much research has explored the interaction methods of virtual windows displayed in the air. Such virtual graphical interfaces are often seen through an HMD and commonly arranged next to the user's body or a device of reference. For example, Personal Cockpit presents virtual windows surrounding the user's body for mobile multitasking [9]. Open Palm Menu provides a virtual menu that shows a list of items around a non-dominant hand that can appear on-demand [3]. These displays enable menus that are accessible anytime because the displays are anchored to the user's body virtually. Depending on the parts that displays are registered to, the design and usability can vary.

Instead of presenting body-referenced windows, the below approaches extend a physical display device with virtual windows, that is, device-referenced windows. These approaches allow for utilization of high-resolution screen and touch controls on the device surface and use virtual sub-windows to extend the original display area. Tangible Spin Cube displays a 3D virtual ring menu around a trackable cube of reference, and this has been evaluated as intuitive and easy to use for novice users [21]. Extending a high-resolution narrow field of view (FOV) smartwatch with a lower resolution wide FOV virtual display is considered a reasonable choice as a hybrid display system that can improve the pointing task performance [12]. Besides the above solutions, world-fixed windows can be useful [27].

Although those seminal works validate the advantages of presenting additional virtual displays, the discussions are often limited to the new design space and task completion time. On the contrary, the aim of our study is to investigate somatosensation when such virtual displays are visually present. Physiological research suggests that the weight of objects influences their importance [1], valuableness [18], and favorability [25]. Thus, we specifically aim to validate the perceived weight of virtual display areas anchored to a holding device; however, these additional virtual displays have no mass.

2.2 Static Appearance & Its Impact on Weight Perception

Humans recognize the characteristics of an object through multiple sensory modalities, and every sensory modality interacts with others [26]. For example, visual information influences the perception of weight. Differences in the brightness [31], size [6], shape [8], and materials [33] of two objects with the same mass allow us to feel the objects' weight differently. Size-weight illusion (SWI) is caused by size differences showing that we perceive a larger object as lighter than a smaller object with the same weight and vice versa [6]. The literature reports that SWI occurs not only in real environments but also in VR [5] and AR [28]. That is, SWI occurs even when the sizes of the two compared objects are simply visually different.

In AREDs, a new virtual display area is superimposed next to a physical device held by the user, hereby using an AR technique. Because such virtual content is registered to the handheld device, the virtual content is indirectly manipulated using the device, which could be recognized either as a whole or two separate displays. In this study, we demonstrate that such configurations induce illusions in weight sensation and that the degree of perceived weight depends on the layout and delay of the extended display area. Further, we discuss how the literature supports these results.

2.3 Pseudo-haptics on Weight Perception

Apart from the constant differences in object's properties, more recent research suggests that dynamically changing manipulated object's properties also induces sensory illusion. Here, pseudo-haptics is a form of haptic illusion that presents pseudo-sensation when the manipulated object's properties are dynamically modified [20]. As a familiar example, we feel haptic resistance when a moving mouse cursor irregularly changes speed, even with no actual force feedback mechanism [32]. Pseudo-haptics also induces illusions in other sensory areas such as weight [24], shape [4], and hardness [2] when virtual objects dynamically change their visual characteristics.

Weight perception could change depending on the control–display ratio, that is, the ratio between actual motion and modified and displayed motion, on a 2D screen [7]. Here, -2% to 5% of motion modifications can change the perceived weight of a lifted object observed through an HMD [29]. Furthermore, there is a technique to induce weight illusion in an AR environment by deforming or shifting a virtual object in response to a collision with a real or virtual object [17, 19]. A motion-manipulated 2D virtual object can affect weight perception in a VR environment [15]. Further, a 200 ms delay while lifting a 3D virtual object makes the object 18% heavier in the user's perception [30]. Those studies indicate that the impact of the illusion in weight varies according to the degree of delays and that such effects vanish when there are excessive delays. This is because human perception is integrated with weighting that depends on the reliability of the information among the activated sensory modalities [11].

AREDs move in accordance with the tracked handheld device the user holds. Thus, poor AR system performance causes a delay from the interaction until the actual display timing, which would eventually change the user's perceived weight of the ARED. In the current paper, we demonstrate the effect and evaluate the impact of such delays on AREDs.

3 Overview

In this study, we explore the pseudo-weight perceived when an ARED that consists of a reference hand-held device and a uniformly colored AR mock-up (Figure 2(a) inset) is periodically waved at a pre-determined time mimicking the interactions of AREDs (Figure 2(b)). In this laboratory setup with uniform color backgrounds, the participants observe the AREDs through a video see-through

