ORIGINAL ARTICLE

Fooling the size–weight illusion—Using augmented reality to eliminate the effect of size on perceptions of heaviness and sensorimotor prediction

Nina Rohrbach[1] · Joachim Hermsdörfer[1] · Lisa-Marie Huber[1] · Annika Thierfelder[1] · Gavin Buckingham[2]

Abstract

Augmented reality, whereby computer-generated images are overlaid onto the physical environment, is becoming significant part of the world of education and training. Little is known, however, about how these external images are treated by the sensorimotor system of the user – are they fully integrated into the external environmental cues, or largely ignored by low-level perceptual and motor processes? Here, we examined this question in the context of the size–weight illusion (SWI). Thirty-two participants repeatedly lifted and reported the heaviness of two cubes of unequal volume but equal mass in alternation. Half of the participants saw semi-transparent equally sized holographic cubes superimposed onto the physical cubes through a head-mounted display. Fingertip force rates were measured prior to lift-off to determine how the holograms influenced sensorimotor prediction, while verbal reports of heaviness after each lift indicated how the holographic size cues influenced the SWI. As expected, participants who lifted without augmented visual cues lifted the large object at a higher rate of force than the small object on early lifts and experienced a robust SWI across all trials. In contrast, participants who lifted the (apparently equal-sized) augmented cubes used similar force rates for each object. Furthermore, they experienced no SWI during the first lifts of the objects, with a SWI developing over repeated trials. These results indicate that holographic cues initially dominate physical cues and cognitive knowledge, but are dismissed when conflicting with cues from other senses.

Keywords Weight illusions · Perception · Sensorimotor control · Virtual reality · Holograms

1 Introduction

Immersive virtual reality (iVR) typically involves allowing an individual to experience, and interact with, a computer-generated environment as if it were the physical environment. Recent technological advances and cost reductions have significantly widened access to this technology, which has become a significant part of the consumer entertainment industry and its potential is rapidly being seen as a training aid for medical education and dangerous industries (Allcoat and Mühlenen 2018; Bideau et al. 2010; Rohrbach

et al. 2019a). A less widespread technology, which has arguably more potential for broad integration into society, is augmented reality (AR). AR systems use either a camera or a transparent screen to provide a live view of the physical environment overlaid with computer-generated images to augment the viewer's perspective of what they are viewing. In its simplest form, a smartphone with a camera is capable of delivering a reasonably compelling AR experience. More sophisticated devices, such as the Microsoft HoloLens, use translucent lenses, external sensors, and holographic projection to overlay individual graphical elements to discrete elements of the physical environment. This technology, while far from widespread, has significant potential to fundamentally alter real-time access to information in the classroom and the workplace (Dey et al. 2018) and has recently been trialled as a way to support clinical populations (Rohrbach et al. 2019b). Further, AR provides a better sense of presence than VR because the user can see their own body parts interacting with the real environment (Al-Issa et al. 2012). Little is known, however, about how the perceptual system

✉ Nina Rohrbach
  nina.rohrbach@tum.de

[1] Department of Sport and Health Sciences, Chair of Human Movement Science, Technical University of Munich, Munich, Germany

[2] Department of Sport and Health Sciences, University of Exeter, Exeter, UK

of the user treats computer-generated virtual cues, nor about how this virtual information interacts with the sensory input from objects in the physical environment.

The objective of this experiment was to determine whether the size–weight illusion (SWI) can be manipulated by augmenting the apparent size of the stimuli with holographic size stimuli. The SWI describes the experience that small objects feel heavier than equally weighted larger objects (Buckingham 2014; Charpentier 1891). Although the precise mechanisms underpinning this effect are still robustly debated (Dijker 2014; Freeman et al. 2019; Peters et al. 2016; Plaisier et al. 2019), the magnitude of the illusion (i.e. how much heavier the small object feels relative to the large object) appears to relate to expectations of heaviness elicited by the apparent volume of the stimuli (Buckingham and Goodale 2013; Buckingham and MacDonald 2016). Indeed, the magnitude of the illusion appears to be related to the reliability of the cue though which object volume is experienced. For example, a recent study has shown that impairing vision with specially designed goggles reduced the degree to which a small set of objects felt heavier than a larger set (Wolf et al. 2018). If one's 'belief' in the size information they are experiencing is a mitigating factor in the SWI, it seems plausible that this effect can be used to provide an insight into how the perceptual system treats novel cues. This has been shown, for example, in the case of human echolocation, whereby unsighted individuals can gain knowledge of their surroundings by interpreting the echoes returned from self-generated vocalizations or clicks. Buckingham et al. (2015) showed that blind echolocators had their experience of object weight affected by the size cues induced by these echoes, which provided new insights into the degree this substituted sense was a genuine replacement for vision.

In the context of augmented reality, the SWI could thus become a proxy to determine how our perceptual systems value AR cues in relation to those generated by the physical environment. It is increasingly well-established that altering visual gain (i.e. increasing of decreasing the speed of the computer-generated object compared to its physical counterpart) can alter the experience of object weight (Von Polanen et al. 2019; Weser and Proffitt 2019). Several studies have already used the SWI in the context of immersive virtual reality through a stereoscopic head-mounted display to show that virtual size cues are capable of affecting how heavy an object feels when lifted (Buckingham 2019; Heineken and Schulte 2007; Kawai et al. 2007). Furthermore, a recent compelling study from van Polanen and Davare (2019) showed that altering the sizes of objects while they are being lifted in a VR environment affects their weight. Interestingly, all of these studies note that visual cues to size delivered through computer-generated images in VR yield a smaller illusion than is experienced when the size cues are experienced through cues from the physical environment. While this might suggest that we prioritize cues derived from the physical environment over computer-generated equivalents, this supposition has never been directly tested. The semi-translucent computer-generated images which are overlaid on the cues from the physical environment in AR provides a unique opportunity to directly examine how these distinct sources of information interact within a single modality.

