The Perception of Affordances in Mobile Augmented Reality

Yu Zhao
Elect. Engr. & Comp. Science
Vanderbilt University
USA
yu.zhao@vanderbilt.edu

Jeanine Stefanucci
Department of Psychology
University of Utah
USA
jeanine.stefanucci@psych.utah.edu

Sarah H. Creem-Regehr
Department of Psychology
University of Utah
USA
sarah.creem@psych.utah.edu

Bobby Bodenheimer
Elect. Engr. & Comp. Science
Vanderbilt University
USA
bobby.bodenheimer@vanderbilt.edu

ABSTRACT

Today, augmented reality (AR) is most easily experienced through a mobile device such as a modern smartphone. For AR to be useful for applications such as training, it is important to understand how people perceive interactions with virtual objects presented to them via mobile AR. In this paper, we investigated two judgments of action capabilities (affordances) with virtual objects presented through smartphones: passing through an aperture and stepping over a gap. Our goals were to 1) determine if people can reliably scale these judgments to their body dimensions or capabilities and 2) explore whether cues presented in the context of the action could change their judgments. Assessments of perceived action capabilities were made in a pre/post-test design in which observers judged their affordances towards virtual objects prior to seeing an AR cue denoting their body dimension/capability, while viewing the cue, and after seeing the cue. Different patterns of results were found for the two affordances. For passing through, estimates became closer to shoulder width in the post-cue compared to the pre-cue block. For gap stepping, estimates were closer to actual stepping capability while viewing the cue, but did not persist when the cue was no longer present. Overall, our findings show that mobile smartphones can be used to assess perceived action capabilities with virtual targets and that AR cues can influence the perception of action capabilities in these devices. Our work provides a foundation for future studies investigating perception with the use of mobile AR with smartphones.

CCS CONCEPTS

• **Human-centered computing** → **Mixed / augmented reality**; **Empirical studies in HCI**; • **Applied computing** → **Psychology**.

KEYWORDS

augmented reality, perception, affordances, mobile devices

ACM Reference Format:
Yu Zhao, Jeanine Stefanucci, Sarah H. Creem-Regehr, and Bobby Bodenheimer. 2021. The Perception of Affordances in Mobile Augmented Reality. In *ACM Symposium on Applied Perception 2021 (SAP '21), September 16–17, 2021, Virtual Event, France.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3474451.3476239

1 INTRODUCTION

AR technology is now widely available at the consumer level with modern smartphones. Mobile AR applications for entertainment, education, healthcare, training, and communication have been developed, and development is both expanding and accelerating [Adhani and Rambli 2012; Siriwardhana et al. 2021]. As these applications become more widespread, a better understanding of user interaction with AR for mobile devices is needed. For example, one opportunity for learning with mobile AR is a situated simulation [Liestøl 2009], in which users are situated in a locale and see an AR simulation of an event through a mobile platform. It will generally be helpful to learning if both virtual and real content be seamless and fluid. One problem with the current technology is that virtual objects may not be perceived in the same way as real world objects. If this happens, then possibilities for action and interaction with real and virtual objects may differ. This difference, called a loss of *perceptual fidelity* [Pointon et al. 2018a], can lead to applications not working as well as intended.

This paper studies the perceptual fidelity of modern augmented reality as perceived through smartphones. We do this by studying how judgments of action possibilities with virtual objects are perceived by users of AR. Possibilities for action in an environment, whether real, virtual, or mixed, are called *affordances*. The concept of affordances was proposed by Gibson [Gibson 1979]. Formally, they are observers' perceptions of their ability to act given the constraints and properties of the environment relative to their own body dimensions and capabilities. Affordances can be used to provide a useful, objective measure of the perceptual fidelity of a virtual or augmented environment [Bhargava et al. 2020b; Creem-Regehr et al. 2019; Gagnon et al. 2020; Geuss et al. 2010; Pointon et al. 2018a,b; Stefanucci et al. 2015].

SAP '21, September 16–17, 2021, Virtual Event, France

ACM ISBN 978-1-4503-8663-0/21/09…$15.00
https://doi.org/10.1145/3474451.3476239

The prior work conducted on affordances and perceptual fidelity is different from the present work in that it is mostly concentrated on virtual or augmented reality as presented through head-mounted displays rather than mobile smartphone displays. While there has been some work looking at perceptual issues in augmented reality on tablets or mobile devices [Liu et al. 2020; Livingston et al. 2009], this work has not evaluated perception of action capabilities, particularly through an affordance lens. Understanding how users perceive action capabilities in smartphones would be a significant advance given that phones are fundamentally different than head-mounted displays and will also be much more common for the foreseeable future. Evaluating their ability to convey affordances to users could lead to better applications in a variety of domains, particularly because smartphones work in a significantly larger number of environments than current head-mounted displays do. Many head-mounted displays need a powerful computer to drive their graphics. And while some displays are mobile, they often have limitations on where they can be used, e.g., the HoloLens does not function well outdoors in bright sunlight. Thus, a second difference between the present work and prior work is that we attempt to evaluate the perceptual fidelity of augmented reality in smartphones in realistic, everyday contexts, i.e., in the wild. This gives our experiments more inherent variance, but through careful study design we can learn about perceptual fidelity of these devices in the situations in which they are most likely to be used.

Specifically, then, in this paper we test two action capabilities in augmented reality as perceived through a smartphone: passing through and stepping over. These are perceived affordances that have been well studied in the real world and in virtual reality (VR). In our tasks, people make affordance judgments about passing through and stepping over virtual objects embedded in the real world as perceived through mobile smartphones. In addition, we allow for training of those judgments by providing users with an augmented reality cue that visually indicates their action capability in the context of the task. We test for whether the cue generalizes to future estimates by comparing judgments of action capability that are made prior to seeing the augmented reality cue versus after experiencing the cue.

This paper thus makes the following contributions:

(1) Quantified judgments of action capabilities for two affordances in mobile augmented reality.
(2) An improved understanding of the role of feedback and training on affordance judgments in mobile augmented reality.
(3) A framework for judging perceptual fidelity of mobile augmented reality in the wild.

2 BACKGROUND

A significant amount of research has investigated the perception of affordances [Gibson 1979] in the real world [Jamone et al. 2018]. How we perceive the environment and the objects within it can be assessed in terms of judgments about possibilities for action [Franchak and Adolph 2014]. These judgments also apply to interactions with everyday objects [Norman 2002]. The general finding is that when people are asked to make judgments about affordances, they scale their judgments (e.g., yes I can perform this action or no I cannot) to the dimensions of their body. For example, an object affords grasping only when it is small enough to fit within one's hand. In this paper, we focus on two previously studied affordances: passing through an aperture [Franchak and Adolph 2012; Stefanucci and Geuss 2009; Warren and Whang 1987] and stepping over a gap [Jiang and Mark 1994; Plumert and Schwebel 1997]. For passing through, participants must take into account their shoulder width when judging whether they can pass through an aperture without turning their bodies [Warren and Whang 1987]. A large body of work has found that participants make judgments that indicate the smallest passable aperture that they can pass through without shoulder rotation is somewhat larger than their shoulder width [Franchak et al. 2012; Higuchi 2013; Warren and Whang 1987]. In contrast, judgments of gap stepping tend to be slightly overestimated, where the largest gap that observers judge that they can just step over is often slightly larger than their actual stride length [Plumert 1995].

