The effect of variability in other objects' sizes on the extent to which people rely on retinal image size as a cue for judging distance

Rita Sousa Research Institute MOVE, Faculty of Human Movement Sciences, Vrije University, The Netherlands

Jeroen B. J. Smeets Research Institute MOVE, Faculty of Human Movement Sciences, Vrije University, The Netherlands

Eli Brenner Research Institute MOVE, Faculty of Human Movement Sciences, Vrije University, The Netherlands

Retinal image size can be used to judge objects' distances because for any object one can assume that some sizes are more likely than others. It has been shown that an increased variability in the size of otherwise identical target objects over trials reduces the weight given to retinal image size as a distance cue. Here, we examined whether an increased variability in the size of objects of a different color, orientation, or shape reduces the weight given to retinal image size when judging distance. Subjects had to indicate the 3D position of a simulated target object. Retinal image size was given significantly less weight as a cue for judging the target cube's distance when differently colored and differently oriented target objects appeared in many simulated sizes but not when differently shaped objects had many simulated sizes. We also examined whether increasing the variability in the size of cubes in the surroundings reduces the weight given to retinal image size when judging distance. It does not. We conclude that variability in surrounding or dissimilar objects' sizes has a negligible influence on the extent to which people rely on retinal image size as a cue for judging distance.

Keywords: depth, size, prior

Citation: Sousa, R., Smeets, J. B. J., & Brenner, E. (2012). The effect of variability in other objects' sizes on the extent to which people rely on retinal image size as a cue for judging distance. *Journal of Vision*, *12*(10):6, 1–8, http://www.journalofvision.org/content/12/10/6, doi:10.1167/12.10.6.

Introduction

When an object's size is known, one can use the size of its retinal image to estimate its distance (Gillam, 1995). When the size of an object is not known, people still use its retinal image size as a cue for distance (Collet, Schwarz, & Sobel, 1991; Lugtigheid & Welchman, 2010; Sousa, Brenner, & Smeets, 2011a; Sousa, Brenner, & Smeets, 2011b), presumably because they consider that the object is more likely to have certain sizes than others in the prevailing context. This is not an unreasonable assumption. For instance, if a cube shaped object is a lamp hanging in a room, it is unlikely to have sides of less than 10 cm or more than 1 m. Retinal image size should be given less weight in distance estimation if assumptions about likely sizes are considered to be less trustworthy (Knill, 2007; Mamassian & Landy, 2001). And indeed, we recently showed that if people are shown cubes of different sizes on consecutive trials, size is given less weight as a cue for distance than if they are shown cubes of the same size on all trials (Sousa, Brenner, & Smeets, 2011a). Thus, priors concerning object size quickly adapt to the statistics of the environment in question.

The influence of variability in object size on the weight given to retinal image size when judging distances (Sousa et al., 2011a) is analogous to findings in slant perception. When judging slant from retinal image shape, presenting differently shaped target objects on successive trials (similar objects with different aspect ratios) led to less weight being given to retinal image shape as a slant cue than if the same target object was presented on all trials (Seydell, Knill, & Trommershäuser, 2010). Seydell et al. (2010) showed that the differently shaped objects influenced the weight given to retinal image shape when judging a target object's slant even if the differently shaped objects could easily be identified by their color, but not if the differently shaped objects were of a different category (ellipses rather than diamonds). They also reported that varying the shapes of objects in the surrounding

doi: 10.1167/12.10.6 Received February 15, 2012; published September 11, 2012 ISSN 1534-7362 © 2012 ARVO

influenced the weight given to image shape as a slant cue. Muller, Brenner, & Smeets (2009), unlike Seydell et al. (2010), found no effect of surrounding objects' shapes when making slant judgments. There were several differences between the methods used in these studies that could have been responsible for the different conclusions. We propose that the nature of the subject's task with respect to the surrounding objects is the critical difference. In the study by Seydell et al. (2010) subjects made judgments about the surrounding objects' shapes on other trials, so changing the surrounding objects was equivalent to changing the previously presented targets. In the study by Muller, Brenner, & Smeets (2009) subjects made judgments for objects of a single shape surrounded by a plane of objects of which the shape was irrelevant, and only the latter objects' shapes were changed between conditions. Together, these results suggest that the weight given to image shape as a slant cue is only influenced by the variability in the shapes of objects of which the shape is relevant to the task at some time.