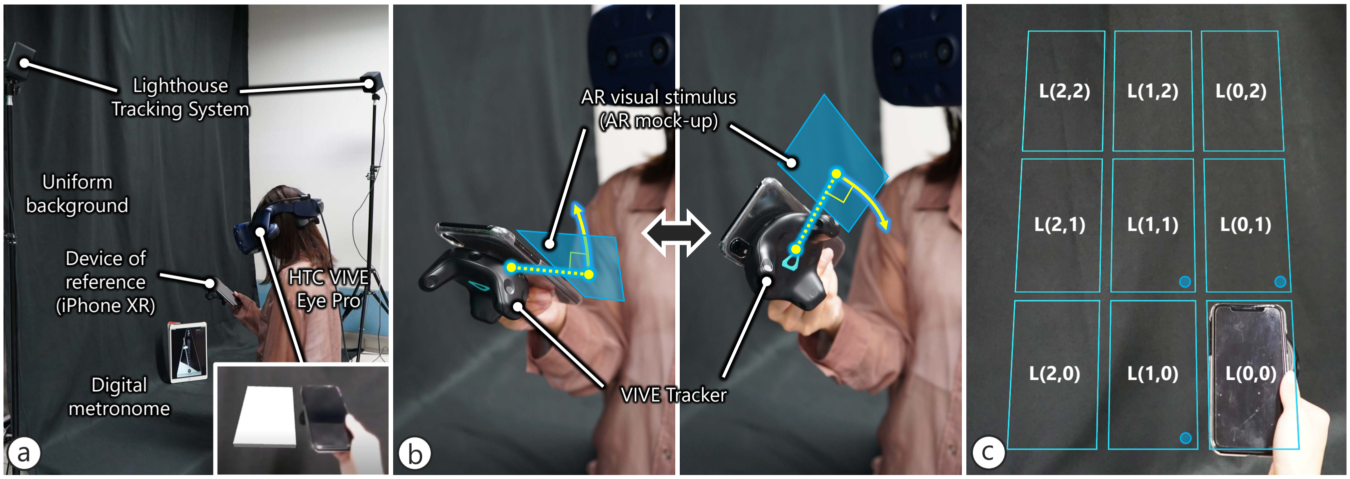


Figure 2: Setup of our experiments. a) The participants wore a video see-through HMD that allowed them to see real scene backgrounds with an AR mock-up overlaid against the background. This experimental ARED system tracked poses of the head and device of reference using an optical tracking system (HTC VIVE lighthouse) in six DOF. b) The participants were asked to actively move the device to feel the pseudo-weight. A VIVE tracker was attached behind the device to track it. The participants were trained to move the device as uniformly as possible. c) We randomly presented the AR visual stimulus at one of nine locations in Experiment 1 and three locations (indicated by blue dots) in Experiment 2, respectively.

HMD. We control the timing of the waving using a metronome. As discussed in the previous section, we investigate the impact on the two factors in AREDs, hence conducting two user studies:

Experiment 1: Layout We first demonstrate that an ARED presents a different weight from that of the sole reference device. We also investigate how dependent such pseudo-weights are on the mock-up locations. To this end, we register the mock-up next to the reference device and vary the mock-up location relative to the reference device in a grid (Figure 2(c)). We let the participants compare the weight of the mock-up for each location.

Experiment 2: Delay Delays caused by limited system throughput are potential sources for changing the perceived weight of AREDs. To investigate the possible delays systematically, we let the participants report how differently they feel the pseudo-weight on each designed delay. We register an AR mock-up as in the first experiment but add artificial delays to the displayed mock-up.

4 Experiment 1

Design We designed a repeated measures within-subjects study to compare the weight of the experimental ARED. We introduced the independent variable *location*, $L(x,y)$ in 3×3 grids (Figure 2(c)). As dependent variables, we collected *ratings for weights*, w_{rating}, of AREDs (the device of reference and AR mock-up) as the participants reported. We asked the participants to move the device held in their hand. The motivation behind this was to induce a confident feeling of weight [14] and to measure the perceived weight during their interactions with an ARED, e.g. menu toggling by gestures [3].

Task We performed a magnitude estimation study. We presented no AR visual stimulus followed by an AR mock-up at random locations per trial. Upon presenting each visual stimulus, the participants were asked to answer the weight in magnitude, while they were asked to set the first visual stimulus of a reference weight (a reference stimulus, no AR mock-up overlaid) to 100. Then, the participants answered the weights w_{rating} of an ARED in magnitude. This procedure was repeated until all visual stimuli had been evaluated.

Apparatus We built a video see-through AR system using an HTC VIVE Pro and VIVE SRWorks SDK[1]. As shown in Figure 2(b), we placed a VIVE tracker (2018) on the backside of a commodity smartphone to track six DOF poses of the device. We used an Apple iPhone XR (6.1 in, $w = 75.7$ mm and $h = 150.9$ mm, 314 g including a VIVE tracker) as the device of reference. The size of the AR mock-up was set to the size of the device of reference. The color of the AR visual stimulus was in a uniform gray to avoid any effect from colors. Considering practical ARED usages [23], we placed the mock-up within the participant's two arms and never placed in the area closer to the body to keep AREDs completely within the HMD's FOV. The device was held by the dominant hand. We covered the background with uniformly colored clothes to avoid disturbance from background clutter. We denote $L(x,y)$ as a relative location of an AR mock-up from a physical display device (Figure 2(c)). Specifically, $L(x,y)$ is (xw,yh) where w and h are the width and height of the physical display device, respectively.

As a practical note, even though we set the exact values in the system using the SDK, we observed a small margin between the AR visual stimulus and the physical device (Figure 2(a) inset). We believe this margin appeared due to registration misalignment; however, we left the margin as is to avoid overlaps between the virtual and real visual stimulus. For reference, we attempted to measure camera system latency, which was approximately 0.3 s. We note that none of the participants reported any physical disorder feeling regarding the system latency both in Experiments 1 and 2.

Participants In the experiment, 11 participants (2 female, $\bar{X} = 21.2$ (SD $= 0.8$) years old, right-handed) volunteered. All participants were university students majoring in computer science, and they have joined some AR-related user studies before.

Procedure After completing a consent form[2], each participant was introduced to a small training session where the participants were taught how to hold and move the device of reference appropriately as the trained experimenter demonstrated. Figure 2(b) illustrates the instructed motion. After wearing an HMD, the participants were asked to keep the device in hand around at the chest height, their elbow bent approximately 90 degrees, and their elbow contacting their body while standing. They were guided to rotate the device with their wrist and move it alternately towards the AR display mock-up's front and back surface normals (Figure 2(b)). The

[1] https://developer.vive.com/resources/vive-sense/sdk/vive-srworks-sdk/

[2] This experiment was conducted before the COVID-19 pandemic.