2.1 Affordance Judgments in VR and AR

The perception of affordances has also been studied in VR and AR as presented through head-mounted displays [Creem-Regehr et al. 2019; Gagnon et al. 2020; Geuss et al. 2010; Jun et al. 2015]. For VR, people make affordance judgments in virtual environments (VEs) displayed with commodity-level, VR hardware in a manner similar to those made in the real world (for exceptions, see [Bhargava et al. 2020a; Ebrahimi et al. 2018]). Although real world comparisons were not conducted in the following studies, immersive VR has also been used to assess judgments of whether a gap is crossable [Creem-Regehr et al. 2019; Geuss et al. 2016; Jun et al. 2015]. Findings suggest that estimates of whether a gap can be stepped across are similar to what has been observed in the real world in prior work, in that action capabilities are somewhat overestimated (i.e., observers believe they can step over gaps that are slightly larger than they are actually capable of crossing).

Less work has been done to assess affordances in AR, and the existing work has mostly been done in optical see-through displays such as the HoloLens [Gagnon et al. 2020; Gagnon et al. 2021; Pointon et al. 2018a,b; Wu et al. 2019]. The comparison of affordance judgments made in AR to the real world is mixed, with some affordance judgments being similar across environments and others being different. For example, Pointon et al. [Pointon et al. 2018b] investigated two different perceived affordances using the Microsoft HoloLens (version 1). In their study, participants were asked to judge whether they could pass through an aperture conveyed by two poles as in Geuss et al. [Geuss et al. 2010] and whether they could step over a gap as in Jun et al. [Jun et al. 2015]. Judgments of whether or not participants believed that they could perform these actions with apertures and gaps of varying widths was assessed in both the real world and in the HoloLens. They found that judgments of passing through were similar in AR and the real world, but that stepping over judgments were underestimated relative to performance in the real world. Recently, Gagnon et al.[Gagnon et al. 2021] found that the reduced field of view of the HoloLens may contribute to the underestimation of stepping affordances observed in AR. Consistent with the mixed nature of affordance findings so far in AR, we test two different affordances in this paper.

2.2 Space Perception in Mobile AR

Findings on accuracy of distance perception in mobile AR are mixed with results varying based on the range of distances tested [Liu et al. 2020; Swan et al. 2017], the environment in which perception was assessed (e.g., indoor vs. outdoor) [Dey et al. 2010; Livingston et al. 2009], the type of device used [Dey et al. 2012], and the AR depth cues available [Berning et al. 2014; Do et al. 2020; Kruijff et al. 2010]. Livingston et al.[Livingston et al. 2009] used a depth matching protocol in mobile AR and found underestimation of distance indoors, but overestimation of distance outdoors. However, Dey et al.[Dey et al. 2010] observed an underestimation of distance to occluded objects in an outdoor environment, which did not replicate the findings of Livingston et al. In further work, Dey et al.[Dey et al. 2012] investigated egocentric, exocentric, and ordinal depth perception in mobile AR on both tablets and mobile phones. They found more severe underestimation of distance on the tablet compared to the mobile phone for egocentric and exocentric judgments. Most relevant to the current experiments is work done by Swan et al.[Swan et al. 2017] that compared distance estimates to real and virtual objects using a tablet AR system with a distance bisection task. Depth distortion was observed in AR compared to the real world, such that observers overestimated distances up to 15 meters, but then underestimated at 30 meters and beyond. Recently, Liu et al. [Liu et al. 2020] also tested users' perception of distances to real and virtual objects presented in both indoor and outdoor environments through an AR platform for a mobile device. They noted that the virtual objects were perceived as farther than the real objects at closer distances, but environmental context did not affect these results. The results from an experiment conducted by Chakraborty et al. [Chakraborty et al. 2021] replicated the findings of Liu et al. [Liu et al. 2020], and their initial findings suggested that added animation does not help participants to estimate egocentric distance to a life-size, virtual avatar in medium-field spaces. Finally, Do et al. [Do et al. 2020] showed that color and luminance play a role in depth perception for 3D objects in mobile AR. To our knowledge, no prior work has assessed affordance judgments on a mobile device such as a smartphone.

2.3 AR Applications for Training

In addition to assessing perception of affordances, the current work is novel in that it investigates whether using a visual cue to signify the size of the user in the display will improve affordance judgments in a pre-test compared to post-test design. Widespread use of AR for training is already underway with mobile devices, e.g., indoor wayfinding in dark environments [Diao and Shih 2018] and troubleshooting for aircraft engine repair [Rios et al. 2013]. Chatzopoulos et al. [Chatzopoulos et al. 2017] provide a more in-depth review of the field citing other mobile AR training applications available on smartphones and tablets.

3 RATIONALE AND HYPOTHESES

Because there is good evidence that AR training is effective in other domains, we decided to test training of action capabilities in the current experiments, in the context of also assessing general performance for making affordance judgments via a mobile smartphone. Across two experiments we employ a pre-cue, cue, and then post-cue design to test whether virtual objects afford passing through (Experiment 1) or stepping over (Experiment 2) as displayed via mobile AR on a smartphone. In the middle block of trials (the cue block), an AR cue is generated in the context of the virtual affordance in order to indicate to users their actual capabilities and potentially improve their judgments. We tested the following hypotheses across the two experiments:

- **H1** Users will reliably estimate both affordances in mobile AR, and patterns of over- or underestimation relative to actual capabilities will parallel what has been observed in prior real world and VR/AR studies.
- **H2** Viewing an AR cue depicting body dimensions or capabilities will move users' estimates for both affordances closer to their actual capabilities.
- **H3** Experience with the cue will persist after the cue is removed, revealing different estimates in the post-cue block compared to baseline performance in the pre-cue block for judgments of both actions (passing through and stepping over).

4 APPLICATION DESIGN

We deployed the applications in a distributed manner through the Apple App Store and Google Play Store. Thus participants could be recruited widely, could run the applications in their own local surroundings, and did not need to come to our laboratory. The applications for this experiment were coded in Unity (v. 2019.2.21f). Unity's cross-platform API ARFoundation (version 2.0.2) was used to support ARCore and ARKit for enabling AR functionality on both the Android and iOS smartphone operating systems. There were two apps, one for each affordance, for each operating system published on the respective app repository. A smartphone that has AR capability was required to run them, e.g., an iPhone running iOS 11 or later, or an Android phone running Android 7.0 (Nougat) or later.

The API provides interfaces to the detection and placement capabilities of ARCore and ARKit, which allowed us to deploy our virtual objects in the proper locations. We did not directly measure the calibration of the detected AR plane to our ground plane in this paper, although others have, finding them to be quite accurate [Hasler et al. 2019; Nowacki and Woda 2020]. Instead, because we were concerned about reliability and accuracy of cue placement across various environments, we tested our applications extensively in different environments (e.g., indoor and outdoor, environments with different light sources, different times of day, etc.) and on different mobile phones before we deployed them broadly. Further, if participants encountered questions or problems when participating, especially any potential issues with the AR features, we told them to contact us for remote support in order to complete the experiment reliably. To retrieve experimental data from participants, we coded in an email account to send data files back as an attachment. Specifically, we used a Gmail account with a 2-step verification process. To insure the robustness of our data retrieval, we used a Google form to save data when there was an exception to the email process.