In the present study we aim to examine the issues discussed in the previous paragraph for size judgments. Does variability in the size of objects of a different color or a different shape influence the weight given to retinal size as a distance cue? Does variability in the size of irrelevant surrounding objects influence this weight? To find out, we compared distance judgments for identical target cubes in a block in which all objects had roughly the same size (*consistent* block) and a block in which there were also objects of other sizes (*mixed* block). We obtained a direct measure of the influence of the target's size on its judged distance from the difference between the judged distances of target cubes that were at exactly the same position but had slightly different sizes. We ran two experiments. In the *different color, orientation, or shape experiment*, only one object was visible in each trial and the target cube was presented interleaved with other target objects. The other target objects were cubes of a different color, cubes with a different orientation or spheres. If the sizes of the other target objects are considered independently of those of the target cubes, because of the difference in color, orientation, or shape, there will be no difference between responses to the target cubes in the *consistent* and *mixed* blocks. If subjects make no distinction between the different types of objects (the target cubes and the other target objects), the other target objects' sizes will influence the responses to the target cubes, so image size will be given more weight in the *consistent* block than in the *mixed* block. In the *context experiment*, a target cube was presented together with one or more other objects on each trial. If the variability in the other objects' sizes is considered, image size will be given more weight in the *consistent* block (in which other objects always had the same size) than in the *mixed* block (in which they had different sizes).

Methods

Subjects

In total, 23 subjects took part in the experiments, but not all subjects took part in all parts. Twelve subjects participated in the *color* and *orientation* sets, and 11 participated in the *shape* set of the *different color, orientation, or shape* experiment. Eight subjects participated in the first two configurations, and 20 participated in the third configuration of the *context* experiment. None of the subjects knew the purpose of the experiments. All of them had normal binocular vision as assessed with the Randot stereo fly test (median stereo acuity of 40 s of arc).

Apparatus

We used the same setup as in our previous study (Sousa, Brenner, & Smeets, 2011a) with mirrors that reflect the images from two CRT monitors (1,096 × 686 pixels, 47.3 × 30.0 cm) to the two eyes to produce simulations of three-dimensional objects (see Figure 1). New images were created for each eye with the frequency of the refresh rate of the monitors (160 Hz). The 3D positions of the subject's head and right index finger were recorded at 250 Hz using Infrared Emitting diodes (IREDs) and an Optotrak 3020 system (Northern Digital, Inc.).

One IRED was attached to the nail of the subject's right index finger and three others to a mouthpiece with a dental imprint. The positions of the subject's eyes relative to the mouthpiece were determined in advance. The measured position and orientation of the mouthpiece was used to adapt the images to the eyes' changing positions. This was necessary because subjects were allowed to move their head freely during the experiments (although they could not move very far since they had to look into the mirrors). The calibration procedure is described in detail elsewhere (Sousa, Brenner, & Smeets, 2010).

Stimuli

The objects were presented in total darkness. The simulated cube's surfaces had Lambertian reflectance with half the simulated illumination being ambient and the other half being from a distant light source above and 30° to the left of the subject. The simulated spheres, used in one block of the different color, orientation, or shape experiment, were self-luminant. The space in which the objects were presented was lower than the subjects' eyes and oriented downwards by about 30° so

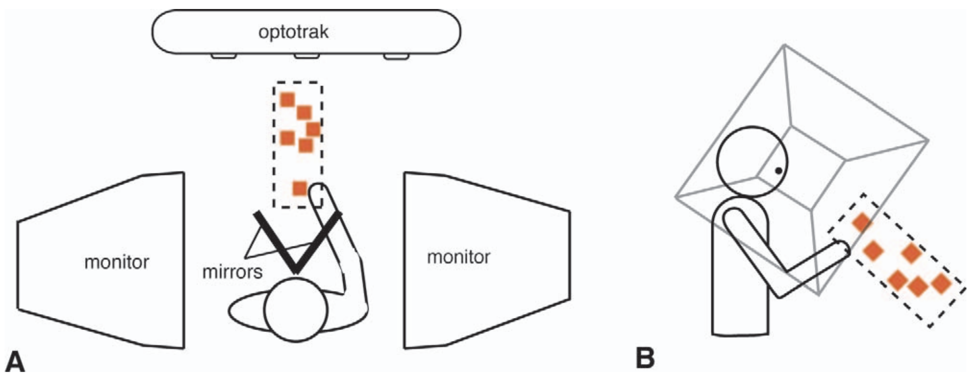

Figure 1. Schematic representation of the setup and of the stimuli (for a block in the context experiment). (A) Top view of the setup. The mirrors reflect the monitors' images, so that virtual stimuli are presented in the area indicated by the dashed rectangle. (B) Lateral view of the setup.

that the subjects pointed at a comfortable height. The space was elongated along the line of sight (depth axis; Figure 1B). Unless mentioned otherwise, all the simulated cubes had the same orientation: their edges were aligned with the edges of the volume of space.

