# Evaluating the Accuracy of Size Perception in Real and Virtual Environments

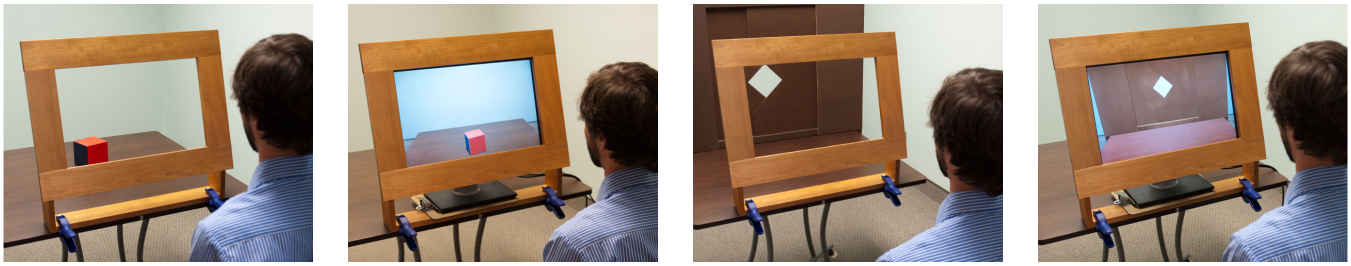
Jeanine K. Stefanucci*
Department of Psychology
University of Utah

David A. Lessard
Department of Psychology
University of Utah

Michael N. Geuss
Department of Psychology
University of Utah

Sarah H. Creem-Regehr
Department of Psychology
University of Utah

William B. Thompson
School of Computing
University of Utah



(a) (b) (c) (d)

**Figure 1:** *Real (a) and virtual (b) grasping tasks. Real (c) and virtual (b) pass-through tasks. Note that the real and virtual views do not look the same since the photographs are not taken from the participants eye point.*

## Abstract

Accurate perception of the size of 3D objects depicted on 2D desktop displays is important for many applications. Whether users perceive objects depicted on a display to be the same size as comparable real world objects is not well understood. We propose using affordances judgments as a way of measuring the perceived size of objects depicted in desktop virtual environments and the real world. The methodology involves indicating whether or not a particular action can be performed in a given environment, making it a flexible measure that can be used across different display technologies. In two studies, we test users' perceptions of size by asking them to make affordance judgments in both the real world and a geometrically matched desktop virtual environment. In the first study, users judge whether they can grasp an object and in the second study, they judge whether they can fit their hand through an opening. In both experiments we show that users perceive the size of objects in the desktop virtual environment to be smaller than in the real world.

**CR Categories:** I.3.3 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual reality H.5.1 [Information Systems]: Multimedia Information Systems—Artificial, augmented, and virtual realities

**Keywords:** virtual environments, perception and action, size perception, human perception and performance, affordances

*e-mail:jeanine.stefanucci@psych.utah.edu

## 1 Introduction

Accurate perception of the scale of spaces conveyed using computer graphics is important in applications ranging from vehicle simulators to visualization. Substantial research has been directed at evaluating the accuracy of distance perception in visually immersive environments (e.g., [Loomis and Knapp 2003; Thompson et al. 2004; Richardson and Waller 2007]). The limited research that has addressed size perception in such environments suggests that size is perceived reasonably accurately [Kenyon et al. 2007; Geuss et al. 2010; Geuss et al. in press]. Almost nothing is known about how well the scale of a depicted reachable space is seen when presented on a desktop monitor. When 3D images are displayed on a screen, there is a dual awareness that is evoked in observers, involving both the screen they are viewing and the three-dimensional scene depicted on the screen [Koenderink et al. 1994]. Understanding how users perceive the size of objects in the scene is complicated by this because users can perceive the physical size of the object on the screen or the perceived representational size of the object as it exists in the real world. In this paper, we present two contributions to the study of perception of scale in desktop virtual environments (DVEs). First, we describe an approach that is able to provide a quantitative evaluation of how well users perceive the representational size of objects presented on a screen. Second, we show two examples of using this technique to evaluate the size of displayed objects and compare these results to matched real-world conditions.