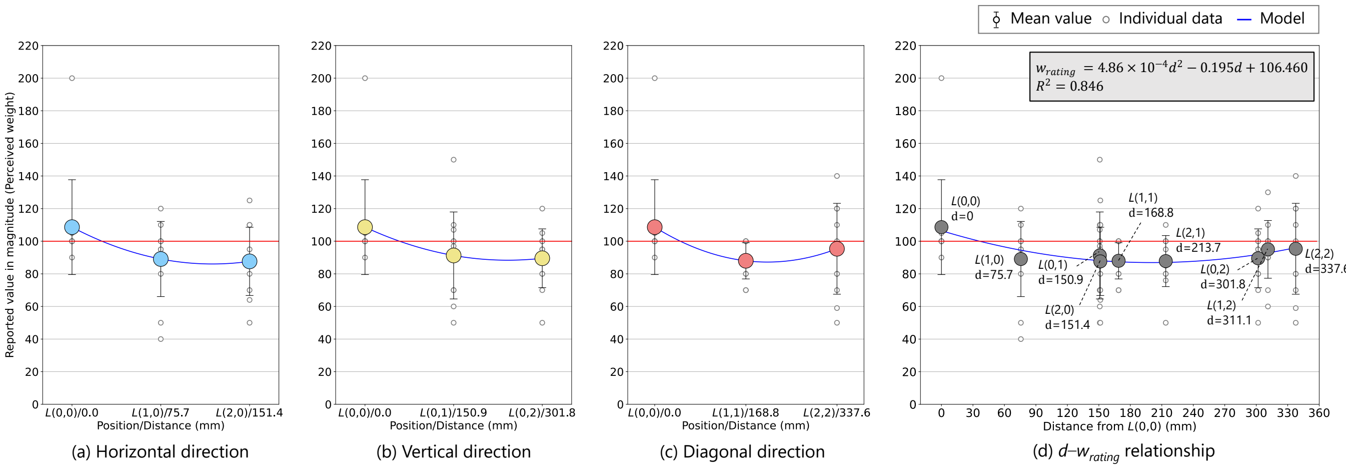


Figure 3: Results of Experiment 1. To evaluate dependencies in the major direction of the AREDs' locations and distances from the hand-held device at $L(0,0)$, we ran a two-way ANOVA on these two factors. We defined three directions, including a) horizontal (i.e., $L(x,0)$, where $(x \in \{0,1,2\})$), b) vertical directions (i.e., $L(0,y)$, where $(y \in \{0,1,2\})$), and c) diagonal (i.e., $L(x,y)$, where $(x \in \{0,1,2\} \cap x = y)$. d) Because the two-way ANOVA suggested the dependency only on distance, we further evaluated our results with respect to the distance of all samples, including the missing samples at $L(1,2)$ and $L(2,1)$. This result clearly shows that as the distance from $L(0,0)$ increases, the weight tends to be perceived as *"lighter"* and the effect seems to saturate at around 200 mm and diminish as the distance increases. Note that there are two overlapping samples around at 150 mm.

alternative rotation was kept at a pitch of 80 bpm to the metronome click sound. Also, the participants needed to restrict their wrist twists to keep the smartphone screen within the HMD's FOV. The participants continued the practice session as long as they wanted.

After finishing the practice session, the main session started. During the main session, the participants could rest whenever they wanted. At every trial, the reference stimulus was presented. The AR mock-up was then presented at one of nine locations $L(x,y)$ $(x,y \in \{0,1,2\})$ in a random order to diminish any order effect. The participants were asked to continue the motion taught in the practice session until they became confident about the presented ARED's weight. After moving the device, the participants rated the ARED's weight in magnitude of each configuration, and reported the score verbally to the experimenter. As explained in the *Task* section, a reference stimulus (i.e., without AR overlay) followed by a visual stimulus to be evaluated (i.e., AR display mock-up overlay) was presented in each trial. This procedure was repeated until all nine locations were rated individually. With 11 participants and nine repetitions, we collected a total of $11 \times 9 = 99$ ratings. At the end of the session, we collected comments from the participants. The entire procedure took approximately 20 minutes. We had sufficient intervals between trials to avoid sensory adaptation.

Hypotheses We expect that (**H1**) the reported weights of the AREDs will vary depending on the relative location with respect to the device of reference. We also assume that (**H2**) the AREDs with an AR mock-up located farther away from the physical device are felt as lighter than those located closer to the device. The literature [14] reports that a virtually extended stick was felt as lighter than a stick with its original size when it was waved. From this, we assume that a greater distance from a device to device-registered visual stimuli would make their weights lighter, even though the real and virtual objects are separately located in the AREDs. We suppose that (**H3**) the perceived weights of the AREDs depend on the direction toward where the mock-up is placed, for example, horizontally, diagonally, or vertically in the grid $L(x,y)$.

Results The score data were analyzed using a two-way ANOVA with *distance* and *direction* factors ($\alpha = 0.05$). The distance factor is in three groups, that is, the origin group: $L(0,0)$, near group: $L(0,1), L(1,0), L(1,1)$, and far group: $L(0,2), L(2,0), L(2,2)$. The

Table 1: Fitted curves in Experiment 1.