In addition to examining how augmented reality might affect the experience of object weight, this paradigm also allows us to examine how these computer-generated cues might affect the fingertip forces used to grip and lift objects. Skilful object manipulation requires the application of appropriate forces. A lifter's expectation of object heaviness influences the nature of this interaction, such that the peak values of grip and load force rates which occur prior or at the time of lift-off serve as a measure for sensorimotor prediction (Li et al. 2011). In the physical environment, a plethora of studies have shown that fingertip force rates (i.e. the maximum of first derivative of a grip force signal or the weight change, respectively) reflect the apparent weight (as signalled by the size) of objects during the initial lifts of objects (Gordon et al. 1991; Nowak and Hermsdoerfer 2009). This reliance of visual expectations from prior experience means that, in a typical SWI paradigm, the large object is gripped and lifted at a higher rate of force than the small object on the first pair of lifts (Buckingham et al. 2011; Davis and Roberts 1976). These long-term expectations are quickly discarded in favour of more short-term evidence based on tactile and visual feedback from the lifts themselves, meaning participants adapt their fingertip force rates to the actual (and thus identical) masses of each object (Flanagan and Beltzner 2000; Grandy and Westwood 2006). In an unpublished Masters thesis, Metcalfe (2007) compared a traditional SWI to a visuo-haptic SWI created in an augmented environment. Despite a lack of vision of the grasping hand and the physical cubes in the augmented environment, participants judged the cubes to feel similarly heavy. The study reaffirmed the robustness of the SWI in both environments. However, despite the persisting perceptual illusion, a steady difference in kinematics between cubes that changed in the same way for all cube sizes was reported. It was suggested that visual and haptic size cues can override sensorimotor memory and that AR is suitable in simulating the natural environment. To our knowledge, neither the initial parameterization of fingertip force rates, nor the subsequent adaptation processes, have been examined in the context of virtual or augmented reality–it is unknown whether the sensorimotor system will respond to computer-generated and augmented environments in a way which reflects performance in the physical environment, which has obvious consequences for the wider uptake of AR in society.

Here, we tested whether computer-generated size cues delivered with an AR system could override the properties of real-world objects. Participants in one group lifted SWI-inducing objects, reporting how heavy they felt after each lift and having their fingertip forces measured during each lift. Participants in another group lifted the same objects which had semi-transparent holograms of identically sized cubes overlaid atop the (still-visible) physical objects. If augmentation with virtual cues overrides with cues information from the real objects, the SWI will be eliminated and participants will experience the identically sized objects as having the same (true) mass. Furthermore, if holographic size cues are utilized by the sensorimotor system at the expense of the size cues from the physical environment, the objects should be gripped and lifted with similar rates of force.

2 Materials and methods

2.1 Participants

A total of 32 healthy young adults (20 women, 12 men), aged 23.6 years (SD = 3.1) took part in the experiment. Participants were recruited from Munich and tested in the native language (German) at the Human Movement Science lab at the Technical University of Munich, in Germany. Eligible subjects were (1) all adults aged 18 to 40 with (2) a normal or corrected-to-normal vision. Exclusion criteria were (1) a history of neurological diseases or (2) upper limb impairment. All but one person in the AR group was right-handed according the Edinburgh test (Oldfield 1971), ten subjects used their visual aids during the experiment (7 in the AR group). None of the participants reported any visual or sensorimotor problem. Nine subjects had previous VR experience (6 in the AR group) but none of them were familiar with the experimental task or the hypotheses being tested. All subjects included in this study gave written informed consent prior to testing. Ethical approval in accordance with the declaration of Helsinki was obtained from the Ethics Committee of the Medical Faculty of the Technical University of Munich (reference number 175/17 S). Participants were randomly assigned into the AR group or the Control group (Table 1).

2.2 Stimuli and equipment

Subjects were instructed to lift two white plastic cubes of equal weight (390 g) but different size (big cube 10.0 cm × 10.0 cm × 10.0 cm; small cube 6.3 cm × 6.3 cm × 6.3 cm). A metal insert was added to the geometric centre of the small cube to raise its weigh to that of the large cube. A magnetic adaptor mounted in the centre top of both cubes served as a removable connector

Table 1 Comparison of the demographics between the AR and control groups

Group	N	Age in years (range)	Gender (m / f)	Handedness (r / l)
AR	16	23.6 (20–32)	10/6	15/1
Control	16	23.6 (20–29)	3/13	16/0

with a 190 g handle that contained force sensors (Li et al. 2011) and facilitated quick exchange between the two cubes. Sandpaper was attached to the two vertical grasping surfaces to prevent slippage (Fig. 1a). The force sensors registered grip forces applied orthogonally to the grasping surfaces (MAK 177, range 0–100 N, accuracy ± 0.1 N, Rieger, Rheinmünster) as well as the load force acting tangentially to the surfaces along the vertical axis of the handle (MAK 177, range ± 50 N, ± 0.1 N). Signals were transmitted wirelessly to a PC with a sampling rate of 125 Hz.

The AR group wore a Microsoft HoloLens (1st generation) device, an optical-see-through system. When the user is looking through the glasses of the display, three-dimensional virtual cubes which matched the dimensions of the large cube (10.0 cm × 10.0 cm × 10.0 cm) appeared on the physical cubes (Fig. 1b). The small cube was placed on a wooden platform (10 cm x 10 cm × 3.7 cm) such that the top surface of the virtual and physical cubes was aligned (Fig. 1c). Exemplary videos demonstrating the experimental setup from the first-person perspective can be accessed via the Open Science Framework (OSF) https://osf.io/fz368/ (Rohrbach et al. 2020a).