The ocular convergence that was required to fixate the object, the motion parallax when the subject moved his or her head, the relative disparity between the edges within the simulated cubes, and the relative disparities between the objects when more than one object was presented in the scene were all consistent with the simulated distance. Positioning the cubes in the above-mentioned volume of space meant that the range of possible heights and lateral positions in the visual field was larger for nearby objects, but more distant objects were not systematically higher in the visual field or further to one side.

Each set or configuration within an experiment was tested with two blocks (*consistent* and *mixed*). Within each block, simulated red target cubes of two sizes were used, with sides of 1.0 or 1.2 cm. Note that these sizes refer to the simulated object, not to the size of the image on the screen (or to the retinal image size). Two different sizes were necessary for calculating the influence of retinal image size, as described in the analysis section. Both simulated target cubes were presented at the same 60 positions. In the *consistent* block, the other objects had sizes of 1.1 cm. In the *mixed* block they had random sizes between 0.5 and 3 cm. Within each set or configuration, the order of the blocks was counterbalanced across subjects. Within each block, the object positions and sizes were presented in random order.

Different color, orientation, or shape experiment

In the *different color, orientation, or shape experiment*, subjects were presented with a single object on each trial. This was either one of the two red target cubes or another target object chosen from one of the three sets of other objects: different *color*, *orientation,* or *shape* (see Figure 2). In the *color* set, the other target objects were blue cubes. In the *orientation* set, the other target objects were red cubes that were rotated so that their diagonals (rather than their sides) were aligned with the volume in which they were placed. In the *shape*

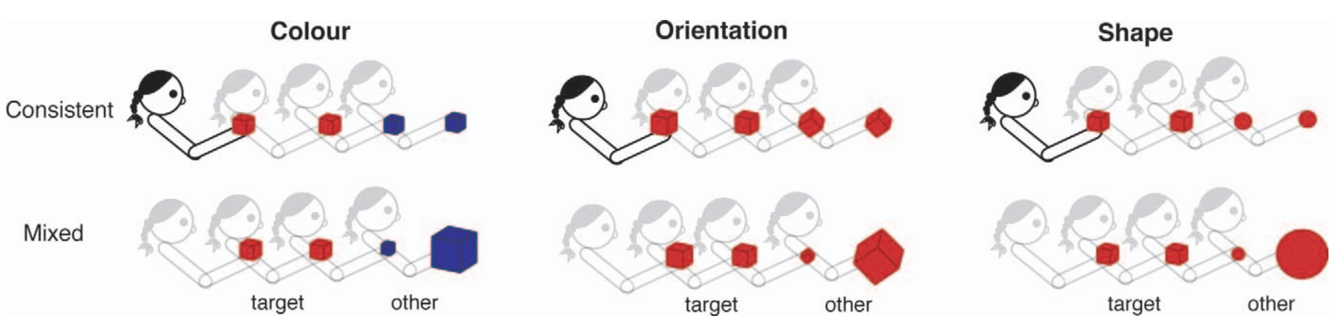

Figure 2. A schematic representation of the three sets in the different color, orientation, or shape experiment. Each set had two blocks: *consistent* and *mixed*. In the *consistent* block all the objects had roughly the same size. In the *mixed* block the size varied. In each panel the two leftmost examples represent the target objects and the two rightmost examples represent the other target objects. Only the latter differ between the six blocks.