Direction	Curve (modeled w_{rating})	R^2
Horizontal	$1.58 \times 10^{-3}d^2 - 0.378d + 108.636$	1.000
Vertical	$3.43 \times 10^{-4}d^2 - 0.167d + 108.636$	1.000
Diagonal	$4.91 \times 10^{-4}d^2 - 0.205d + 108.636$	1.000
All	$4.86 \times 10^{-4}d^2 - 0.195d + 106.460$	0.846

main effect of the distance factor was significant ($F(2,99) = 6.332, p = 0.002$, $\eta_p^2 = 0.113$, $1-\beta = 0.916$); hence **H1** is supported. The direction factor is in three groups, that is, the horizontal group, $L(x,0)$, where $x \in \{0,1,2\}$, vertical group, $L(0,y)$, where $y \in \{0,1,2\}$, and diagonal group, $L(x,y)$, where $x \in \{0,1,2\} \cap x = y$. Contrarily, the statistical analysis revealed that the main effect of the direction factor was not significant ($F(2,99) = 0.049, p = 0.952$, $\eta_p^2 = 0.001$, $1-\beta = 0.057$); therefore, **H3** is not supported. Also, the interaction was not significant ($F(4,99) = 0.126$, $p = 0.972$, $\eta_p^2 = 0.005$, $1-\beta = 0.077$). Figure 3 summarizes (a-c) the scores in each direction group and (d) scores of all the data.

Because we found the significance in the distance factor, we further conducted a post-hoc test. Two-tailed multiple comparisons by t-test using the pooled standard deviation (SD) as a post-hoc test for the distance factor were performed. The p-value was adjusted by using Bonferroni's method. The statistical analysis found a significant difference ($t(105) = 3.244$, adjusted $p = 0.004$) between the origin ($\bar{X} = 108.75$) and near groups ($\bar{X} = 90.361$). Furthermore, a significant difference was found ($t(105) = 3.078$, adjusted $p = 0.008$) between the origin ($\bar{X} = 108.75$) and far group ($\bar{X} = 91.306$). Here, notice that the mean values of both near and far groups are lower than that of the origin group and reference of 100. However, the far group mean value is larger than that of the near group. These results suggest that illusions in weight sensation would degrade as the AR visual stimulus is placed greatly far from the device of reference. Overall, we conclude that **H2** is supported but partially.

For further detailed analyses, we fit a quadratic function to approximate the relationships between the distance $d = ||L(0,0) - L(x,y)||_2$ and perceived weight in each direction. Table 1 shows the fitted curves. As represented by the data at $L(2,2)$ in Figure 3(c), the data do not show a monotonic decrease.

Figure 3(d) shows the relationship between the distance d to the AR display mock-up and perceived weight w_{rating} of all the collected data. We found that a quadratic function best fits the data instead of other functions, e.g., a logarithmic function. From the fitted curve, the perceived weight monotonically decreases (AREDs are felt as lighter) until the distance from the device of reference grows up to around 200 mm, which is approximately at $L(2,1)$. The effect saturates towards the peak and then grows back to 100 in weight.

Discussion The two-way ANOVA revealed that the main effect of the direction factor was not significantly different. The main reason for this result might be the difficulties in rotating the device in a specific direction. Some participants commented as follows: "In the horizontal direction, the device is easy to rotate," and "I felt the motion of the display did not appear clearly especially when I rotate the AREDs vertically." However, from the results, we could not conclude yet that the different characteristics in the resulting perceived weight could be attributed to the AR mock-up's visibility.

In contrast, the distance factor showed a significant difference. In other words, the distance from the device of reference to the superimposed AR display mock-up influences the perceived weight. We formulate the distance–weight relationship as a quadratic function (Table 1 All). The literature [14] has demonstrated that a virtually extended rod was perceived as lighter than that with no visual edit. This result is consistent with ours if we interpret that an ARED with a greater distance d is a single extended object. Although in conditions $L(1,0)$ and $L(0,1)$ the AR mock-up was nearly attached next to the smartphone, in $L(0,2)$ and $L(2,2)$, the AR overlay was apart from the handheld device. The result showing pseudo-weight even with separate objects but virtually connected is considered new, which draws our interest the most in this experiment.

We observe huge variances in some conditions. Especially for $L(2,2)$, half of the participants reported that the AREDs were heavier. Those participants stated the following: "When the virtual object is located far, I felt that the object was not registered to my holding device," and "I felt such a virtual object does not belong to the device anymore." From these comments, the attribution of virtual objects to a handheld device could be a contributor to the variances of weight sensation. We also refer to Section 6.3 for deeper discussions.

5 Experiment 2

Design We carried out a repeated measures within-subjects study to analyze the impact of tracking delay on perceived weights in AREDs. To this end, we placed an AR mock-up at either $L(1,0)$, $L(1,1)$, or $L(0,1)$ (Figure 2(c)) and artificially set up a delay (See the *Designing Delay* section). We introduced an independent variable *delay*, t_{delay}, and a dependent variable *perceived weight*, w_{rating}. In a similar way as in Experiment 1, the perceived weight was evaluated as ratings in magnitudes of the relative weight of the presented AREDs against the reference visual stimulus, which was assumed to be 100. Here, the reference visual stimulus was an ARED overlaid at the same location with no delay so that the participants evaluated their perceived weights as affected only by the designed delay.

Task A magnitude estimation experiment was conducted. We presented a reference visual stimulus with no delay at a random location and then an ARED with a random delay t_{delay} at the same location, per trial. We asked the participants to assume that the weight of the reference visual stimulus was 100 at each randomly selected location. Then, the participants answered weight w_{rating} of the presented ARED with a given delay t_{delay}. This procedure was repeated until all visual stimuli had been evaluated.

Apparatus We used the same AR system and AREDs, as in the previous experiment. The only difference between the experimental setups is the delay artificially given to the AR mock-up in AREDs.

Designing Delay A preliminary experiment was conducted to determine the range of the delay. We used the same experimental setups as those in Experiment 1 (i.e., the same hardware and experimental controls). Five participants (1 female, $\bar{X} = 21.4$ (SD $= 0.5$) years old, right-handed) volunteered to join the preliminary experiment. Among these participants, two participants had taken part in Experiment 1.