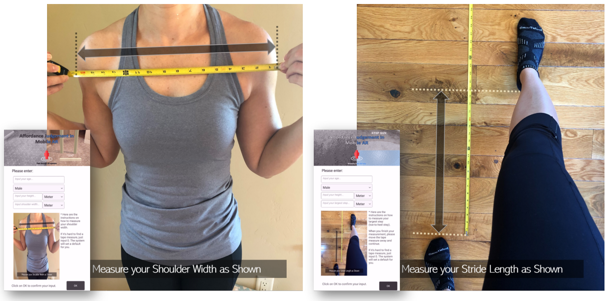

Figure 1: Instructions for measuring shoulder width(left) and maximum stride length (right) as shown exactly to participants while performing the task.

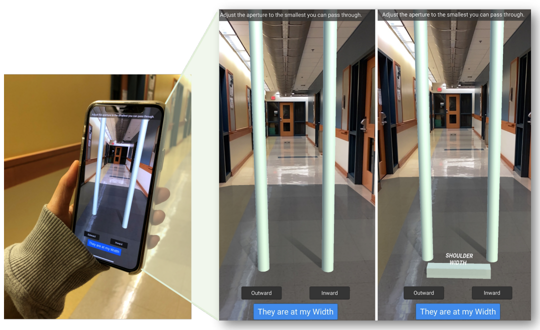

Figure 2: Screenshot of passing through apertures scenario (without the AR cue on left and with the AR cue on the right).

5 EXPERIMENT 1: PASSING THROUGH

5.1 Participants

We recruited participants through email. Thirty-two unpaid volunteers (17 females, 15 males) with a mean age of 32.15 years (range 19 to 68 years) completed the experiment. All gave informed consent for their participation via an online consent form.

5.2 Materials

Participants accessed and completed the experiment via an online application downloaded to their smartphone. Their smartphone needed to be recent enough that it supported the AR SDK for either iOS or Android as described in Section 4. Instructions within the app asked participants to acquire a tape measure to enter into the app their relevant body dimension for the action judgment (in this case, their shoulder width) before any trials began. In Experiment 1, there were 5 subjects who used an Android device and 27 subjects who used an iOS device to run the application. The Unity game engine (v. 2019.2.21f) was used to create the AR application.

5.3 Procedure

Once participants downloaded the app and consented to participate, they were given a brief overview of the experiment. They were then asked to stand in front of a mirror and to measure the distance between the edge of each shoulder as precisely as possible by extending the tape measure across their chest in the widest area of their shoulders. Participants were given explicit instructions on how to measure this width, and shown a photograph illustrating the procedure they were to follow (Figure 1). They were then asked to input their shoulder width in the app, as well as their age, gender and height. Shoulder width and height measurements could be entered in either meters, inches, feet, or yards. If for some reason the participant could not find a tape measure, the system set a default value for their shoulder width to 0.5m (which is the average shoulder width observed in prior experiments on judgments of passing through as reported by Pointon et al. [Pointon et al. 2018b] as well as the average observed in this experiment). Before participants started the experiment, they were also required to find a clear open space that extended several meters, was well lit, and did not have a shiny floor. Once they found a suitable place, they were told to stand still to start the experiment. We asked participants to make judgments in a static pose since some work suggests that calibration of affordances occurs with basic movements that generate optic flow, even if the movements are not specific to the affordance task at hand [Mark et al. 1990; Stoffregen et al. 2009; Yu et al. 2010]. We did not want to include additional optic flow as a cue to help make the judgments. Before any perceptual judgments were made, participants were asked to scan the ground plane with their phones until visual recognition of the ground was complete. The detected ground plane was then used to generate a pair of virtual poles (each 3.6m in height and 0.1m in diameter) two meters in front of the participant (Figure 2). Therefore, passing through affordance judgments will be made from action space distances [Cutting and Vishton 1995], i.e., greater than reaching distance, which is also typical for this action in the real world [Warren and Whang 1987].

Participants' task in the experimental trials was to adjust the two poles to be the width that they believed they could just pass through without turning their shoulders or body in any way. First, participants completed a block of baseline, pre-cue trials which consisted of four pass-through judgments to assess baseline ability to make judgments (see Figure 2). The four adjustment trials in this pre-cue block included two "ascending" trials (A; the poles started at 70% participants' input shoulder width) and two "descending" trials (D; the poles started at 180% participants' input shoulder width), always following the same order: A, D, A, D. We employed this design to prevent potential bias from always having started with narrow or wide gaps. To address hysteresis effects—where starting wide tends to result in a wider aperture than starting narrow—we alternated the ascending and descending trials and averaged across them in the analysis. For each pre-cue trial, the aperture width could be adjusted by clicking on either the button labeled "Inward" or "Outward" to make the distance between the poles wider or narrower. The width between the poles would decrease or increase 0.02m on each click. Once the width of the poles was adjusted to the smallest width that participants believed that they could just pass through without turning their body, they hit a button to confirm the adjustment was done and this moved them to the next trial.

After completing the pre-cue block of trials, participants were asked to stay in their current standing location and were then presented with a virtual cue to help them better understand their shoulder width (see Figure 2, right). They then completed four trials in this cue feedback trial block. The cue block also followed the pattern: A, D, A, D as in the pre-cue block. The only difference in this block of trials was the presence of an AR cue in front of the camera that represented the input shoulder width of the participant. This shoulder width cue was attached to the recognized ground plane and could be dragged along the horizontal plane by participants when they put their finger on it. The cue was placed on the ground; initial designs with the cue shoulder height led to difficulties in perceiving the distance and size of the cue. Participants could use the AR cue to help make their judgements about the width of the poles that were just passable for them. As in the previous block of trials, the aperture width could be adjusted by clicking on either the button labeled "Inward" or "Outward" to make the distance between the poles wider or narrower for each trial.

To assess whether the virtual cue helped participants recalibrate their perceptual judgments of just-passable aperture width, a third block of post-cue trials was conducted. To begin this block of trials, participants were instructed to take four steps forward and then to turn around and face the opposite direction. This turning was required in order to reduce reliance on or use of real world references in order to make more accurate estimates. The walking and turning also allowed us to assess whether training with the cue generalized to a new location in the environment. In order to relocate the camera of the phone after walking four steps forward and turning around, participants re-scanned the environment to detect the ground plane. Again, this post-cue block of trials followed the A, D, A, D pattern, but no AR cue was present. For each post-cue trial, the aperture width could be adjusted by clicking on either the button labeled "Inward" or "Outward" to make the distance between the poles wider or narrower.

In total, participants completed 12 trials over the 3 blocks. The second block (cue) provided a shoulder width AR cue, while the first (pre-cue) and third block (post-cue) did not. After finishing all trials, participants were asked to subjectively rate the usefulness of the AR cue presented in the second block of trials. A five-point Likert scale (ranging from extremely useless to extremely useful) was given. Finally, participants clicked on the "Click to Finish" button to send their recorded data back to the experimenter.

6 EXPERIMENT 2: STEPPING OVER

6.1 Participants

Thirty-one unpaid volunteers (17 females, 14 males) with mean age of 32.19 years (range 19 to 68) completed the experiment. All gave informed consent for their participation via an online consent form.