2.3 Augmented reality

For the AR group, the physical cubes were superimposed with slightly transparent cubic holograms. Holograms consist of light points that are projected into the user's field of view. In this article, a hologram refers to the perception of a computer-generated object through stereo imaging. The application for the experiment was developed in Unity 3D version 2017.4 (Unity Technologies 2019). Vuforia (Vuforia Engine 2018) was integrated into the Unity Framework and used to superimpose the virtual cube onto the physical cube. Vuforia offers several functionalities, including target tracking, i.e. the tracking of predefined images. An important aspect for our research question was to develop virtual objects that (1) convey the impression that they were present in the real environment and (2) still allow the subjects to easily see the physical cube. To do so, the holograms which were basic cube primitives, were adjusted in colour, structure, and brightness to create a strong contrast and to optimize the perception of the presented cubes being three-dimensional but also appeared
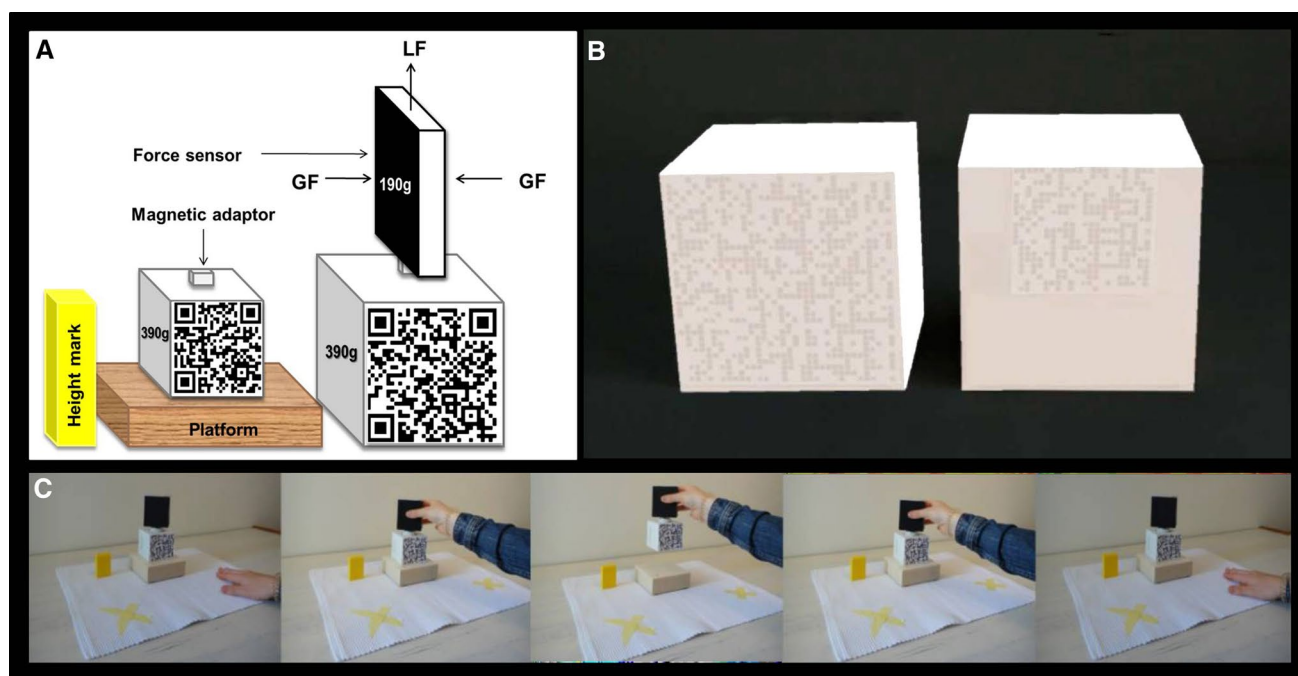
Fig. 1 **a** Schematic of the boxes which were lifted by participants, **b** the view from the participant's perspective in the AR condition, with the identically sized virtual boxes overlaid atop the physical boxes and **c** photographs of the task in action

slightly transparent to enable users to see through. Further, we wanted the virtual objects to remain overlaid atop the physical cube when being moved to increase the sense of presence. We therefore developed patterns, based on QR codes, which were attached to the physical cubes and loaded into the Unity project. This allowed us to precisely scale and align the holographic cubes relative to the cube throughout the experimental trials. The full project code is available at GitHub https://github.com/athierfelder/size-weight-illusion (Rohrbach et al. 2020b).

2.4 Study protocol

During testing, subjects were seated in front of a table and a white wall. First, participants were shown both cubes and asked to estimate the cubes' weights by verbally indicating a self-chosen number (pre-liftoff rating = T0). Note that participants of the AR group did not wear the HMD while giving these pre-liftoff ratings. This was the only moment the boxes were seen simultaneously by the subjects. Participants were then familiarized with the task with five lifting trials solely with the handle, followed by nine pairs of experimental lifts during which they lifted the cubes in alternation.

Prior to each trial, they were asked to close their eyes, at which point a single object was placed on the table in front of them. Following an acoustic signal, they were instructed to open their eyes and lift the cube with their dominant hand utilizing their thumb, middle, and index finger about

up to 5 cm in a smooth and rapid way, hold it steady for 4 s until a second acoustic signal indicated to return back to the starting position. Between trials, while the cubes were exchanged, the subjects were first instructed to close their eyes, and then verbally informed about the actual size of the upcoming physical cube to be lifted (i.e. 'please lift the large/small box').

After the first and the second lifts of the experimental trials (first pair of lifts = T1) and, respectively, after the third (T2), sixth (T3), and ninth pair (T4) participants were again asked to report the felt weight in relation to the value given on the previous trial. Estimations were always given directly after one cube was lifted (i.e. small cube–weight estimation–large cube–weight estimation).

Participants in the AR group carried out a further three pairs of 'retention' lifts, where no augmented boxes were overlaid on the physical boxes, to determine whether any transient effects of AR remained after its removal. Heaviness ratings were further given after the first (T5) and third pair (T6) of these retention lifts. Lifting order was counterbalanced between subjects, and reversed every three pairs of lifts.

At the end of the lifting trials the intervention group was further asked to fill out a questionnaire measuring presence in AR environments (Regenbrecht and Schubert 2002) comprising seven questions that were answered on a 7-point Likert scale. It examines elements of realness (component 1, Q1-3), elements of spatial presence (component 2, Q4-5)