First, the participants wore the HMD and grasped the smartphone at chest height. They were asked to remember the reference weight of the AREDs with an AR mock-up superimposed either at $L(1,0)$, $L(0,1)$, or $L(1,1)$ without a delay ($t_{delay} = 0$), while twisting the device at 80 bpm as they were standing. The participants were then asked to tell the experimenter the timing when they agreed with the following conditions the most: 1-1) "I have started to feel the delay of the ARED," and 1-2) "I have noticed that the ARED did not seem to be attached to the handheld device anymore." Referring to the literature [15], the participants examined delays ranging from 0.0 s to 0.8 s at every 0.05 s in ascending order. In addition, the participants were asked to tell the experimenter on which timing they most agreed with the following conditions: 2-1) "The ARED started to follow the hand-held device, and 2-2) "The delay completely diminished." The participants examined delays ranging from 0.0 s to 0.8 s at every 0.05 s in descending order.

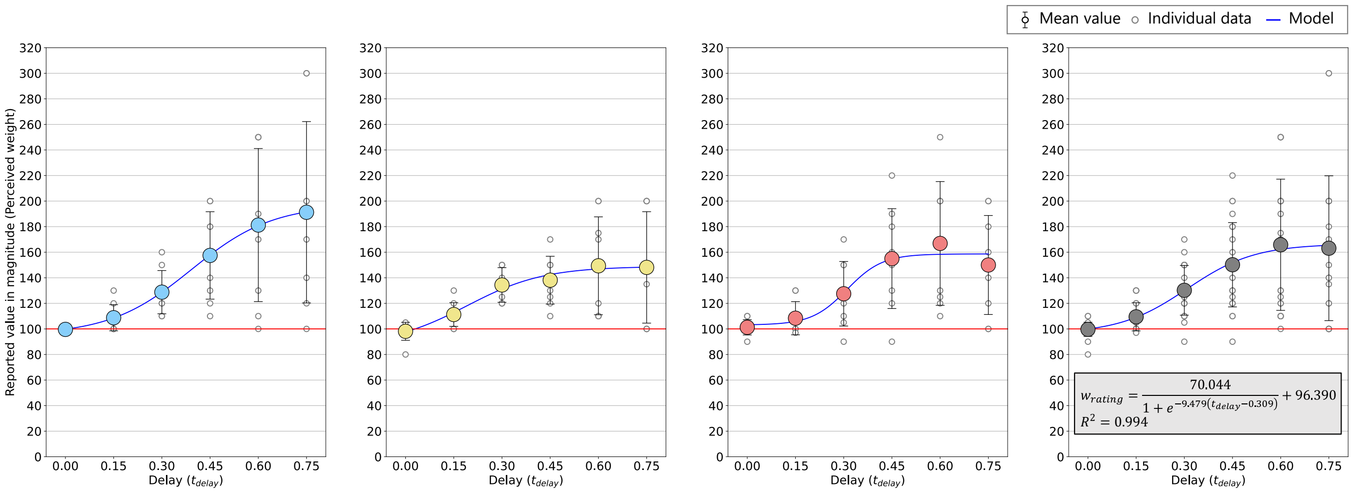
The results of the preliminary experiment showed that the median values for conditions 1-1 and 1-2 were 0.10 s (min. 0.08 s, max. 0.18 s) and 0.44 s (min. 0.24 s, max. 0.60 s), respectively. Also, the median values for items 2-1 and 2-2 were 0.38 s (min. 0.30 s, max. 0.50 s) and 0.14 s (min. 0.08 s, max. 0.18 s), respectively. Based on the results, we decided to use the following six delays: $t_{delay} \in \{0,\ 0.15,\ 0.30,\ 0.45,\ 0.60,\ 0.75\}$ s. Namely, we set the shortest delay to 0.15 (> 0.14) s where the participants started to feel the delay and sampled delays by every 0.15 s until the upper-bound of 0.8 s [15].

Participants We collected data from eight participants (4 female, $\bar{X} = 22.1$ (SD $= 0.8$) years old, right-handed). Much like in Experiment 1, all participants were university students majoring in computer science, and they have joined some AR-related user studies before. Among all of the participants, three, two, and one participant had joined the preliminary experiment, Experiment 1, and both, respectively.

Procedure We paid the highest attention to hygiene since this experiment was conducted during the COVID-19 pandemic.

After filling out a consensus form, a small training session was conducted for each participant much like in Experiment 1. After this session, each participant evaluated the weight of a reference visual stimulus at a randomly selected location, $L(x,y)$, compared with an ARED visual stimulus at the same location with one of the six different delays. The participants moved the ARED until they were confident about the felt weight of each visual stimulus. Then, the perceived weight in magnitude was reported to the experimenter verbally. The procedure was repeated until all the configurations has been evaluated. The participants could take a break if they wanted to. The participants provided comments at the end of the session. The session took approximately 20 min. With eight participants and six different delays for three directions, we collected a total of $8 \times 6 \times 3 = 144$ ratings. We had sufficient intervals between trials to avoid sensory adaptation.

Hypothesis It is a well-known effect that delays in the movement of the controlled object on a 2D screen induce the feeling of weight [15]; Delays in the motion of a controlling device make the controlling person feel the device is heavier. We, therefore, assume that (**H4**) the delay affects weight perception in AREDs, i.e., such pseudo-weight would change depending on the given time length of the delay. Finally, despite the results of Experiment 1, we expect that (**H5**) weight perception depends on the direction factor.