6.2 Materials

The materials were the same as in Experiment 1, except that the application tested participants' abilities to judge whether they could step over gaps of varying widths. Instructions in the app asked participants to acquire a tape measure to enter their relevant body dimension for the action judgment (in this case, their trailing toe to leading heel step length for the largest step they could take without having either foot off the ground) into the app before any trials began. There were 5 subjects who used an Android device and 26 subjects who used an iOS device to run the application.

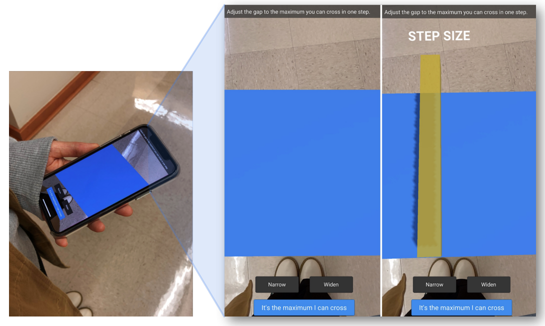

Figure 3: Screenshot of stepping over a gap scenario (without AR cue on left and with AR cue on the right).

6.3 Procedure

The procedure was similar to Experiment 1. In the stepping affordance judgment task, besides entering their age, gender, and height, participants were told to measure their maximum step length from trailing toe to leading heel. If for some reason the participant could not find a tape measure, the system set their step length to 0.88m as the default, which is the average step length for adults observed in prior experiments on judgments of stepping over a gap [Creem-Regehr et al. 2019].

There were also 12 trials over 3 blocks (pre-cue, cue, post-cue) in the experiment. The first block contained four pre-cue trials as in Experiment 1 to gauge baseline judgments of stepping over. Before making any judgment, participants were asked to scan the ground plane with their phones to generate a blue virtual gap at their feet (see Figure 3). The virtual gap was 2m horizontally in length and began as 50% of the input stride width on ascending trials and 140% of the input stride width on descending trials. In each trial, participants were instructed to click on either "Narrow" or "Widen" to adjust the gap to be the width that they believed they could step over with one step. Gap width would increase or decrease 0.05m on each click. Once the width of the gap had been adjusted to what participants believed they could just cross with one stride, they hit a button to record the width and move to the next trial. In the cue block, an AR cue that represented the extent of their maximum step size was shown on one side of the ground plane to help participants make more accurate judgments for stepping (see Figure 3, right). As in Experiment 1, before the beginning of the post-cue block, participants were asked to take 4 steps, turn around and face the opposite direction before completing the remaining 4 trials of the experiment. The change of viewing direction was implemented to limit participants' abilities to use other cues in the environment to make their estimates after seeing the AR cue, and also involved re-scanning the ground plane. When all 12 trials were completed, participants rated the usefulness of the AR cue with the same 1 to 5 Likert Scale in Experiment 1 and then clicked the "Finish" button to send their data back to the experimenter.

7 RESULTS

7.1 Experiment 1: Passing Through

We averaged the four trials from each block and calculated ratios of mean estimated just-passable aperture width to actual measured shoulder width. Actual measured shoulder width ranged from 0.36 m to 1.02 m (M = 0.49 m). One participant was excluded from analysis because the measured shoulder width (1.02 m) was greater than 3 SD above the mean. After removing the outlier, the mean of measured shoulder widths was M =0.47m (SD =0.08 m). Seven participants recorded the default 0.5 m width. (Supplementary analyses with the seven participants removed from the dataset are presented in the Appendix). Scaling to measured shoulder width gives us a measure of how accurate participants' perception of passing through was relative to their own body widths. A ratio greater than 1.0 indicates that an observer judges that the aperture needs to be larger than their shoulder width to pass through; a ratio below 1.0 indicates that observers judge they could pass through an aperture that they could not physically through. Means for each block with standard errors are shown in Figure 4. In support of our first hypothesis, ratios in the baseline pre-cue block were about 9% greater than measured shoulder width, consistent with the tendency to overestimate the width of an aperture that is needed to be just-passable.

We ran a repeated-measures ANOVA on the ratios comparing the three blocks (pre-cue, cue, post-cue). There was a significant effect of block, $F(2, 60) = 3.74, p < .03, \eta_p^2 = .11$ (see Figure 4). Planned contrasts comparing the cue and post-cue blocks to the pre-cue showed that ratios were not different in the cue (M = 1.07) compared to the pre-cue (M = 1.09) block, $F(1, 30) = .35, p = .55$, but that ratios were lower in the post-cue (M = .99) block compared to the pre-cue block, $F(1, 30) = 6.99, p < .02, \eta_p^2 = .19$. These results suggest that providing the cue reduced estimates for width of the aperture needed to be just passable when the cue was removed in the last block of trials. This finding supports our third hypothesis, which stated that training with the AR cue would transfer to the post-cue block performance, bringing the estimates closer to shoulder width. Regarding our second hypothesis, that the AR cue would make estimates closer to shoulder width when it was present, we did not find a significant difference between estimates in the AR cue block compared to the pre-cue block. Reasons for lack of a significant change from estimates in the pre-cue block to the cue block are discussed further in the General Discussion.

We also examined the cue usefulness ratings. Most participants found the cue to be useful (5 rated the cue as neutral with a score of 3, and 22 rated the cue as useful with a score of 4 or 5), but 4 participants rated the cue as useless (score of 1 or 2). We tested the correlation between the cue usefulness rating and the ratios at the cue and post-cue block. There was a significant negative correlation between usefulness rating and ratio in the cue block, $r(31) = -.41, p < .03$. As usefulness rating increased, ratios decreased, consistent with the overall effect of the cue. There was no significant correlation between the rating and the post-cue block.

7.2 Experiment 2: Stepping Over

One concern in analyzing the data is whether the placement of the AR cue is reliable relative to the detected location of the ground

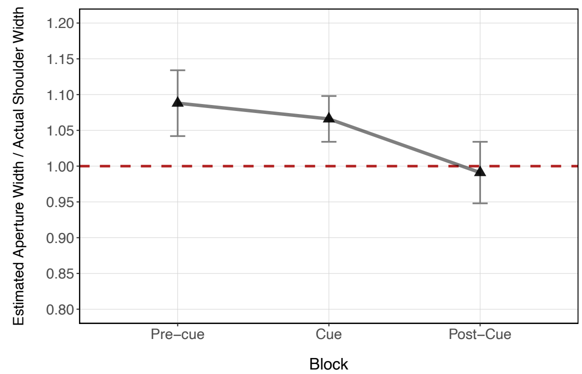

Figure 4: Average passability ratios for each block. Error bars show one standard error above and below the mean ratio.

plane. If the cue is not accurately or reliably placed, then its perceived size may be somewhat distorted either within observers or across them (i.e., closer gaps will look larger compared to gaps of the same size that are displayed farther away). To alleviate potential uncertainty in the placement of the cue relative to the detected ground plane across distinct contexts that were used for testing, the application recorded the position data of the AR gap plane in each trial. We checked the position of the AR gap plane relative to the initial phone camera position to confirm participants had reasonable views during the experiment. If the average altitude of the AR gap plane was disproportionate to the participant's input height, the gap was considered misplaced and data from the participant would be marked invalid. We adopted a criteria for evaluating this disproportion by dividing the AR gap plane altitude for each participant by their reported height. This allowed us to account for people naturally holding their smartphones at different locations around the mid-line of their bodies. Ratios (|Altitude|/Height) that were outside of the range from 0.92 (phone held at approximately eye height) to 0.47 (phone held near the waist height)[Sanchez-Lite et al. 2013] were accepted as valid. We removed one participant's data whose ratio was low (0.38). The height and absolute value of the average altitude of the AR gap plane across 12 trials for participants that were included in further analysis are shown here (see Figure 5). To further explore this issue, we analyzed the mean altitude of the AR gap plane per participant in the pre-cue block versus the post-cue block, after the ground plane had been re-scanned by the user when they changed viewing orientation. A paired sample t-test on these means revealed no significant difference between these altitudes.