Affordances (as proposed by Gibson [1979]) are action possibilities in an environment, which are directly related to an observer's capabilities. Affordance judgments are decisions about one's capability to act in the environment, without requiring the actual action to be performed. For example, an object is seen as affording grasping behavior if the object's width is less than the observers' hand size. Many studies conducted in the real world show that observers reliably perceive whether an environment affords certain actions such as sitting, reaching through, or walking through [Ishak et al. 2008; Mark 1987; Warren and Whang 1987]. Recently, researchers have begun to test whether affordances are also reliably perceived in immersive virtual environments (IVEs). Loomis and Knapp [2003]


UU IR Author Manuscript

UU IR Author Manuscript

assessed observers' perceptions of the size of a virtual aperture in an IVE by asking them to adjust its extent until they believed they could just walk through the aperture without turning their bodies. They found that participants adjusted the size of the aperture to be slightly larger than their shoulder width. However, they did not directly compare real world adjustments and IVE adjustments in their experiment. Geuss et al. [2010] asked participants to judge whether a virtual aperture afforded passage in both a real and immersive virtual environment. They found no differences between judgments made in the real world and the visually matched virtual environment, suggesting that affordance judgments are valid estimates of the size of virtual objects in IVEs.

In two experiments, we tested whether users accurately perceived the size of objects in a DVE by assessing their judgments about possible interactions with objects conveyed in the display. Such affordance judgments may be especially useful as a perceptual measure of size across display type because 1) they are more task-relevant in that they require the user to consider acting on the object, 2) they do not require direct action towards the object which would be constrained by the presence of the screen, and 3) they are well calibrated as users make these decisions on a daily basis. This study highlights the usefulness of affordance judgments as a measure of perceived space in VEs. In addition, the current experiments extend prior work on affordance judgments in large scale virtual environments to perception of object size in spaces that are close to within reach. The current experiments are the first to assess whether these judgments are made on DVEs similarly to how they are made in the real world. In the first experiment, we asked participants to judge whether or not they could grasp and lift cubes of different sizes off a table with one hand. In the second experiment, participants judged whether or not their hand was able to fit through an opening, based on the methodology of [Ishak et al. 2008]. For both experiments, judgments are made in a real environment and in a visually matched DVE. How observers perceive pictorial, 2D displays with respect to body-based judgments about action capabilities has not been examined, and may be relevant for applications intended to display act-on-able objects or environments such as in architectural design.

## 2 Using Affordances to Estimate the Size of Cubes

Participants judged whether they could grasp cubes of varying dimensions in either a real environment or a geometrically matched DVE. All participants viewed both the real and virtual environments. In each environment, participants made 36 decisions about whether they could grasp cubes of various sizes with their hand and lift them off a table. Ten University of Utah students (4 male, 6 female) participated for class credit or $10. All participants had normal or corrected to normal vision and were right-handed.

### 2.1 Design and Procedure.

Participants judged their ability to grasp nine multicolored (red, green, and blue) cubes varying in size from 6 cm to 22 cm on a side at 2 cm intervals. Virtual cubes were modeled with the textures from the real cubes. Cubes were viewed either through a frame on a real table (real world condition) or through a frame and desktop monitor (DVE condition; see Figure 1). The table was 91.5 cm square. The frame was 66 cm by 47.5 cm on the outside and 48.5 cm by 30 cm (22" diagonal) on the inside and centered at the front edge of the table. The monitor was a Dell brand 56 cm (22") size diagonal and, when present, was centered inside the frame. Each participant sat in a chair that was adjusted so that their eyes were 30 cm vertically and 60 cm horizontally from the front of the frame. A guide was suspended from the ceiling so that participants always kept their viewpoint in the same position. This was to ensure that the visual geometry displayed in the real world and DVE was identical.

A within-subjects design was used for display condition (real-world and desktop display) and for distances to the cube (50 cm and 70 cm from front edge of table). Participants saw each cube twice at each location (36 trials) in each display condition, for 72 total trials. Two locations were used to test whether differences in affordance estimates were present or not across a range of distances. All cubes were oriented at 15° from straight on so that three sides were clearly visible. The order of condition was randomized and counterbalanced across participants. The order of presentation of trials (cube size and distance) within each display condition block was randomized. For this experiment and the following, there were no significant main effects of order of presentation of condition, so order was excluded from further analyses and discussion.