Fig. 2 Average normalized heaviness ratings for the pre-liftoff ratings (T0), the experimental trials (T1-4) and the retention trials (T5,6). Error bars indicate standard error of the mean. The dark blue (black) bars represent the large object and the light blue (grey) bars represent the small object. (Color figure online)
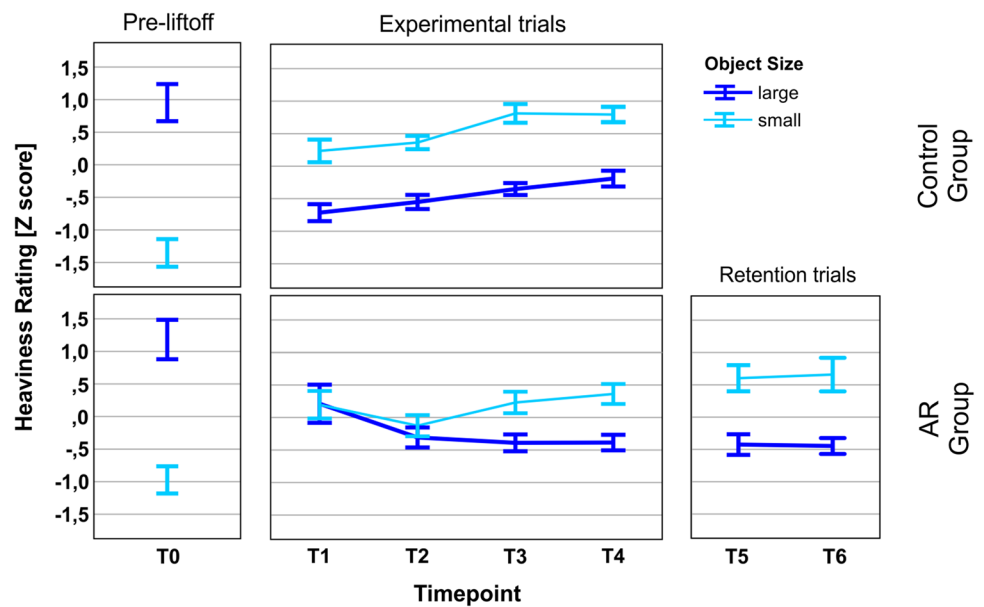
and elements of the experience of perceptual stress (component 3, Q6-7). The detailed questions are accessible via OSF https://osf.io/fz368/ (Rohrbach et al. 2020a).

2.5 Analysis

Numerical ratings of heaviness were standardized to a z-distribution based on the mean and standard deviation of an individual's ratings throughout the course of the experiment (T0 to T4). For the retention trials, which were only conducted in the AR group, means and standard deviation of T0 to T4 in this group were used to calculate the Z-scores of T5 and T6. These ratings were analysed with a mixed design $2 \times 2 \times 4$ ANOVA with one between-group factor (group: AR, Control) and two within-subject factors (object size: large, small) and timepoint (T1, T2, T3, T4). The data from T5, T6 in the AR group were analysed with a separate within-subject 2×2 ANOVA with factors of object size (large, small) and timepoint (T5, T6). The pre-liftoff (T0) ratings were examined with a mixed 2×2 ANOVA with object size (large, small) as the within factor and group (AR, Control) as the between-group factor.

Customized software (GFWin, MedCom, Munich) collected and analysed the data. The grip force was averaged from the signals of the two grasping surfaces and the load force was normalized to a pre-liftoff baseline. The values were differentiated to yield their rates of change. The peak grip force rates (GFR) and peak load force rates (LFR) before liftoff on each trial were used as the dependent variables to examine sensorimotor prediction. If a clear peak was detected (defined as a force rate increase longer than 50 ms to more than 5 N/s and a drop of more than 25% of the peak value before the next increase) before a second higher

peak, the value of the first peak was considered to represent prediction. A single trial from one participant in the control group's LFR data was removed due to an early liftoff and thus uncertainty regarding the baseline normalization. These were examined with a mixed design $2 \times 2 \times 9$ ANOVA with one between-group factor (group: AR, Control) and two within-subject factors (object size: large, small and trial pair: T1-9). The data from the three retention trial pairs were examined in a separate within-subject 2×3 ANOVA with factors of object size (large, small) and trial pair (T10, T11, T12).

All data were examined with Mauchly's test of sphericity prior to statistical analysis. Significant main effects and interactions were followed up with paired tests comparing the large-small ratings/forces at each timepoint within each group. All analyses were conducted in Jamovi version 1.21.

3 Results

3.1 Size–weight illusion

Before lifting the objects (T0), we observed overall higher ratings for the large object than the small object (i.e. a significant main effect of Object Size; $F(1,30) = 114.9$, $p < 0.001$, $\eta_p^2 = 0.79$), but no main effect of Group ($F(1,30) = 1.1$, $p = 0.31$, $\eta_p^2 = 0.03$) or interaction between these variables ($F(1,30) = 0.13$, $p = 0.72$, $\eta_p^2 < 0.01$). Participants in both groups thus experienced normal pre-liftoff ratings of heaviness (Fig. 2).

In the timepoints examined during the experimental trials (T1-4), all main effects and interactions involving the

Timepoint variable failed to meet the assumption of Sphericity, so tests involving these factors had their degrees of freedom adjusted with the Greenhouse–Geisser correction. We observed a significant main effect of Object Size ($F(1,30) = 37.9$, $p < 0.001$, $\eta_p^2 = 0.59$), but no significant effect of Timepoint ($F(2,63.0) = 0.97$, $p = 0.056$, $\eta_p^2 = 0.09$) or of Group ($F(1,30) = 1.1$, $p = 0.31$, $\eta_p^2 = 0.03$). There was a significant interaction between Timepoint and Group ($F(2,63.0) = 5.6$, $p = 0.005$, $\eta_p^2 = 0.16$) and between Object Size and Group ($F(1,30) = 7.6$, $p = 0.01$, $\eta_p^2 = 0.20$). The three-way interaction was not significant ($F(1,61.3) = 1.29$, $p = 0.08$, $\eta_p^2 = 0.04$). Due to the presence of significant interactions, we compared the ratings given to the large and small objects across each trial within each group separately. In the Control group's experimental trials, paired t tests (with a Bonferroni-adjusted threshold of 0.0125 for statistical significance) comparing the heaviness ratings given to the large object compared to the small object at each timepoint found significant differences at Timepoint 1 ($t(15) = 5.1$, $p < 0.001$, $d = 1.27$), Timepoint 2 ($t(15) = 6.2$, $p < 0.001$, $d = 1.54$), Timepoint 3 ($t(15) = 8.0$, $p < 0.001$, $d = 2.0$) and Timepoint 4 ($t(15) = 4.9$, $p < 0.001$, $d = 1.23$). In the AR group's experimental trials, these tests revealed no differences in these ratings on Timepoint 1 ($t(15) = 0.03$, $p = 0.97$, $d < 0.01$) and Timepoint 2 ($t(15) = 0.8$, $p = 0.42$, $d = 0.2$), but robust differences on Timepoint 3 ($t(15) = 3.1$, $p = 0.008$, $d = 0.76$) and Timepoint 4 ($t(15) = 5.1$, $p < 0.001$, $d = 1.3$). These patterns of data indicate that the Control group experienced a normal, unchanging SWI throughout the experiment. The AR group, by contrast, experienced no SWI in early trials, but a normal-looking SWI emerged across repeated lifts.