As in Experiment 1, we calculated ratios of estimated just-crossable gap to actual measured maximum stride. Actual measured stride ranged from 0.36 m to 1.83 m (M = 0.86 m). Seven participants recorded the default 0.88 m stride, found to be the average stride for adults in prior work [Creem-Regehr et al. 2019]. Supplementary analyses with the 7 participants removed from the dataset are presented in the Appendix. A ratio greater than 1.0 indicates that an observer overestimated the gap that can be crossed and a ratio less than 1.0 indicates that an observer underestimated their ability to cross the gap. We excluded 3 participants from the analysis for

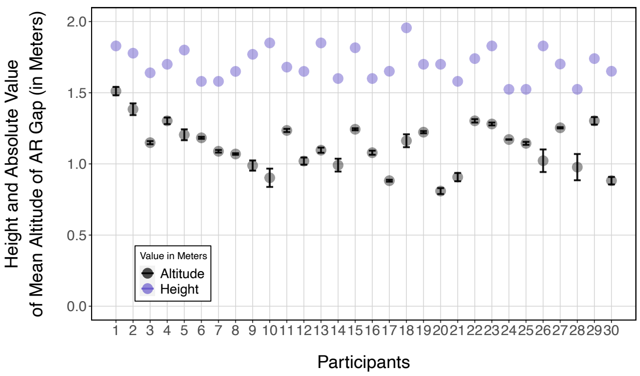

Figure 5: Height and absolute value of the mean altitude across 12 trials for participants included in analysis for Experiment 2. Error bars show one standard error above and below the mean altitude.

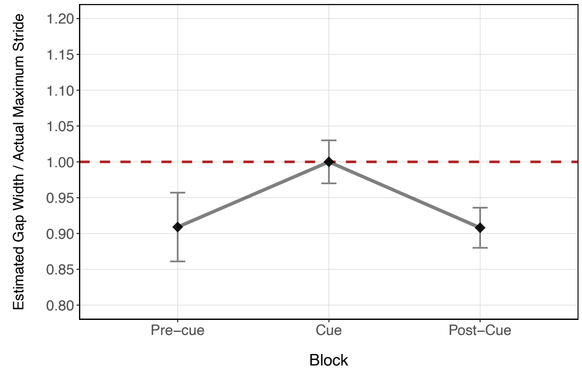

Figure 6: Average gap stepping ratios for each block. Error bars show one standard error above and below the mean ratio.

the following reasons: failure of the app to place the gap on the ground plane (1 participant), a measured stride 3 SD above the mean (1 participant), and estimated/actual ratios falling 3 SD below the mean (2 participants including the 1 failed ground plane). After removing the outliers, the mean of measured stride lengths was M =0.83m (SD =0.19m). Means of the ratios by block, together with standard errors, are shown in Figure 6. For our first hypothesis, the mean ratio in the baseline pre-cue block is about 9% below actual maximum stride, which is inconsistent with prior work in the real world and VR that shows overestimation of judgments for stepping over gaps.

We ran a repeated-measures ANOVA on the ratios comparing the three blocks (pre-cue, cue, post-cue). There was a marginally significant effect of block, $F(2, 54) = 3.12, p = .052, \eta_p^2 = .10$ (see Figure 6). Planned contrasts comparing the cue and post-cue blocks to the pre-cue showed that estimates in the cue block ($M = 1.00$) were marginally greater than in the pre-cue block ($M = .91$), $F(1, 22) = 3.70, p = .065, \eta_p^2 = .12$. This finding weakly supports our second hypothesis, that users estimates would become more accurate when the AR cue was present. However, there was no difference between the post-cue block ($M = .91$) and the pre-cue block, $F(1, 22) = 0.0, p = .99$. The mean ratio values in the pre- and post-cue blocks were essentially the same. These results suggest that while the visible cue increased estimations during the cue block itself, the effect of the cue did not generalize when the cue was no longer present in the post-cue block. Thus, our third hypothesis was not supported here.

Results for the cue usefulness ratings were similar to Experiment 1. Most participants found the cue to be useful: 3 participants rated the cue as useless (score of 1 or 2), 5 rated the cue neutrally (score of 3), and 18 rated the cue as useful (score of 4 or 5). We tested the correlation between the cue usefulness rating and the ratios at the cue and post-cue block. There were no significant correlations between the usefulness rating and the ratios at either block.

8 GENERAL DISCUSSION

Affordance judgments are becoming more widely used to assess perception in VR and AR, but to our knowledge no work has assessed users' perceptions of affordances on mobile smartphones, a pervaise AR platform. This paper had two goals. First, we aimed to determine whether people could reliably and consistently make affordance judgments for actions presented via a mobile smartphone. We therefore tested the perception of two affordances on mobile AR via smartphones: passing through and stepping over. These tasks were chosen because they are feasible to complete given the size of smartphone screens and because there is prior work to benchmark the findings here against in both optical see-through AR [Pointon et al. 2018b] and immersive VR [Geuss et al. 2010; Jun et al. 2015]. In initial judgements prior to seeing an AR cue, users were conservative in their estimates of the capabilities. For passing through, users left a margin of error by adjusting the aperture to be wider than their shoulder width, and for stepping over users adjusted the gap to be smaller than they could actually step over. These results suggest that passability judgments in mobile AR are similar to that observed in other virtual and augmented environments. But, judgments for stepping over were more conservative (underestimation of capabilities) than reported in prior work. The findings provide a foundation for future work assessing affordances via mobile smartphone AR with other tasks.

In addition to providing a baseline set of studies for how people perceive affordances via smartphones, our second goal was to investigate whether perception of action capabilities would change when providing a cue corresponding to body dimensions or capabilities and whether effects of the cue would persist when the cue was removed. With regard to the AR cues, estimates of the width of the aperture needed to pass through became closer to the viewers' actual shoulder widths in the post-cue trials compared to the pre-cue trials, but there was no significant difference when comparing the cue to the pre-cue block. These results supported our third hypothesis (transfer of cue training) but not our second (effect of the cue while viewing). In contrast, the size of the gap estimated to be crossable increased to more closely match the viewer's actual stride when the cue was present, but did not extend to estimates in the post-cue block. These findings supported our second hypothesis, but not our third. Taken together, the findings suggest that cues that provide information about body size or capabilities in AR have the potential to influence a user's estimates of affordances towards virtual objects overlaid in real environments via mobile smartphones. However, the persistence of the training with the cue and its effectiveness when displayed in the context of the affordance clearly depends on the task at hand and the affordance being assessed. Thus, the usefulness of AR cues for training perceived affordances in future work needs more assessment to ensure they are effective for all tasks and that the information provided by the cue successfully transfers to future trials.