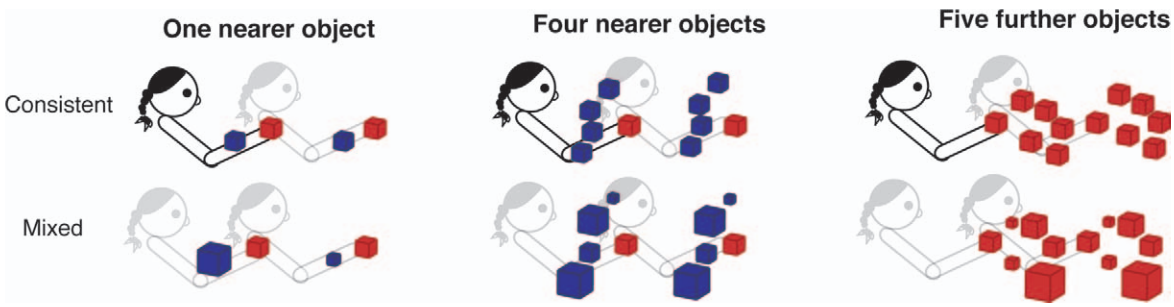

Figure 3. A schematic representation of three configurations in the context experiment. Each configuration had two blocks: *consistent* and *mixed*. In the *consistent* block all the cubes had roughly the same size. In the *mixed* block the context objects' sizes varied. The target cube was always red. When the configuration was *five further objects*, all six cubes were red. In that case, the target was identified by being the nearest cube.

set, the other target objects were red spheres. As mentioned above, the two red target cubes were each presented at 60 positions. Both these positions (120 trials) and the other target objects' positions (120 other trials) were the same for the *consistent* and *mixed* blocks (but presented in different random orders). All objects were positioned within a volume of space of 8 × 8 × 20 cm (width × height × depth) that was centered about 45 cm from the subject's eyes.

Context experiment

In the context experiment the target was always a red cube with sides of 1.0 or 1.2 cm. The surrounding items varied across the three configurations (*one nearer context object*; *four nearer context objects*; *five further context objects*) and two blocks (*consistent* or *mixed*; see Figure 3). We tested three different configurations because each has its own advantages and disadvantages. Having more objects makes the distinction between surroundings with objects of the same size and ones with objects of many sizes more evident, but it introduces additional cues for distance, so that retinal image size may altogether be given less weight. Similarly, placing objects further than the target gives us more freedom to vary the distances and the sizes of the surrounding items, but Sousa et al. (2010) showed that disparity relative to the furthest object is used as a cue to distance, so with many distant objects retinal image size may be given less weight when judging distance. The nearer context objects were blue cubes that were at least 2.5 cm nearer than the target. The further context objects were red cubes that were at least 5 cm further away than the target. In the latter case the distance could be larger because such cubes could be placed beyond reach, making it unnecessary to color the context cubes differently than the target cubes.

The target cubes were positioned within a volume of space of 8 × 8 × 20 cm, 22 × 8 × 10 cm, and 8 × 12 × 20 cm (width × height × depth), respectively for the *one nearer context object*, *four nearer context objects,* and *five further context objects* sets. All volumes were centered about 44 cm from the subject's eyes. The context cubes were positioned in slightly different volumes. The one nearer context cube was presented at random positions in a volume of space of 8 × 8 × 23 cm (width × height × depth) and its average distance was 35.5 cm. The four nearer context cubes were presented at random positions within a single plane about 39 cm from the subject's eyes (22 × 8, width × height). The five further context cubes were positioned in a volume of space of 8 × 12 × 55 cm that was centered about 60 cm from the subject's eyes. The cubes never overlapped laterally. The context cubes were placed at the same (semirandomly chosen) positions for each pair of trials in which the two targets were presented at the same position.

Procedure

Subjects started each pointing movement with their hand near their body. They were instructed to move their unseen index finger to the center of a specific object. In the different color, orientation, or shape experiment they were to move to whatever object was presented. In the context experiment they were either to move to the red cube or to the nearest one. The pointing movement was considered to have ended if the hand had moved less than 1 mm in 300 ms and was within 30 cm of the centre of the volume of possible target object positions. At that moment the finger position was saved (as was that of the eyes) and the objects disappeared. The next object or group of objects only appeared after the subject had brought the hand back near the body.

Analysis

The influence of the variability in the objects' sizes on the use of retinal image size to judge the target cube's