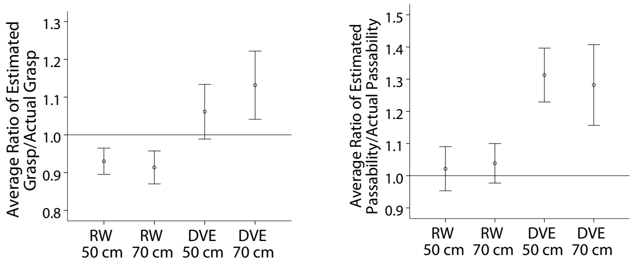
Participants were told that they would see cubes of varying sizes and that they should judge whether or not they could grasp each cube. Grasping behavior was defined as the use of the forefinger and thumb to pick up the cube across the face of the cube. Participants were instructed to respond with a "yes" when they thought the cube was small enough to perform this action and to respond with a "no" when they thought the cube was too large to perform this action. The experimenter emphasized that it was more important to think about grasping the cube than whether or not the cube was reachable, since some cubes would be out of reach. Participants' positions were monitored throughout the experiment in order to maintain consistent viewing conditions. Participants kept both hands in their lap at all times so that they did not attempt to grasp the cubes at any point and could not use the visual size of their hand as a reference. In the real world condition, participants closed their eyes between judgments so that they did not see the experimenter move the cubes. In the desktop display condition, the display turned black for 5 s (to match the inter-trial interval in the real world condition). After completing all trials in one condition, participants judged their ability to grasp cubes in the other condition. After both conditions were completed, the experimenter recorded the participant's hand size by measuring from the tip of the thumb to the tip of the forefinger and their actual largest grasp by recording the largest cube that the participant could actually pick up.

### 2.2 Results

Cross-over points were calculated as the average of the largest cube judged graspable and the smallest cube judged as not graspable across the two trials for each cube size. Ratios were then created by dividing the cross-over point for each participant by that participant's actual grasp size for each distance (50 cm, 70 cm) and condition (real world, DVE). A ratio greater than 1 means participants estimated that they could grasp cubes that were larger than their actual grasp; a ratio smaller than 1 means participants estimated that the largest cube they could grasp was smaller than their actual grasp.

A 2 distance (50 cm, 70 cm) x 2 condition (real world, DVE) repeated measures ANOVA with the ratios as the dependent variables revealed a main effect of environment such that in the DVE condition, participants judged that they could grasp larger cubes (M = 1.096, SE = 0.058) than in the real world (M = 0.922, SE = 0.038), F(1,9) = 5.48, p = 0.044, MSE = 0.056, $\eta_p^2$ = 0.379. There was no main effect of distance, F(1,9) = 1.774, p= 0 .216, MSE = 0.004, $\eta_p^2$ = 0.165. In addition, there was an interaction of environment and location such that the effect of distance was greater in the desktop condition (50 cm M = 1.061, SE = 0.072; 70 cm M = 1.132, SE = 0.090) than in the real world (50 cm M = 0.930, SE = 0.035; 70 cm

UU IR Author Manuscript




(a) (b)

**Figure 2:** *(a) Average of the ratios for estimated grasp/actual grasp for all displayed cubes in the real world (RW) and desktop virtual environment (DVE) at 50 and 70 cm distances. (b) Average of the ratios for estimated passability/actual passability for all of the apertures in the RW and DVE at 50 and 70 cm distances. For both graphs, bars represent +- 1 standard error.*

M = 0.914, SE = 0.044, F(1,9) = 5.908, p = 0.038, MSE = 0.003, $\eta_p^2$ = 0.396. (See Figure 2(a)).

## 3 Using Affordances to Estimate the Ability to Reach Through

Experiment 1 showed that objects were perceived to be smaller in the DVE when compared to the real world. This underestimation of size led to an overestimation of grasping ability in the DVE compared to the real world. In Experiment 2, we tested whether we would replicate the pattern of results in Experiment 1 by asking participants to make a different affordance judgment that was related to perceived size. Participants judged whether their hand could fit through an aperture in both a real environment and a geometrically matched DVE. Given the results of Experiment 1, we hypothesized that users would judge the smallest aperture they could pass their hand through in the DVE condition to be larger than in the real world. Ten University of Utah students (6 male, 4 female) participated for class credit or $10. All participants had normal or corrected to normal vision and were right-handed.

### 3.1 Design and Procedure.

All stimuli and apparati were the same as in Experiment 1 except that participants viewed apertures instead of cubes. The aperture was a hole in a wooden frame with sides that could be adjusted to change its size (see Figure 1). The frame measured 99 cm by 67 cm. The size of the aperture varied in size from 2 cm square to 30 cm at 2 cm intervals (15 sizes). The aperture was constructed based on the one used in [Ishak et al. 2008].

A within-subjects design was used for display condition (real world and DVE) and for distances to the target (50 cm and 70 cm from the front of the table). Participants saw each aperture size twice at each location (60 trials) in each environment for 120 trials total. The order of presentation of environment was randomized and counterbalanced. The order of presentation of aperture size and distance within each block of trials for environment was also randomized.