Why did the effect of the AR cue transfer to judgments without the cue for only one affordance? For passing through, we found little change during the cue block, but reduced estimations of width after the cue was removed. For stepping over, the cue had an effect only during the cue block. One possible explanation could be the region of space in which the action is performed. Whereas passing through an aperture typically involves viewing and acting on an environmental feature that is farther from the viewer in action space, the gap for the crossing judgments was located in personal space. Different visual depth cues are available in these different regions [Cutting and Vishton 1995]. The underlying perception mechanism of depth and layout might fundamentally be different in the two affordance tasks, which might also affect the use of visual AR cues in the estimations.

Differences in strategies for how the cue was used in different affordance contexts could also be attributed to the differences in the effect of AR cues. For passing through, the cue displayed the shoulder width of participants, and a good strategy was to align the cue between the poles to check if it "fit." Doing so would have led to a needed margin of error to get the cue to fit between the poles. Users adopting this strategy would have resulted in a slight overestimation with the cue present compared to when it was absent (i.e., in the post-cue trials), which is what we observed in Experiment 1. For stepping over, the cue may have changed judgments during the cue-block because the stepping over judgments were harder to see in the cameras of smartphones for some users. For example, if a user had a particularly large actual step size, then the largest virtual gaps generated (which were calculated relative to actual capability) would have been difficult to fully view in the camera "window" without having to scan to see the extent of gap (via tilting the camera). Such a scanning process could have required more working memory resources to integrate the gap extent across views and would have been more cumbersome in terms of manipulation of the smartphone. With the cue present, scanning may have become easier given the cue would not extend out of the viewing window. It is also possible that participants may have relied more on real world cues on the ground near them when judging stepping over the gaps and then continued this strategy in the post-cue block instead of relying on the feedback from the cue block.

In addition, prior work on distance perception with mobile AR has shown that the phone camera leads to a perceptual expansion of space [Liu et al. 2020; Swan et al. 2017], particularly for closer distances. This perceptual distortion could explain the initial underestimation of capability for stepping over (i.e., if the extent looks farther, then the stepping capability judgment would be reduced) and may also have been difficult to ignore in the post-cue block.

9 LIMITATIONS AND FUTURE WORK

We also acknowledge a number of limitations associated with our distributed data collection approach using mobile AR via smartphones. First, we necessarily had less control over measurement of the body dimension (shoulder width) or capability (stepping) compared to a traditional laboratory study where these dimensions are carefully measured by trained experimenters. However, we deliberately tried to control for variability in self-measurements as much as possible by providing photos and instructions for how to complete the measurement. Despite these instructions, some participants appeared to lack the means to measure their dimensions/capabilities and opted for the default measure provided. We show in the additional analyses in the Appendix that while the pattern of results is similar in the sample analyzed with only the participants who used self-measurements, the absolute values are different. Some steps in future studies that require body measurement could be taken to mitigate errors — for example, the participants could submit a photograph of them measuring shoulder width and stride length so they could be validated, i.e., in that way experimenters can verify that a) they measured correctly and b) they entered the measurement correctly. Second, there are differences among the devices, especially the screen sizes of smartphones, that may have affected our results in current experiments. Different screen sizes could influence perception of the same AR stimulus and might interact with the tasks at hand in different spaces. Nonetheless, generalizations of AR training effects across difference devices will be important to assess in future work. Third, we had no control over the real world environments that participants viewed the affordance objects and AR cues in for each task. It is possible that the virtual apertures or gaps would be perceived differently in different contexts, such as outdoors versus indoors, or cluttered versus sparse environments. Future work could record this information by requesting that participants take a photo of their environment, or control this variable by requiring that a specific type of environment be used. But, it is also important to consider the advantages of testing participants in the wild, i.e., across varying environments and with different devices. We traded experimental control in measurements and environment for ecological validity. We think this trade-off was warranted given that AR apps are already in use in numerous contexts and will continue to be widely employed moving forward. Thus, our results are important in that they show that we were able to find reliable effects of providing cues even in the face of this inherent variability and loss of some experimental control.

The applications and experiments presented here provide an important framework for work assessing perception with objective measures (such as perceived affordances) and training that perception with mobile AR smartphones. However, what exactly should be trained and how may depend on the application and task at hand. Given a few participants did not find the AR cue "useful," we should evaluate how to best portray this information to users in future work. Specifically, an AR cue that represents shoulder width might be more helpful if displayed at shoulder height, or labeled differently to convey more meaning. Further, future work could assess whether the AR cue should train the aperture size that would be needed if participants intended to actually *walk through* the virtual poles (rather than statically viewing them). The dynamics of walking includes body sway, which necessitates a wider aperture width in order to successfully fit through [Franchak et al. 2012]. Thus, future work could change the AR cue to facilitate judgments of dynamic actions (i.e., walking through the poles) rather than just training participants on their static body size. Comparisons of affordance judgments in mobile AR in a more controlled lab setting to the real world would be a starting point for further investigation.

10 CONCLUSION

This paper provides a foundation for using applications with smartphones for testing users' perceived action capabilities with virtual objects in mobile AR. We show that data collection with smartphones is feasible and reliable. In two experiments we find that users can perceive their action capabilities, and in the case of passing through, they can change the perception of their abilities with virtual targets given training with an AR cue. The lack of transfer of training with judgments of stepping over may have been due to factors related to the cameras in the phones, their field of view, or the means by which scanning is employed to make the judgment. Future work will assess when and how AR cues may improve perceptions of action capabilities with virtual targets across different tasks, but the findings here suggest the use of mobile AR with smartphones for assessing and training perception is possible.

ACKNOWLEDGMENTS

The authors thank the reviewers for their constructive comments. This material is based in part upon work supported by the Office of Naval Research under Grant No. N00014-18-1-2964 and by the National Science Foundation under Grant 1763996.

REFERENCES

Nur Intan Adhani and Dayang Rohaya Awang Rambli. 2012. A survey of mobile augmented reality applications. In *1st International conference on future trends in computing and communication technologies*. 89–96.

Matthias Berning, Daniel Kleinert, Till Riedel, and Michael Beigl. 2014. A Study of Depth Perception in Hand-Held Augmented Reality using Autostereoscopic Displays.

Ayush Bhargava, Kathryn M Lucaites, Leah S Hartman, Hannah Solini, Jeffrey W Bertrand, Andrew C Robb, Christopher C Pagano, and Sabarish V Babu. 2020a. Revisiting affordance perception in contemporary virtual reality. *Virtual Reality* (2020), 1–12.

Ayush Bhargava, Hannah Solini, Kathryn Lucaites, Jeffrey W Bertrand, Andrew Robb, Christopher C Pagano, and Sabarish V Babu. 2020b. Comparative Evaluation of Viewing and Self-Representation on Passability Affordances to a Realistic Sliding Doorway in Real and Immersive Virtual Environments. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 519–528.