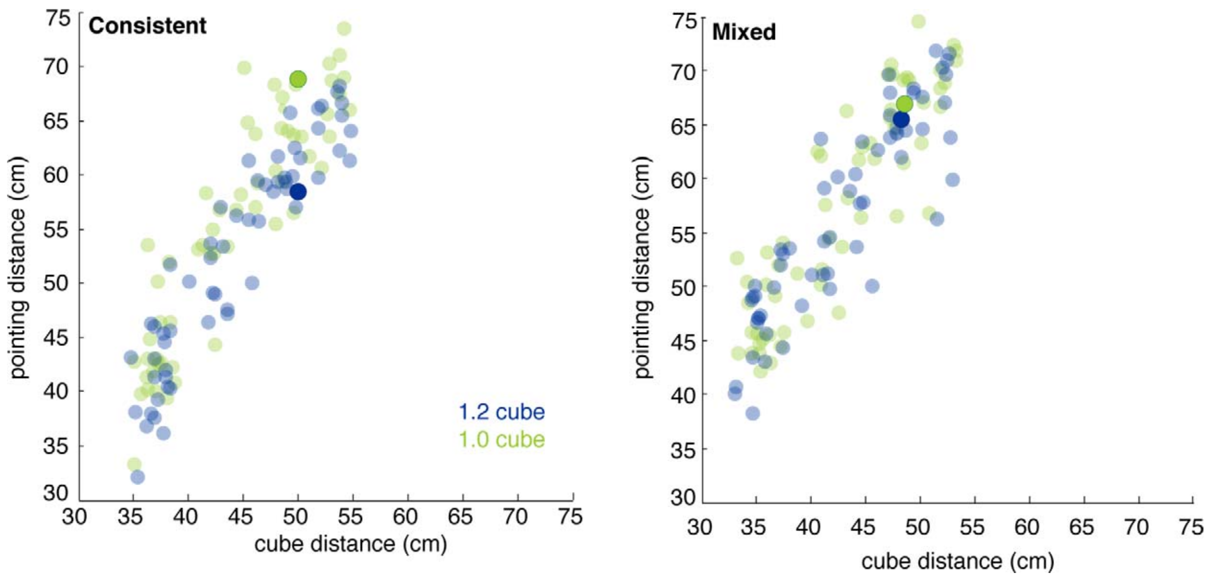

Figure 4. One subject's pointing distances for the 1.0 cm cubes and the 1.2 cm cubes for the *consistent* and *mixed* blocks (data from the color set in the different color, orientation, or shape experiment). Each point represents one trial. The highlighted dots in each panel show a pair of matched trials. The cube distance is not completely identical for the matched cubes because although the cube was at exactly the same position in space, the eyes were not always at exactly the same place.

distance was evaluated by averaging the differences between the pointing distances for the matched 1.0 and 1.2 cm target cubes over all 60 positions. Since the cubes were presented at the same 60 positions for the two target sizes, the average difference between the pointed distances for the 1.0 cm cubes and the 1.2 cm cubes provides a direct measure of the influence of the target's size on its judged distance (see Figure 4). We calculated the average difference in pointing distance separately for the *mixed* and *consistent* blocks and tested whether