(a) Horizontal direction (b) Vertical direction (c) Diagonal direction (d) t_{delay}–w_{rating} relationship

Figure 4: Results of Experiment 2. We evaluated delays that could be caused by the AR system throughput. Much like in Experiment 1, we placed the mock-up AR overlay a) horizontally at $L(1,0)$, b) vertically at $L(0,1)$, and c) diagonally at $L(1,1)$ and added systematic delays artificially. Although the delays in the horizontal setup show clearer effects on the weight than the other configurations, similar to Experiment 1, statistical analysis revealed the dependency only on delays. d) The figure summarizes all the results, regardless of the mock-up locations. The results show that as the delay increases, the reported weight monotonically increases. However, the effect saturates at around 0.60 s. Recall that the reference visual stimulus in Experiment 2 and those at the same locations (i.e., $L(1,0)$, $L(0,1)$, and $L(1,1)$) in Experiment 1 are different as the results appear differently in the plots (the latter shows higher values above 100). Concretely, in Experiment 1, the participants compared weights of AREDs without and with AR mock-up overlay, while in Experiment 2, participants compared the weights of the AREDs without and with delays.

Table 2: Significance of the delay combinations. The highlighted values show significant differences ($p < 0.05$).

	0.00	0.15	0.30	0.45	0.60
0.15	1.000	N/A	N/A	N/A	N/A
0.30	0.058	0.719	N/A	N/A	N/A
0.45	< 0.001	0.002	0.847	N/A	N/A
0.60	< 0.001	< 0.001	0.012	1.000	N/A
0.75	< 0.001	< 0.001	0.028	1.000	1.000

Table 3: Fitted curves in Experiment 2.

Direction	Curve (modeled w_{rating})	R^2
Horizontal	$\frac{101.821}{1+e^{-7.895(t_{delay}-0.391)}}+95.345$	0.999
Vertical	$\frac{61.724}{1+e^{-8.323(t_{delay}-0.191)}}+87.154$	0.982
Diagonal	$\frac{55.581}{1+e^{-18.206(t_{delay}-0.310)}}+103.071$	0.955
All	$\frac{70.044}{1+e^{-9.479(t_{delay}-0.309)}}+96.390$	0.994

Results The reported weight ratings were analyzed using a two-way ANOVA with *delay* and *direction* factors ($\alpha = 0.05$). The delay factor is in six groups, that is the six different delays determined in the preliminary study (please refer to the *Designing Delay* section). The direction factor is in three groups, that is horizontal group: ratings at $L(1,0)$, vertical group: ratings at $L(0,1)$, and diagonal group: ratings at $L(1,1)$. The statistical analysis found that the main effect of the delay factor was significantly different ($F(5,126) = 14.378$, $p < 0.001$, $\eta_p^2 = 0.363$, $1-\beta = 0.999$); therefore, **H4** is supported but partly due to the saturation with extensive delay. However, no significant difference was found in the main effect of the direction factor ($F(2,126) = 2.049$, $p = 0.133$, $\eta_p^2 = 0.032$, $1-\beta = 0.466$), therefore **H5** is not supported. The interaction effect was not significant ($F(10,126) = 0.797$, $p = 0.631$, $\eta_p^2 = 0.006$, $1-\beta = 0.461$).

For the delay factor where we found significant differences, we conducted two-tailed multiple comparisons with a t-test using pooled SD as a post-hoc test. Table 2 summarizes the results. These results show that the AREDs were perceived as heavier when the delays were given to the AR display mock-up. To formulate the relationships between t_{delay} and w_{rating}, we fit sigmoid functions to the data of each direction group and to all the data (Table 3). We found that all three direction groups and a set of all data well fit the sigmoid functions ($R^2 > 0.95$). The fitted curves in Figure 4(a-d) illustrate that the perceived weight monotonically increases (i.e., being felt heavier) with respect to the delay, regardless of the directions.

Discussion The statistical analysis neither showed significant differences for the direction factor nor interaction with the delay and direction factors. Related to these results, the participants stated the following: "It is easier to swing the device in the horizontal direction," and "it is harder to see the AR display mock-up at the vertical condition moving toward me than that at the horizontal or diagonal conditions." These comments were the same as in Experiment 1. The average scores for the 0.60 s delay condition in each direction are 181.25 (horizontal), 166.875 (diagonal), and 149.375 (vertical) in descending order. Due to the nature of a test of delays, the participants had to observe the AR display mock-up carefully to guarantee that they were observing the delayed motion. Therefore, difficulties in controlled swing motions and the visibility of the AR visual stimulus would have had stronger effects on the end perceived weight.

Figure 4(a-d) show that the weight ratings increase monotonically up to around a delay of 0.45 s or 0.6 s. However, beyond the delay, the growth in weight starts diminishing. $t_{delay} = 0.44$ is the median value where the participants in the preliminary study started to report that the AR visual stimulus did not follow the motion of the device. As such, many participants in the 0.60 s and 0.75 s conditions were not able to recognize the AR visual stimulus as tracked or registered to the holding device. In particular, the variances in both of the conditions are rather higher than those of the other conditions (0.60 s: SD 51.33 and 0.75 s: SD 56.64 while 0.45 s: SD 32.99, see also Figure 4(d)). With the two specific conditions, we observed two categories of participants whose scores either increased or decreased. Here, there were two participants who joined both the preliminary experiment and Experiment 2. One of them commented that in the 0.3 s condition, the AR display mock-up was not following the motion of their moving device, and the scores of this participant decreased in the 0.6 s and 0.75 s conditions. Consequently, we would conclude that if the participants recognized that the AR display mock-up tracked their motion, then those participants reported increased ratings for longer delays, while if they felt the tracking was lost, then they reported decreased ratings for the same longer delays. That is, one of the key factors to the pseudo-weight in AREDs would be unity as an ARED under the holder's subconsciousness.