In the timepoints examined during the retention trials (T6, T7), we observed higher ratings for the small object than the large object (i.e. a significant main effect of object size; $F(1,15) = 31.7$, $p < 0.001$, $\eta_p^2 = 0.68$), but no main effect of Timepoint ($F(1,15) = 0.01$, $p = 0.91$, $\eta_p^2 < 0.01$) or interaction between the variables ($F(1,15) = 0.2$, $p = 0.66$, $\eta_p^2 = 0.01$). Paired t tests (with a Bonferroni-adjusted threshold of 0.0125 for statistical significance) comparing the heaviness ratings given to the large object compared to the small object at each timepoint noted significantly higher ratings for the small object at Timepoint 6 ($t(15) = 4.7$, $p < 0.001$, $d = 1.2$) and Timepoint 7 ($t(15) = 5.6$, $p < 0.001$, $d = 1.4$). Participants in the AR group thus experienced a normal SWI after removing the AR glasses.
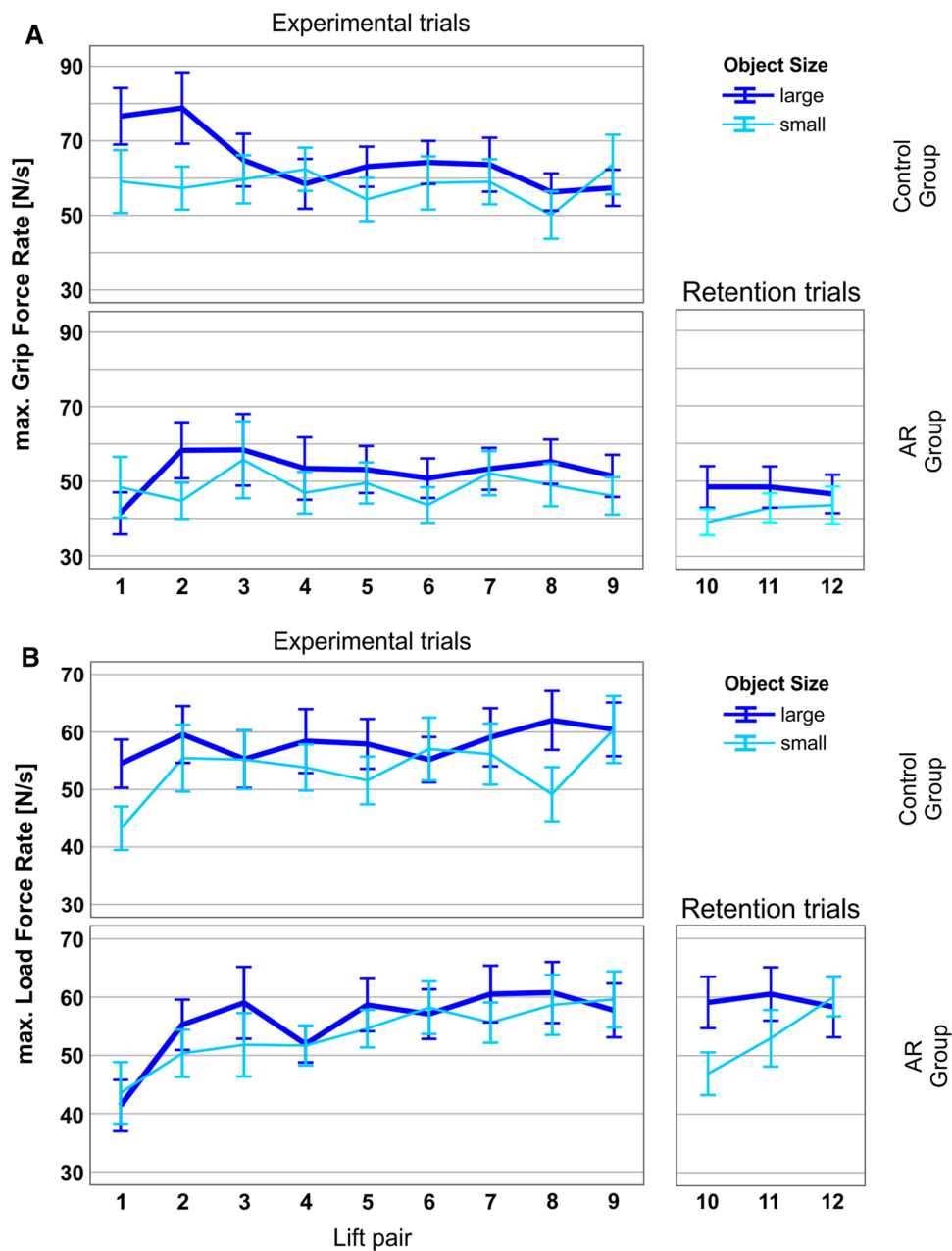
3.2 Fingertip forces