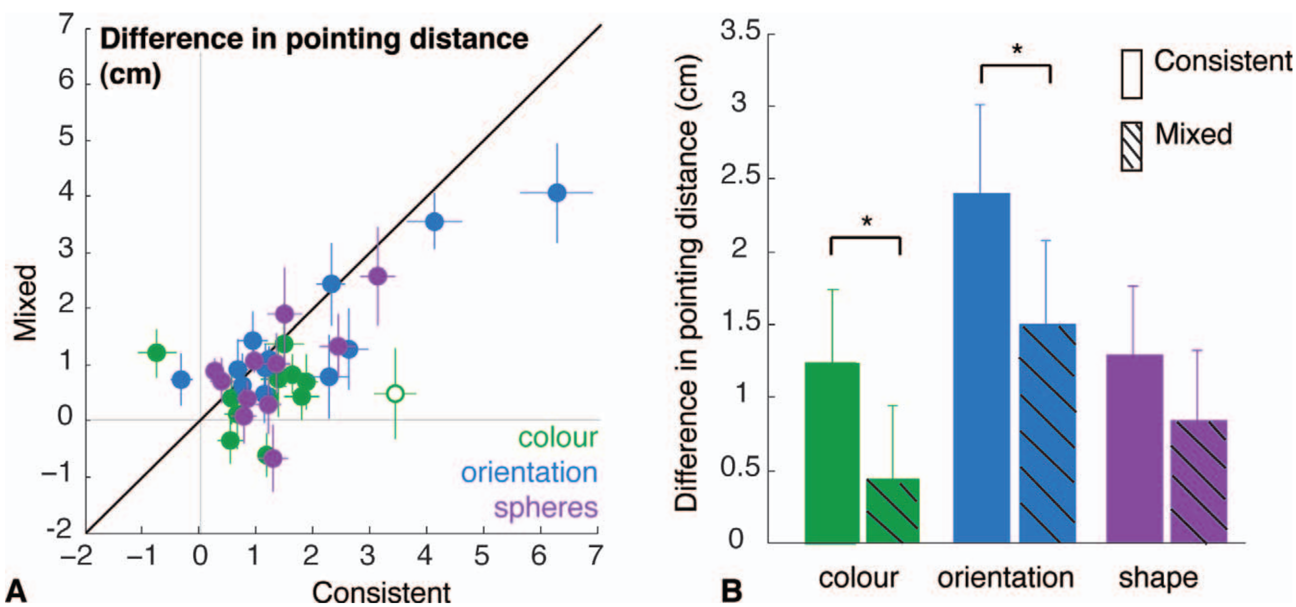

Figure 5. Influence of retinal image size on distance judgments in the different color, orientation, or shape experiment. (A) The average difference in pointing distance for the *mixed context* block as a function of the average difference in pointing distances for the *consistent context* block. Each point represents one subject. The error bars are standard errors across the 60 red target cube positions. (B) The mean across subjects of the average difference in pointing distance for the *consistent* and *mixed* blocks, for the *color*, *orientation,* and *shape* sets. The error bars are standard errors across participants. The difference for the *consistent* block was significantly smaller than for the *mixed* block in the color set ($t_{11} = 4.4$, $p = 0.001$) and in the orientation set ($t_{11} = 2.8$, $p = 0.016$), but not in the shape set ($t_{10} = 2.0$, $p = 0.076$).

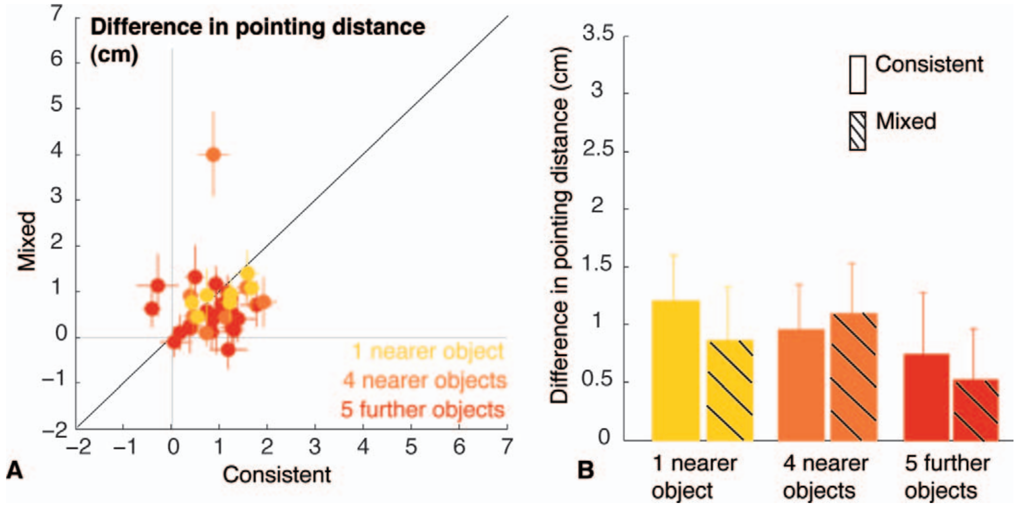

Figure 6. Influence of retinal image size on distance judgments in the context experiment. (A) The average difference between the pointing distances for the 1.0 and 1.2 cm cube target in the *mixed* block as a function of the average difference between such pointing distances in the *consistent* block. Each point represents one subject. The error bars are standard errors across target positions. (B) The mean across subjects of the average difference in pointing distance for the *consistent* and *mixed* blocks, for the three sets. The *consistent* blocks are not significantly smaller than the *mixed* blocks in the *one object nearer* than the target set ($t_7 = 1.9$, $p = 0.096$), in the *four objects nearer* than the target set ($t_7 = -0.3$, $p = 0.775$), or in the *five objects further* than the target set ($t_{19} = 1.3$, $p = 0.195$). The error bars are standard errors across participants.

this difference was reliably smaller (across subjects) for the *mixed* block than for the *consistent* block with paired, one-tailed *t*-tests. We determined the mean and standard error of the average differences (across subjects) to summarize the effect.

We also calculated the slopes of pointing distance as a function of target distance to determine to what extent subjects were using informative distance cues for estimating the distances.