Soumyajit Chakraborty, Jeanine K. Stefanucci, Sarah Creem-Regehr, and Bobby Bodenheimer. 2021. Distance Estimation with Mobile Augmented Reality in Action Space: Effects of Animated Cues. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 144–147. https://doi.org/10.1109/VRW52623.2021.00034

D. Chatzopoulos, C. Bermejo, Z. Huang, and P. Hui. 2017. Mobile Augmented Reality Survey: From Where We Are to Where We Go. *IEEE Access* 5 (2017), 6917–6950.

Sarah H Creem-Regehr, Devin M Gill, Grant D Pointon, Bobby Bodenheimer, and Jeanine Stefanucci. 2019. Mind the gap: Gap affordance judgments of children, teens, and adults in an immersive virtual environment. *Frontiers in Robotics and AI* 6 (2019), 96.

James E. Cutting and Peter M. Vishton. 1995. Perceiving Layout and Knowing Distance: The Integration, Relative Potency and Contextual Use of Different Information about Depth. In *Perception of Space and Motion*, William Epstein and Sheena Rogers (Eds.). Academic Press, New York, 69–117.

Arindam Dey, Andrew Cunningham, and Christian Sandor. 2010. Evaluating depth perception of photorealistic mixed reality visualizations for occluded objects in outdoor environments. In *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology*. ACM, 211–218.

Arindam Dey, Graeme Jarvis, Christian Sandor, and Gerhard Reitmayr. 2012. Tablet versus phone: Depth perception in handheld augmented reality. In *2012 IEEE international symposium on mixed and augmented reality (ISMAR)*. IEEE, 187–196.

Pei-Huang Diao and Nj Shih. 2018. MARINS: A mobile smartphone AR system for pathfinding in a dark environment. *Sensors* 18 (10 2018), 3442. https://doi.org/10.3390/s18103442

Tiffany Do, Jr LaViola, and Ryan McMahan. 2020. The Effects of Object Shape, Fidelity, Color, and Luminance on Depth Perception in Handheld Mobile Augmented Reality. *ArXiv preprint* (08 2020).

Elham Ebrahimi, Andrew Robb, Leah S Hartman, Christopher C Pagano, and Sabarish V Babu. 2018. Effects of anthropomorphic fidelity of self-avatars on reach boundary estimation in immersive virtual environments. In *Proceedings of the 15th ACM Symposium on Applied Perception*. 1–8.

John M Franchak and Karen E Adolph. 2012. What infants know and what they do: Perceiving possibilities for walking through openings. *Developmental psychology* 48, 5 (2012), 1254–1261.

J. M. Franchak and K. E. Adolph. 2014. Affordances as probabilistic functions: Implications for development, perception, and decisions for action. *Ecological Psychology* 26 (2014), 109–124.

John M Franchak, Emma C Celano, and Karen E Adolph. 2012. Perception of passage through openings depends on the size of the body in motion. *Experimental Brain Research* 223, 2 (2012), 301–310.

Holly C. Gagnon, Dun Na, Keith Heiner, Jeanine Stefanucci, Sara Creem-Regehr, and Bobby Bodenheimer. 2020. The Role of Viewing Distance and Feedback on Affordance Judgments in Augmented Reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 922–929.

Holly C. Gagnon, Yu Zhao, M. Richardson, Grant Pointon, Jeanine K Stefanucci, S. Creem-Regehr, and B. Bodenheimer. 2021. Gap Affordance Judgments in Mixed Reality: Testing the Role of Display Weight and Field of View. In *Frontiers in Virtual Reality*.

Michael Geuss, Jeanine Stefanucci, Sarah Creem-Regehr, and William B. Thompson. 2010. Can I Pass?: Using Affordances to Measure Perceived Size in Virtual Environments. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization* (Los Angeles, California) *(APGV '10)*. ACM, New York, NY, USA, 61–64. https://doi.org/10.1145/1836248.1836259

Michael N. Geuss, Michael J. McCardell, and Jeanine K. Stefanucci. 2016. Fear similarly alters perceptual estimates of and actions over gaps. *PLoS one* 11, 7 (2016), e0158610.

James J. Gibson. 1979. *The ecological approach to visual perception*. Houghton Mifflin, Boston, MA.

O. Hasler, S. Blaser, and S. Nebiker. 2019. Implementation and First Evaluation of an Indoor Mapping Application Using Smartphones and AR Frameworks. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLII-2/W17 (2019), 135–141. https://doi.org/10.5194/isprs-archives-XLII-2-W17-135-2019

Takahiro Higuchi. 2013. Visuomotor control of human adaptive locomotion: understanding the anticipatory nature. *Frontiers in Psychology* 4, 277 (2013).

Lorenzo Jamone, Emre Ugur, Angelo Cangelosi, Luciano Fadiga, Alexandre Bernardino, Justus Piater, and José Santos-Victor. 2018. Affordances in Psychology, Neuroscience, and Robotics: A Survey. *IEEE Transactions on Cognitive and Developmental Systems* 10, 1 (2018), 4–25. https://doi.org/10.1109/TCDS.2016.2594134

Yang Jiang and Leonard S Mark. 1994. The effect of gap depth on the perception of whether a gap is crossable. *Perception & Psychophysics* 56, 6 (1994), 691–700.

Eunice Jun, Jeanine K. Stefanucci, Sarah H. Creem-Regehr, Michael N. Geuss, and William B. Thompson. 2015. Big Foot: Using the Size of a Virtual Foot to Scale Gap Width. *ACM Trans. Appl. Percept.* 12, 4 (2015), 16:1–16:12.

Ernst Kruijff, J. Edward Swan, and Steven Feiner. 2010. Perceptual issues in augmented reality revisited. In *2010 IEEE International Symposium on Mixed and Augmented Reality*. 3–12. https://doi.org/10.1109/ISMAR.2010.5643530

Gunnar Liestøl. 2009. Situated simulations: A prototyped augmented reality genre for learning on the iPhone. *International Journal of Interactive Mobile Technologies (iJIM)* 3 (2009), 24–28.

Jingjing (May) Liu, Gayathri Narasimham, Jeanine Stefanucci, Sarah Creem-Regehr, and Bobby Bodenheimer. 2020. Distance Perception in Modern Mobile Augmented Reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 196–200.

Mark A Livingston, Zhuming Ai, J Edward Swan, and Harvey S Smallman. 2009. Indoor vs. outdoor depth perception for mobile augmented reality. In *2009 IEEE Virtual Reality Conference*. IEEE, 55–62.

Leonard S Mark, James A Balliett, Kent D Craver, Stephen D Douglas, and Teresa Fox. 1990. What an actor must do in order to perceive the affordance for sitting. *Ecological Psychology* 2, 4 (1990), 325–366.

Donald A Norman. 2002. *The design of everyday things*. Basic books.

Paweł Nowacki and Marek Woda. 2020. Capabilities of ARCore and ARKit Platforms for AR/VR Applications. In *Engineering in Dependability of Computer Systems* *and Networks*, Wojciech Zamojski, Jacek Mazurkiewicz, Jarosław Sugier, Tomasz Walkowiak, and Janusz Kacprzyk (Eds.). Springer International Publishing, Cham, 358–370.