In the experimental trials (T1-9), in terms of peak GFR (Fig. 3a) all main effects and interactions involving the Trial variable failed to meet the assumption of Sphericity, so tests involving these factors had their degrees of freedom adjusted

with the Greenhouse–Geisser correction. We observed a main effect of Object size ($F(1,29) = 15.2$, $p < 0.001$, $\eta_p^2 < 0.001$), but no main effect of Trial ($F(4.03, 116.9) = 1.03$, $p = 0.39$, $\eta_p^2 = 0.03$) or Group ($F(1,29) = 2.08$, $p = 0.16$, $\eta_p^2 = 0.07$). The interaction between Object size and Group ($F(1,29) = 0.42$, $p = 0.52$, $\eta_p^2 = 0.01$) and Object Size and Trial ($F(4.2,121.8) = 1.75$, $p = 0.14$, $\eta_p^2 = 0.06$) were not significant. The interactions between Trial and Group ($F(4.0,116.9) = 2.14$, $p = 0.08$, $\eta_p^2 = 0.07$), as well as the 3-way interaction ($F(4.2,121.8) = 2.15$, $p = 0.08$, $\eta_p^2 = 0.07$), both failed to reach significance due to the corrections for sphericity. As these interactions were borderline, we conducted the within-group analysis reported above for the perceptual ratings of heaviness. In the Control group's lifts, paired t tests (with a Bonferroni-adjusted threshold of 0.0056 for statistical significance) comparing peak GFR of the large object compared to the small object on each trial found that participants used significantly higher rate of force used to grip the large object on Trial 2 ($t(15) = 3.88$, $p = 0.001$, $d = 0.97$). By contrast, in the AR group's lifts, no significant differences emerged (all p values > 0.054). In terms of peak LFR (Fig. 3b) on the experimental trials, all main effects and interactions met the assumption of sphericity. We observed a main effect of Object size ($F(1,28) = 6.17$, $p = 0.019$, $\eta_p^2 = 0.18$) and a main effect of Trial ($F(8,224) = 5.26$, $p < 0.001$, $\eta_p^2 = 0.16$), but no main effect of Group ($F(1,28) = 0.05$, $p = 0.83$, $\eta_p^2 = 0.002$). As no other interactions were significant or borderline (all p values > 0.30), no further post hoc analysis was conducted. In summary, there is some indication that overlaying computer-generated identically sized objects atop physical objects disrupts the normal tendency to grip heavy-looking large objects at a higher rate of force than light-looking small objects.

In the AR group's retention trials (T10-12), all main effects and interactions met the assumption of sphericity. With the GFR data, we observed a significant main effect of Object size ($F(1,15) = 5.06$, $p = 0.04$, $\eta_p^2 = 0.25$), but no main effect of Trial ($F(2,30) = 0.35$, $p = 0.71$, $\eta_p^2 = 0.02$) and no interaction between the variables ($F(2,30) = 1.04$, $p = 0.37$, $\eta_p^2 = 0.07$). For the LFR data, we observed a significant main effect of Object size ($F(1,15) = 5.28$, $p = 0.04$, $\eta_p^2 = 0.26$), but no main effect of Trial ($F(2,30) = 1.19$, $p = 0.32$, $\eta_p^2 = 0.07$). The interaction between the variables was significant ($F(2,30) = 6.03$, $p = 0.006$, $\eta_p^2 = 0.29$), so we compared the force rates used to lift the large and small object on each trial of this phase of the experiment. These paired t tests, with a threshold of 0.017 to achieve statistical significance, revealed that the large object was lifted at a higher rate of force than the small object on Trial 1 ($t(15) = 3.89$, $p = 0.001$, $d = 0.91$), but not the later trials (all p values > 0.04). This analysis suggests that the sensorimotor memories learned

Fig. 3 a Peak grip force rate and **b** peak load force rate for the experimental trials (T1-9) across both groups, and for the retention trials (T10-12) for the AR group. Error bars indicate standard error of the mean. The dark blue (black) bars represent the large object and the light blue (grey) bars represent the small object. (Color figure online)

3.3 Subjective experience of augmented elements

The experienced presence of augmented objects in the physical environment was assessed using a questionnaire (Regenbrecht and Schubert 2002). The results of the presence questionnaire (Table 2) reveal that our augmented cubes were judged low in realness and seemed to be not integrated well with the real objects (component 1, Mode [Q1&Q2] = 1, Mode [Q3] = 2). Spatial presence of our augmented objects was rated high (component 2). Subjects had the impression that the augmented cubes were located in space and experienced them as three-dimensional (Mode [Q4] = 5, Mode [Q5] = 6). Perceptual stress in our sample size was moderate (component 3). The difference between real and virtual drew the subject´s attention (Mode [Q6] = 2), but the perception of the augmented cubes did not need a lot of effort (Mode [Q7] = 4).

when seeing the computer-generated objects does not transfer to interactions with the physical objects upon which they were overlaid.

Table 2 Results of the presence questionnaire

	Q1	Q2	Q3	Q4	Q5	Q6	Q7
Median	1	2	2	4	5,5	2	4
Mode	1	1	2	5	6	2	4
Mean (SD)	1.6 (1.3)	2.4 (1.9)	2.6 (1.8)	4.0 (1.3)	4.8 (1.7)	2.3 (1.9)	4.2 (1.5)
Minimum	0	0	0	0	1	0	0
Maximum	4	6	6	5	6	6	6

4 Discussion

This study sought to examine how computer-generated cues to an object's volume, delivered via AR glasses, might affect perceptions of heaviness and the predictive application or fingertip forces in the context of the SWI. Participants lifted and judged the weight of a pair of boxes with the same mass but different volumes. These stimuli typically induce a robust and unchanging perceptual effect whereby the small object feels heavier than the large object, and a transient sensorimotor prediction whereby small objects are lifted at a lower rate of force than larger objects during the initial lifts. Half of the participants lifted these boxes with normal vision, whereas the other half wore AR glasses which overlaid images of identically sized boxes atop the physical stimuli. The Control group, who lifted the boxes without the AR glasses, experienced a strong SWI throughout the experimental trials, and gripped the large box at a higher rate of force than the small box on early trials. Despite being informed about the real physical size differences the AR group, by contrast, experienced no SWI in the early trials, with a robust illusion emerging over repeated lifts. Furthermore, they showed no evidence that the physical size of the cubes affected the way that they gripped and lifted the object. This propensity to favour the AR cues over the cues from the physical environment was further evidenced through retention trials, where participants in the AR group removed their AR glasses and lifted the objects several more times. Here, despite having lifted the boxes enough times to adapt their forces to the true mass, they lifted the large object with a higher rate of force than the small object–as if they had not undertaken the experimental lifts at all.