6 Interpretations and Future Directions

In this section, we give an overview of our findings. Finally, we provide interpretations of our study, limitations, and future directions in weight perception of AREDs.

6.1 Layout and Delay Matter to ARED's Weight

Under our experimental conditions, we have confirmed two factors that are relevant to AREDs' weight perception: the *distance* between the device of reference the user holds and AR display area registered to the device and the *delay* of the AR display area when the held device is moved. In both Experiments 1 and 2, we observed a significant difference between the main effects to the weight sensation neither in *direction*s nor their interactions between *direction*s and the abovementioned factors. The collected data enabled us to approximately quantify the relationships between the expected perceived weight and each of the two factors, regardless of to which *direction* an AR display area is placed. Contrary to this, interconnections between the two factors remain a topic for future work.

While our findings are statistically supported, we recognize that our research is user study-driven, and therefore, we should not forget the limited variety in gender and age. We are eager to collect more participants to improve the solidity of our findings after the COVID-19 pandemic ends. We should also point out our symbolized setups, that is, the uniformly colored AR mock-up and background, and regularized movement as interactions. While such a setup is in line with previous studies, investigating the weight sensation of AREDs with more practical applications remains our future work.

6.2 How to Design Heavier and Lighter AREDs

A participant who joined both experiments mentioned that the heaviness from the delay was more pronounced than the lightness from the virtual display layout. Because the reference visual stimuli were different in Experiments 1 and 2, we understand that we cannot compare these two results directly. However, it seems to stand out that the ratings in Experiment 2 grew to doubled the reference stimulus. As we discussed in Section 2.1, such controllable heaviness of AREDs will bring benefits to its interface design.

We successfully formulated weight-distance and weight-delay relationships under the experimental setup (e.g., the formulas in the teaser figure). Although we used an AR display mock-up with the same size as the device of reference, some ARED applications utilize larger virtual screens [12]. According to SWI research, a larger virtual display area would make the user feel as though these were lighter. Therefore, to design a lighter ARED, additional parameters including the virtual display layout and size would have to be taken into account.

6.3 Handheld Device and AR Screen as a Whole

We observed the reduced pseudo-weight effects for the excessive greater *distance* and longer *delay* become closer to the reference weight. As such, each such effect has a certain effective range. Humans can extend their body image when using a tool and use the tool as if it was a part of their body [16]. Our results suggest that how to feel an ARED's weight seems to depend on how sequaciously the AR display area follows the user's motion (refer to *Discussion* in Sections 4 and 5). Therefore, pseudo-weight in AREDs would depend on how well the user extends their body image when holding the ARED.

As future work, we are interested in having a user study that takes two factors (*distance* and *delay*) at once to see how the effective range would behave. Especially, we expect that the saturation of the effects would occur sooner as observed in $t_{delay} = 0.45$ s of Figure 4(c). Contradictory to the usual method of user interface design, in AREDs, there would be a chance to have a richer design space regarding the pseudo-weight by laying out the virtual content far from the place where the usability is the most enhanced, and doing so on purpose.

7 Conclusion

In this study, we have explored the pseudo-weight perceived in AREDs consisting of a handheld device of reference and an AR display mock-up with no weight. Through two user studies, we first demonstrated that (1) pseudo-weight can present different weights than that of the original device in AREDs. We also found that such pseudo-weight is dependent on (2) the distance from the physical device to the AR mock-up and (3) the delay, with which the AR mock-up follows the motion of the physical device later. We further clarified that (4) the greater distance induces the lighter pseudo-weight and (5) the longer delay induces the heavier pseudo-weight in AREDs although the effects diminish gradually as distances and delays grow extensively. Finally, (6) we formulate both distance–weight and delay–weight relationships with data-fitted curves (see the formulas in the teaser figure). We would enhance that our findings are similar to but different from results in previous pseudo-haptics research to the point that we found such behavior in pseudo-weight for a real device and virtual visual stimuli in the air.

Although we used a uniformly colored virtual object, using different colors and content for the virtual visual stimulus would be an interesting extension. Also, having variations in the source display form factor could result in different results (e.g., tablets, smartwatches, etc.). How much the user is familiar with AR experience would affect the weight illusion, but this remains a topic for future work. In addition to the future work discussed in Section 6, we are interested in combining our findings in a more direct way, that is, presenting physical stimulus to the user's skin [13, 22].

Acknowledgments

This work was partly supported by JSPS KAKENHI Grant Numbers JP 21K11947.

References

[1] J. M. Ackerman, C. C. Nocera, and J. A. Bargh. Incidental haptic sensations influence social judgments and decisions. *Science*, 328(5986):1712–1715, 2010.

[2] F. Argelaguet, D. A. G. Jáuregui, M. Marchal, and A. Lécuyer. Elastic Images: Perceiving local elasticity of images through a novel pseudo-haptic deformation effect. *ACM Trans. on Applied Perception*, 10(3):1–14, 2013.

[3] T. Azai, M. Otsuki, F. Shibata, and A. Kimura. Open Palm Menu: A virtual menu placed in front of the palm. In *Proc. of the 9th Augmented Human International Conference*, pp. 1–5, 2018.

[4] Y. Ban, T. Kajinami, T. Narumi, T. Tanikawa, and M. Hirose. Modifying an identified curved surface shape using pseudo-haptic effect. In *Proc. IEEE Haptics Symposium*, pp. 211–216, 2012.

[5] G. Buckingham. Examining the size–weight illusion with visuo-haptic conflict in immersive virtual reality. *Quarterly Journal of Experimental Psychology*, 72(9):2168–2175, 2019.

[6] A. Charpentier. Experimental study of some aspects of weight perception. *Arch. Physiol. Normales Pathologiques*, 3:122–135, 1981.