All procedures were the same as in Experiment 1 except that participants judged whether they thought their hand could fit through the aperture. They were told to think about putting their dominant hand through the aperture by holding it flat with the thumb beside the palm (without scrunching up or folding their hand) and that their hand could not touch sides of the aperture. Participants were instructed to respond with a "yes" when the aperture was large enough for them to pass their hand through the aperture without touching the walls and to respond with a "no" answer when they thought that the aperture was too small to pass through. The experimenter emphasized that it was more important to think about whether or not their hand was small enough to pass through the hole rather than whether or not they could reach the hole since some apertures would be out of reach. At the end of the experiment, participants' actual smallest passable aperture was measured by having the participants insert their dominant hand into increasingly smaller apertures until they were no longer able to pass their hand through.

### 3.2 Results

Cross-over points were calculated as the average of the smallest aperture participants judged passable and the largest aperture they judged as not passable. A ratio greater than 1 means the participants estimated that the smallest passable aperture was larger than their actual smallest passable aperture; a ratio smaller than 1 means the participants estimated that the smallest estimated passable aperture was smaller than their actual smallest passable aperture. Analyses were run as the ratio of cross-over point divided by actual smallest passable aperture. A 2 distance (50 cm, 70 cm) x 2 condition (real world, DVE) repeated measures ANOVA with the ratios as the dependent variables revealed a main effect of environment such that in the DVE condition, participants judged the smallest aperture they could pass through (M = 1.30, SE = 0.103) to be larger than the smallest aperture they could pass through in the real world (M = 1.03, SE = 0.064), F(1, 9) = 6.33, p = .033, MSE = 0.113, $\eta_p^2$ = 0.413. (See Figure 2(b)). There was no main effect of distance, F(1,9) = 0.07, p = 0.797, MSE = 0.007, $\eta_p^2$ = 0.008, nor an interaction of environment and location, F(1, 9) = 0.512, p = 0.493, MSE = 0.011, $\eta_p^2$ = 0.054.

## 4 General Discussion

The results of these experiments suggest that users are less accurate at perceiving the representational size of objects on a desktop display than they are at perceiving the size of objects in the real world. In Experiment 1, users judged they could grasp larger objects on a desktop display and in Experiment 2 they judged that they needed a larger aperture in order to fit their hand through. These results point to the same effect, that the sizes of objects displayed on a desktop monitor are perceived as smaller, informing estimates of action capabilities of grasping and reaching-through that are less conservative than the same judgments made in the real world.

One likely account for the underestimation of size revealed by the affordance judgments in our two experiments is an underestimation of the distance to the displayed objects on the monitor. The size-distance invariance hypothesis [Epstein 1973] claims that perception of the size of an object is affected by the relationship between its its retinal size and the perceived distance to that object. Though this connection between distance perception and size perception has been debated [Epstein 1977], it is possible that the underestimation of representational size observed in these studies was due to an underestimation of the distance to the objects. Perceiving the distance to objects depicted on nearby flat surfaces is confounded by conflicts between the pictorial cues of the image, particularly perspective and familiar size, and the depth cues associated with the display surface itself, such as stereo, motion parallax, and the context of the surrounding space [Rogers 1995; Thompson et al. 2011]. Often the result is that the location of depicted objects appears to be closer to the screen than would be indicated by the pictorial cues alone. If the size-distance invariance hypothesis holds in our situation, underestimation of distance will result in underestimation of size. In addition, desktop displays present a non-immersive environment in which users are aware of viewing images on a screen. There is some evidence that this may introduce bias in perceptual judgments of spatial relationships [Yang et al. 1999].

Future work could address these issues in several ways. First, a stereo-tracked, screen-based display could be used to evaluate whether the addition of stereo as a cue for depth moderates the bias to perceive objects as smaller on the monitor, thus improving the accuracy of the judgments. Second, our design featured the use of a frame in both the desktop condition and the real world environment, which when eliminated may increase the confound associated with perception of the screen surface, thus further increasing the underestimation of size. Monocular viewing of the screen could also affect underestimation of size. A mismatch between the support surface in the virtual world and the actual table might likewise affect the results. Finally, the same size judgment tasks could be explored in an immersive virtual environment, such as that depicted in a head-mounted display (HMD). Such a test would be appealing because it also would control for the dual awareness that arises from viewing objects on a screen. Thus, immersive displays could test for whether the addition of stereo cues changes the perception of representational size in graphical images, while also reducing the possibility that the dual awareness of the display and the depicted 3D scene may contribute to the observed effects.