The findings from this work help better understand the factors which drive the SWI. First, they highlight the important role that visual cues play in the induction of the SWI, with participants (on early trials) appearing to value the cues to volume over the explicitly delivered 'high-level' information about the size of the object on the upcoming lift, which is consistent with recent work showing that the size of a container completely overwhelms cues to how full the container is (Saccone et al. 2019). Similarly, semantic cues are not sufficient to induce an expectation-driven weight illusion (Naylor et al. 2020). The emergence of the SWI in later trials is also interesting and could suggest that the lack of a

SWI on early trials was simply a consequence of participants being distracted by the novel visual cues. This hypothesis is not, however, consistent with recent work showing that the SWI is not reduced in the presence of a secondary cognitive task (Freeman et al. 2019). One possibility is that the emergence of the SWI might be related to participants' down-weighting the visual cues which did not elicit strong feelings of realness (Table 2) and instead relying on the objects' centre of mass as a cue to size. The perceptual quality of our augmented cues might further have influenced the size perception. Optical-see-through displays can exhibit underestimation of size in augmented objects and is affected by the visualization techniques (Ahn et al. 2019). As both of the lifted objects were cubes, participants would have had access to their physical volume through the moment of their inertia tensor, experienced thought slight deviations from a perfectly vertical lift. This cue is well-established as a way to affect perceptions of heaviness (Amazeen and Turvey 1996; Valdez and Amazeen 2008), but the dynamic switch in dominance from vision to haptic size cues over repeated interactions has not, to our knowledge, been reported. Future work undertaking the opposite paradigm, with differently sized objects overlaid atop objects with identical physical volumes, and stimuli which dynamically alter their visual properties during and between trials (van Polanen and Davare 2019), might help disentangle these possibilities. Increasing the sense of realness by integrating a 3D scanned mesh model of the physical cubes might further influence the size perception and potentially the experienced SWI.

Arguably more important, however, are what these findings mean in the context of AR. This technology is being used widely across society, from the resurgence of heads-up displays to overlay key information atop drivers' perspective of the road in cars, to so-called 'smart glasses' such as those employed in the current work. And, while this study is not the first to show that computer graphics can influence perception enough to induce the SWI (Buckingham 2019; Heineken and Schulte 2007; Kawai et al. 2007), it is the first indication that these graphical elements can take precedence over visible cues from the physical environment for a period of time, eliminating an otherwise robust perceptual illusion. The findings related to the fingertip forces during initial lifts is also particularly noteworthy. The control group showed broadly the expected pattern of data, using higher rates of

fingertip forces to interact with the large object than the small object on early trials. The AR group, by contrast, lifted both boxes with very similar rates of force from the initial trials, and continued to do so until the end of the study. When the AR goggles were removed, they lifted the boxes as if they were doing so for the first time (i.e. analogous to the control group's initial interactions). Together, these data suggest that the computer-generated objects displayed through the AR goggles were treated by participants' sensorimotor systems as if they were real, in the sense that the artificial cues were prioritized to drive this form of dextrous behaviour. Designers of AR content should take heed that the images overlaid on the physical environment can affect this ostensibly automatic behaviour.

Data Availability All project files are publicly available at the Open Science Framework website: https://osf.io/fz368/ (Rohrbach et al. 2020a). This includes the perceptual data and fingertip force data as well as exemplary videos demonstrating the experimental setup from the first-person perspective.

Code availability The full project code is available on GitHub https://github.com/athierfelder/size-weight-illusion (Rohrbach et al. 2020b).

Declarations

Conflicts of interest The authors declare that they have no competing interests with regards to the submitted work.

Ethics approval Ethical approval for the study was given by the Ethics Committee of the Medical Faculty of the Technical University of Munich (reference number 175/17 S). All participants gave written informed consent.

References

Ahn J, Ahn E, Min S, Choi H, Kim H, Kim GJ (2019) Size perception of augmented objects by different AR displays. In C. Stephanidis (ed.), HCI International 2019–Posters (pp. 337–344). Springer International Publishing. doi: https://doi.org/10.1007/978-3-030-23528-4_46

Al-Issa H, Regenbrecht H, Hale L (2012) Augmented reality applications in rehabilitation to improve physical outcomes. Phys Therapy Rev 17(1):16–28

Allcoat D, von Mühlenen A (2018) Learning in virtual reality: Effects on performance, emotion and engagement. Res Learn Technol. https://doi.org/10.25304/rlt.v26.2140

Amazeen EL, Turvey M (1996) Weight perception and the haptic size–weight illusion are functions of the inertia tensor. J Exp Psychol Hum Percept Perform 22(1):213–232. https://doi.org/10.1037/0096-1523.22.1.213

Bideau B, Kulpa R, Vignais N, Brault S, Multon F, Craig C (2010) Using virtual reality to analyze sports performance. IEEE Comput Graph Appl 30(2):14–21. https://doi.org/10.1109/MCG.2009.134

Buckingham G (2014) Getting a grip on heaviness perception: A review of weight illusions and their probable causes. Exp Brain Res 232(6):1623–1629. https://doi.org/10.1007/s00221-014-3926-9

Buckingham G, Goodale MA (2013) Size matters: a single representation underlies our perceptions of heaviness in the size-weight illusion. PLoS ONE 8(1):e54709. https://doi.org/10.1371/journal.pone.0054709

Buckingham G, MacDonald A (2016) The weight of expectation: Implicit, rather than explicit, prior expectations drive the size–weight illusion. Q J Exp Psychol 69(9):1831–1841. https://doi.org/10.1080/17470218.2015.1100642

Buckingham G, Milne JL, Byrne CM, Goodale MA (2015) The size-weight illusion induced through human echolocation. Psychol Sci 26(2):237–242. https://doi.org/10.1177/0956797614561267

Buckingham G, Ranger NS, Goodale MA (2011) The role of vision in detecting and correcting fingertip force errors during object lifting. J Vision. https://doi.org/10.1167/11.1.4

Buckingham G (2019) Examining the size-weight illusion with visuo-haptic conflict in immersive virtual reality. Quarterly J Experim Psychol (2006) 72(9):2168–2175. https://doi.org/10.1177/1747021819835808

Charpentier A (1891) Analyse expérimentale quelques éléments de la sensation de poids. Archives de Physiologie Normales et Pathologiques 3:122–135

Davis CM, Roberts W (1976) Lifting movements in the size-weight illusion. Percept Psychophys 20:33–36