Jodie M Plumert. 1995. Relations between children's overestimation of their physical abilities and accident proneness. *Developmental Psychology* 31, 5 (1995), 866.

Jodie M Plumert and David C Schwebel. 1997. Social and temperamental influences on children's overestimation of their physical abilities: Links to accidental injuries. *Journal of experimental child psychology* 67, 3 (1997), 317–337.

Grant Pointon, Chelsey Thompson, Sarah Creem-Regehr, Jeanine Stefanucci, and Bobby Bodenheimer. 2018a. Affordances as a Measure of Perceptual Fidelity in Augmented Reality. In *2018 IEEE VR 2018 Workshop on Perceptual and Cognitive Issues in AR (PERCAR)*. 1–6.

Grant Pointon, Chelsey Thompson, Sarah Creem-Regehr, Jeanine Stefanucci, Miti Joshi, Richard Paris, and Bobby Bodenheimer. 2018b. Judging action capabilities in augmented reality. In *Proceedings of the 15th ACM Symposium on Applied Perception*. 1–8.

Horacio Rios, Eduardo González, Ciro Rodriguez, Hector R. Siller, and Manuel Contero. 2013. A Mobile Solution to Enhance Training and Execution of Troubleshooting Techniques of the Engine Air Bleed System on Boeing 737. *Procedia Computer Science* 25 (2013), 161 – 170. https://doi.org/10.1016/j.procs.2013.11.020 2013 International Conference on Virtual and Augmented Reality in Education.

Alberto Sanchez-Lite, Manuel Garcia, Rosario Domingo, and Miguel Angel Sebastian. 2013. Novel ergonomic postural assessment method (NERPA) using product-process computer aided engineering for ergonomic workplace design. *PloS one* 8, 8 (2013), e72703.

Yushan Siriwardhana, Pawani Porambage, Madhusanka Liyanage, and Mika Ylianttila. 2021. A Survey on Mobile Augmented Reality With 5G Mobile Edge Computing: Architectures, Applications, and Technical Aspects. *IEEE Communications Surveys Tutorials* 23, 2 (2021), 1160–1192. https://doi.org/10.1109/COMST.2021.3061981

Jeanine K Stefanucci, Sarah H Creem-Regehr, William B Thompson, David A Lessard, and Michael N Geuss. 2015. Evaluating the accuracy of size perception on screen-based displays: Displayed objects appear smaller than real objects. *Journal of Experimental Psychology: Applied* 21, 3 (2015), 215–223.

Jeanine K. Stefanucci and Michael N. Geuss. 2009. Big people, little world: The body influences size perception. *Perception* 38, 12 (2009), 1782–1795.

Thomas A Stoffregen, Chih-Mei Yang, M Russell Giveans, Moira Flanagan, and Benoît G Bardy. 2009. Movement in the perception of an affordance for wheelchair locomotion. *Ecological Psychology* 21, 1 (2009), 1–36.

J Edward Swan, Liisa Kuparinen, Scott Rapson, and Christian Sandor. 2017. Visually Perceived Distance Judgments: Tablet-Based Augmented Reality Versus the Real World. *International Journal of Human–Computer Interaction* 33, 7 (2017), 576–591.

W. H. Warren and S. Whang. 1987. Visual Guidance of Walking Through Apertures: Body Scaled Information for Affordances. *Journal of Experimental Psychology: Human Perception and Performance* 13 (1987), 371–383.

Hansen Wu, Haley Adams, Grant Pointon, Jeanine Stefanucci, Sarah Creem-Regehr, and Bobby Bodenheimer. 2019. Danger from the deep: A gap affordance study in augmented reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 1775–1779.

Yawen Yu, Benoît G Bardy, and Thomas A Stoffregen. 2010. Influences of head and torso movement before and during affordance perception. *Journal of Motor Behavior* 43, 1 (2010), 45–54.

A SUPPLEMENTARY ANALYSES

Given that seven participants did not measure their shoulder width or actual stride in each of the apps, we conducted the same ANOVAs as in the main paper but with only the sample of participants who had measured dimensions/capabilities. This allowed us to further explore the impact of including the data that did not have direct body measurements. For Experiment 1, passing through judgments, this resulted in a sample of 24 participants (excluding 1 outlier as in the primary analysis). We ran a repeated-measures ANOVA on the mean ratios (estimated just-passable aperture/actual measured shoulder width) comparing the three blocks (pre-cue, cue, post-cue). There was a significant effect of block, $F(2,46) = 3.91, p < .03, \eta_p^2 = .15$. Planned contrasts comparing the cue and post-cue blocks to the pre-cue block showed that ratios were not different in the cue ($M = 1.05$) compared to the pre-cue ($M = 1.05$) block, $F(1,23) = .004, p = .95$, but that ratios were lower in the post-cue ($M = .96$) block compared to the pre-cue block, $F(1,23) = 6.27, p < .02, \eta_p^2 = .21$. These effects are consistent with those presented in the primary analysis.

For Experiment 2, stepping over judgments, we analyzed a sample of 21 participants who had measured their maximum stride (after removing the same three outliers as in the primary analysis). We ran a repeated-measures ANOVA on the mean ratios (estimated just-crossable gap/actual measured maximum stride) comparing the three blocks (pre-cue, cue, post-cue). There was a significant effect of block, $F(2,40) = 3.66, p < .04, \eta_p^2 = .16$. Planned contrasts comparing the cue and post-cue blocks to the pre-cue block showed that ratios were not different in the cue ($M = 1.02$) compared to the pre-cue ($M = .95$) block, $F(1,20) = 1.51, p = .23$, and there was no difference between the post-cue block ($M = .89$) block and the pre-cue block, $F(1,20) = 1.36, p < .26$. A follow-up contrast confirmed that the effect of block was driven by a significantly lower estimate in the post-cue block compared to the cue block, $F(1,20) = 16.23, p < .01, \eta_p^2 = .45$. The pattern of effects is generally consistent with the the results presented in the primary analysis, although as noted, the planned comparison between the cue and pre-cue block did not reach significance (it was marginally significant in the primary analysis at p = .052).

We ran one other additional analysis for Experiment 2, to address the known variability in measuring maximum stride (see Creem-Regehr et al. [2019]) and the smaller sample size that resulted from excluding seven participants, using ratios of estimated just-crossable gap to reported body height. Although this ratio does not give us a measure of accuracy relative to actual step, it allows for scaling to a body dimension that does not rely on a performed capability and is correlated to stride length. Prior work has used eye height in an analogous way [Geuss et al. 2016; Jiang and Mark 1994]. We ran a repeated-measures ANOVA on the mean ratios (estimated just-crossable gap/actual height) comparing the three blocks (pre-cue, cue, post-cue) including all 31 participants. There was a significant effect of block, $F(2,60) = 3.39, p < .04, \eta_p^2 = .10$. Planned contrasts comparing the cue and post-cue blocks to the pre-cue block showed that ratios increased in the cue ($M = .47$) compared to the pre-cue ($M = .42$) block, $F(1,30) = 8.47, p < .01 \eta_p^2 = .22$, and there was no difference between the post-cue block ($M = .43$) block and the pre-cue block, $F(1,30) = .16, p = .69$. This analysis supports the same pattern of results as in the primary analysis, i.e., an increase in estimates in the cue-block, but no transfer to the post-cue.