It is important to note that the current studies were run using one virtual environment, with two different affordance judgments made in that environment (grasping and reaching through). Both judgments required estimating actions related to the hand and its size. Future work should test whether other actions that can be performed in the space near the body, such as reaching, are estimated differently in desktop virtual environments as compared to the real world. Further, though we have asked participants to estimate locomotor affordances, such as passing-through an aperture, in immersive environments that depict large scale spaces in previous work [Geuss et al. 2010], it would be interesting to test whether estimates of those types of actions are possible on a smaller display like a desktop monitor. We believe it is essential to ask all participants to make affordance judgments in analogous real world environments in order to be able to interpret the direction of any effects observed in representational size when viewing a desktop virtual environment.

## 5 Summary

One goal of graphical interfaces is to portray 3D environments on 2D displays, such as computer monitors. An important question to ask of these displays is how well they portray the real physical size of objects to users. Affordance judgments measure perceived size by assessing what users believe they can or cannot do with objects in virtual environments. Thus, they can be used as an index of how well graphical displays portray the physical size of 3D objects. They are also advantageous because they can be used across a variety of displays. We tested the perceived size of objects displayed in a desktop virtual environment with two different affordance estimates, grasping and reaching through. For both judgments, we found that users underestimated the size of the objects when judging their action capabilities for the desktop virtual environment in comparison to a real environment.

## Acknowledgments

This material is based on work supported by the National Science Foundation under Grants No. 0914488 and 1116636. Ian King produced the virtual environments used in this work and Garrett Allen provided valuable technical support for this project.

## References

EPSTEIN, W. 1973. The process of "taking-into-account" in visual perception. *Perception 2*, 3, 267–285.

EPSTEIN, W. 1977. Stability and constancy in visual perception: Mechanisms and processes.

GEUSS, M., STEFANUCCI, J., CREEM-REGEHR, S., AND THOMPSON, W. B. 2010. Can I pass?: Using affordances to measure perceived size in virtual environments. In *Proc. Symposium on Applied Perception in Graphics and Visualization*, 61–64.

GEUSS, M. N., STEFANUCCI, J. K., CREEM-REGEHR, S. H., AND THOMPSON, W. B. in press. Effect of viewing plane on perceived distances in real and virtual environments. *Journal of Experimental Psychology: Human Perception and Performance*.

GIBSON, J. J. 1979. *The Ecological Approach to Visual Perception*. Houghton Mifflin, Boston.

ISHAK, S., ADOLPH, K. E., AND LIN, G. C. 2008. Perceiving affordances for fitting through apertures. *Journal of Experimental Psychology: Human Perception and Performance 34*, 6, 1501–1514.

KENYON, R. V., SANDIN, D., SMITH, R. C., PAWLICKI, R., AND DEFANTI, T. 2007. Size-constancy in the CAVE. *Presence: Teleoperations and Virtual Environments 16*, 2, 172–187.

KOENDERINK, J. J., VAN DOORN, A. J., AND KAPPERS, A. M. L. 1994. On so-called paradoxical monocular stereoscopy. *Perception 23*, 5, 583–594.

LOOMIS, J. M., AND KNAPP, J. 2003. Visual perception of egocentric distance in real and virtual environments. In *Virtual and Adaptive Environments*, L. J. Hettinger and M. W. Haas, Eds. Erlbaum, Mahwah, NJ, ch. 2, 21–46.

MARK, L. S. 1987. Eyeheight-scaled information about affordances: A study of sitting and stair climbing. *Journal of Experimental Psychology: Human Perception and Performance 13*, 3, 361–370.

RICHARDSON, A. R., AND WALLER, D. 2007. Interaction with an immersive virtual environment corrects users' distance estimates. *Human Factors 49*, 3, 507–517.

ROGERS, S. 1995. Perceiving pictorial space. In *Perception of Space and Motion*, W. Epstein and S. Rogers, Eds., vol. 5 of *Handbook of Perception and Cognition*. Academic Press, ch. 4, 119–163.

THOMPSON, W. B., WILLEMSEN, P., GOOCH, A. A., CREEM-REGEHR, S. H., LOOMIS, J. M., AND BEALL, A. C. 2004. Does the quality of the computer graphics matter when judging distances in visually immersive environments? *Presence: Teleoperators and Virtual Environments 13*, 5, 560–571.

THOMPSON, W. B., FLEMING, R. W., CREEM-REGEHR, S. H., AND STEFANUCCI, J. K. 2011. *Visual Perception from a Computer Graphics Perspective*. CRC Press.

WARREN, JR., W. H., AND WHANG, S. 1987. Visual guidance of walking through apertures: Body scaled information for affordances. *Journal of Experimental Psychology: Human Perception and Performance 13*, 371–383.

YANG, T. L., DIXON, M. W., AND PROFFITT, D. R. 1999. Seeing big things: Overestimation of heights is greater for real objects than for objects in pictures. *Perception 28*, 445–467.