Dey A, Billinghurst M, Lindeman RW, Swan JEI (2018) A systematic review of 10 years of augmented reality usability studies: 2005 to 2014. Frontiers in Robotics and A I:5. https://doi.org/10.3389/frobt.2018.00037

Dijker AJM (2014) The role of expectancies in the size-weight illusion: A review of theoretical and empirical arguments and a new explanation. Psychon Bull Rev. https://doi.org/10.3758/s13423-014-0634-1

Flanagan JR, Beltzner MA (2000) Independence of perceptual and sensorimotor predictions in the size-weight illusion. Nat Neurosci 3(7):737–741. https://doi.org/10.1038/76701

Freeman CG, Saccone EJ, Chouinard PA (2019) Low-level sensory processes play a more crucial role than high-level cognitive ones

in the size-weight illusion. PLoS ONE 14(9):e0222564. https://doi.org/10.1371/journal.pone.0222564

Gordon AM, Forssberg H, Johansson RS, Westling G (1991) Visual size cues in the programming of manipulative forces during precision grip. Exp Brain Res 83(3):477–482

Grandy MS, Westwood DA (2006) Opposite Perceptual and Sensorimotor Responses to a Size-Weight Illusion. J Neurophysiol 95(6):3887–3892. https://doi.org/10.1152/jn.00851.2005

Heineken E, Schulte FP (2007) Seeing size and feeling weight: The size-weight illusion in natural and virtual reality. Hum Factors 49(1):136–144

Kawai S, Henigman F, MacKenzie CL, Kuang AB, Faust PH (2007) A reexamination of the size-weight illusion induced by visual size cues. Exp Brain Res 179(3):443–456. https://doi.org/10.1007/s00221-006-0803-1

Li Y, Randerath J, Goldenberg G, Hermsdörfer J (2011) Size–weight illusion and anticipatory grip force scaling following unilateral cortical brain lesion. Neuropsychologia 49(5):914–923. https://doi.org/10.1016/j.neuropsychologia.2011.02.018

Metcalfe RW (2007) The size-weight illusion in a natural and augmented environment with congruent and incongruent size information. Doctoral Dissertation, School of Kinesiology, Simon Fraser University

Naylor CE, Power TJ, Buckingham G (2020) Examining Whether Semantic Cues Can Affect Felt Heaviness When Lifting Novel Objects. J Cogn 3(1):3. https://doi.org/10.5334/joc.93

Nowak DA, Hermsdoerfer J (2009) Sensory control of object manipulation. Cambridge University Press, In Sensorimotor Control of Grasping. https://doi.org/10.1017/CBO9780511581267

Oldfield RC (1971) The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia 9(1):97–113

Peters MAK, Ma WJ, Shams L (2016) The Size-Weight Illusion is not anti-Bayesian after all: A unifying Bayesian account. PeerJ 4:e2124. https://doi.org/10.7717/peerj.2124

Plaisier MA, Kuling IA, Brenner E, Smeets JBJ (2019) When does one decide how heavy an object feels while picking it up? Psychol Sci 30(6):822–829. https://doi.org/10.1177/0956797619837981

Regenbrecht H, Schubert T (2002) Measuring presence in augmented reality environments: design and a first test of a questionnaire. In: Proceedings of the Fifth Annual International Workshop Presence 2002. http://citeseerx.ist.psu.edu/viewdoc/citations;jsessionid=EE4E7F62FF8FBBC62112677A08670CAA?doi=10.1.1.581.2484

Rohrbach N, Buckingham G, Hermsdörfer J, Thierfelder A, Huber L-M (2020a) Project files for "Fooling the size-weight-illusion – Using augmented reality to eliminate the effect of size on perceptions

of heaviness and sensorimotor prediction". Open Science Framework. Retrieved from https://osf.io/fz368

Rohrbach N, Buckingham G, Hermsdörfer J, Thierfelder A, Huber L-M (2020b) Project code for "Fooling the size-weight-illusion – Using augmented reality to eliminate the effect of size on perceptions of heaviness and sensorimotor prediction". GitHub repository. Retrieved from https://github.com/athierfelder/size-weight-illusion

Rohrbach N, Chicklis E, Levac DE (2019a) What is the impact of user affect on motor learning in virtual environments after stroke? A scoping review. J Neuroeng Rehabil 16(1):79. https://doi.org/10.1186/s12984-019-0546-4

Rohrbach N, Gulde P, Armstrong AR, Hartig L, Abdelrazeq A, Schröder S, Neuse J, Grimmer T, Diehl-Schmid J, Hermsdörfer J (2019b) An augmented reality approach for ADL support in Alzheimer's disease: A crossover trial. J Neuroeng Rehabil 16(1):66. https://doi.org/10.1186/s12984-019-0530-z

Saccone EJ, Goldsmith RM, Buckingham G, Chouinard PA (2019) Container size exerts a stronger influence than liquid volume on the perceived weight of objects. Cognition 192:104038. https://doi.org/10.1016/j.cognition.2019.104038

The Jamovi project (2020). *jamovi* (Version 1.21.6) [Computer Software]. Retrieved from https://www.jamovi.org

Therapy Lens. *Augmented Reality in Rehabilitation.* Retrieved November 12, 2020, from http://www.therapylens.com/en/home/

Valdez AB, Amazeen EL (2008) Sensory and perceptual interactions in weight perception. Percept Psychophys 70(4):647–657. https://doi.org/10.3758/PP.70.4.647

van Polanen V, Davare M (2019) Dynamic size-weight changes after object lifting reduce the size-weight illusion. Scientif Rep 9(1):15697. https://doi.org/10.1038/s41598-019-52102-y

van Polanen V, Tibold R, Nuruki A, Davare M (2019) Visual delay affects force scaling and weight perception during object lifting in virtual reality. J Neurophysiol 121(4):1398–1409

Weser V, Proffitt DR (2019) Making the visual tangible: substituting lifting speed limits for object weight in VR. *PRESENCE: Virtual and Augmented Reality*, *27*(1), 68–79.

Wolf C, Tiest WMB, Drewing K (2018) A mass-density model can account for the size-weight illusion. PLoS ONE 13(2):e0190624. https://doi.org/10.1371/journal.pone.0190624