[7] L. Dominjon, A. Lécuyer, J.-M. Burkhardt, P. Richard, and S. Richir. Influence of control/display ratio on the perception of mass of manipulated objects in virtual environments. In *Proc. IEEE Virtual Reality*, pp. 19–25, 2005.

[8] F. B. Dresslar. Studies in the psychology of touch. *The American Journal of Psychology*, 6(3):313–368, 1894.

[9] B. M. Ens, R. Finnegan, and P. P. Irani. The Personal Cockpit: A spatial interface for effective task switching on head-worn displays. In *Proc. ACM Conf. on Human Factors in Computing Systems*, pp. 3171–3180, 2014.

[10] M. O. Ernst. A bayesian view on multimodal cue integration. *Human Body Perception from the Inside Out*, 131:105–131, 2006.

[11] M. O. Ernst and M. S. Banks. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870):429–433, 2002.

[12] J. Grubert, M. Heinisch, A. Quigley, and D. Schmalstieg. Multifi: Multi fidelity interaction with displays on and around the body. In *Proc. ACM Conf. on Human Factors in Computing Systems*, pp. 3933–3942, 2015.

[13] T. Han, F. Anderson, P. Irani, and T. Grossman. Hydroring: Supporting mixed reality haptics using liquid flow. In *Proc. ACM Symp. on User Interface Software and Technology*, pp. 913–925, 2018.

[14] S. Hashiguchi, S. Mori, M. Tanaka, F. Shibata, and A. Kimura. Perceived weight of a rod under augmented and diminished reality visual effects. In *Proc. ACM Symp. on Virtual Reality Software and Technology*, pp. 1–6, 2018.

[15] T. Honda, N. Hagura, T. Yoshioka, and H. Imamizu. Imposed visual feedback delay of an action changes mass perception based on the sensory prediction error. *Frontiers in Psychology*, 4:760, 2013.

[16] A. Iriki, M. Tanaka, S. Obayashi, and Y. Iwamura. Self-images in the video monitor coded by monkey intraparietal neurons. *Neuroscience Research*, 40(2):163–173, 2001.

[17] P. Issartel, F. Guéniat, S. Coquillart, and M. Ammi. Perceiving mass in mixed reality through pseudo-haptic rendering of newton's third law. In *Proc. IEEE Virtual Reality*, pp. 41–46, 2015.

[18] N. B. Jostmann, D. Lakens, and T. W. Schubert. Weight as an embodiment of importance. *Psychological Science*, 20(9):1169–1174, 2009.

[19] T. Kawai, F. Ohta, S. Kim, and H. Morikawa. Pseudo-haptic by stereoscopic images and effects on muscular activity. *Electronic Imaging*, 2017(5):80–84, 2017.

[20] A. Lécuyer, S. Coquillart, A. Kheddar, P. Richard, and P. Coiffet. Pseudo-haptic feedback: can isometric input devices simulate force feedback? In *Proc. IEEE Virtual Reality*, pp. 83–90, 2000.

[21] H. Lee and W. Woo. Tangible spin cube for 3D ring menu in real space. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 4147–4152. 2010.

[22] T. Murakami, T. Person, C. L. Fernando, and K. Minamizawa. Altered touch: Miniature haptic display with force, thermal and tactile feedback for augmented haptics. In *Proc. ACM SIGGRAPH 2017 Emerging Technologies*, pp. 1–2. 2017.

[23] E. Normand and M. McGuffin. Enlarging a smartphone with AR to create a handheld VESAD (virtually extended screen-aligned display). In *Proc. Symp. on Mixed and Augmented Reality*, pp. 123–133, 2018.

[24] K. L. Palmerius, D. Johansson, G. Höst, and K. Schönborn. An analysis of the influence of a pseudo-haptic cue on the haptic perception of weight. In *Proc. Conf. on Human Haptic Sensing and Touch Enabled Computer Applications*, pp. 117–125, 2014.

[25] B. Piqueras-Fiszman, V. Harrar, J. Alcaide, and C. Spence. Does the weight of the dish influence our perception of food? *Food Quality and Preference*, 22(8):753–756, 2011.

[26] A. Pusch and A. Lécuyer. Pseudo-haptics: From the theoretical foundations to practical system design guidelines. In *Conf. on Multimodal Interfaces*, pp. 57–64, 2011.

[27] C. Reichherzer, J. Fraser, D. C. Rompapas, and M. Billinghurst. Secondsight: A framework for cross-device augmented reality interfaces. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–6, 2021.

[28] N. Rohrbach, J. Hermsdörfer, L.-M. Huber, A. Thierfelder, and G. Buckingham. Fooling the size–weight illusion—using augmented reality to eliminate the effect of size on perceptions of heaviness and sensorimotor prediction. *Virtual Reality*, 25:1–10, 2021.

[29] Y. Taima, Y. Ban, T. Narumi, T. Tanikawa, and M. Hirose. Controlling fatigue while lifting objects using pseudo-haptics in a mixed reality space. In *Proc. IEEE Haptics Symposium*, pp. 175–180, 2014.

[30] V. van Polanen, R. Tibold, A. Nuruki, and M. Davare. Visual delay affects force scaling and weight perception during object lifting in virtual reality. *Jour. of Neurophysiology*, 121(4):1398–1409, 2019.

[31] P. Walker, B. J. Francis, and L. Walker. The brightness-weight illusion. *Experimental Psychology*, 2010.

[32] K. Watanabe and M. Yasumura. Visualhaptics: Generating haptic sensation using only visual cues. In *Proc. Conf. on Advances in Computer Entertainment Technology*, pp. 405–405, 2008.

[33] H. K. Wolfe. Some effects of size on judgments of weight. *Psychological Review*, 5(1):25, 1898.