Results

Figure 5 shows the influence of cube size for all the subjects in the three sets of the different color, orientation, or shape experiment in which only a single object was present in each trial. In Figure 5A, each dot corresponds to a subject's average difference in pointing distance for the red target cubes in the two blocks. The open dot corresponds to the subject whose data is presented in Figure 4. Figure 5B compares the mean differences in pointing distance for the *consistent* and *mixed* blocks. When the red target cubes and the other target objects differed in color or orientation, the difference in pointing distance was significantly larger ($p < 0.05$) when the other target objects were of roughly the same size (*consistent* block) than when they had many sizes (*mixed* block). The difference between the *consistent* and *mixed* blocks was not significant when the other target objects differed in shape (spheres). The average slope of pointing distance as a function of simulated distance in the *different color, orientation, and shape experiment* was 0.81 with a standard deviation of 0.28. The slopes did not differ significantly between the consistent and mixed blocks.

Figure 6 shows the influence of cube size for all the subjects in the three configurations of the context experiment, in which the target cube was always presented with one or more context cubes. In Figure 6A each dot corresponds to one subject's average difference in pointing distance. Figure 6B compares the subjects' average differences in pointing distance for the *consistent* and *mixed* blocks when *one context cube* was nearer, *four context cubes* were nearer, and *five context cubes* were further than the target. The influence of cube size when the context cubes are of roughly the same size (*consistent* block) is never significantly different from when the context cubes have many sizes (*mixed* block). The smaller difference in pointing distance when there are five further objects suggests that retinal image size is indeed altogether given less weight as a distance cue in this condition due to the additional disparity cue (Sousa et al., 2010), but due to the large variability between subjects we cannot be sure about this. The average slope of pointing distance as a function of simulated distance in the *context experiment* was 0.83 with a standard deviation of 0.36. Again, the slopes did not differ significantly between the consistent and mixed blocks.

Discussion

Our findings concerning how variability in the size of similar target objects influences the extent to which people rely on size as a distance cue are globally consistent with findings about the influence of other objects' shapes on slant judgments, as described in the Introduction. Even when the other objects that one interacted with had a different color or were orientated differently than the target cube, variability in their sizes decreased the weight given to retinal image size as a distance cue, as one would expect of objects that were regarded as being from the same category as the red target cube. This can be seen in the systematic differences between the *consistent* and *mixed* blocks in Figure 5B. When the other target objects had a different shape (spheres rather than cubes), variability in their size did not significantly affect the use of size as a cue for distance. It seems that the clearer the distinction between the red target cubes and the other target objects, the less influence other objects' sizes have in the response to the red target cubes. Thus the prior that we are manipulating is not for size in general, but for the size of a cube.

As Seydell et al. (2010) found for slant, we find that people consider objects of different colors but not objects with different shapes to belong to the same class when considering dimensions. We show that it is the object's shape that matters, not the retinal image shape, because rotating the object was not enough to make it be considered a different object. As Muller et al. (2009) found for shape, we find that varying the size of the surrounding objects did not significantly influence the weight given to retinal image size when judging distance. We conclude from these findings that the object that one interacts with has a much stronger influence on the prior than surrounding objects. This is consistent with the absence of significant effects in the context experiment and in Muller et al. (2009). It is also consistent with the reported influence of surrounding items in Seydell et al. (2010), because in that study the participants interacted with the surrounding objects: participants had to make slant judgments for all the nine items in the array. The influence of other items was identical in such a nine items array (their fourth experiment) and when each item was presented in isolation (their fifth experiment).

Thus, variability in the surrounding or dissimilar objects' sizes has a negligible influence on the extent to which people rely on retinal image size as a cue for judging distance. We already knew that people update their prior expectations about objects properties (Adams, Graf, & Ernst, 2004; Körding & Wolpert, 2004) and their confidence in them (Seydell et al., 2010; Sousa et al., 2011a) as a result of exposure during experiments. We here show that the update is specific in that only the size of objects that are similar to the target object and directly relevant for the task at hand are taken into consideration.

Acknowledgments

This research was financed by a FCT (Portuguese Foundation for Science and Technology) and FSE (European Social Fund) PhD grant to RS.

Commercial relationships: none.
Corresponding author: Rita Sousa.
Email: ritass@gmail.com.
Address: Research Institute MOVE, Faculty of Human Movement Sciences, Vrije University, The Netherlands.

References

Adams, W. J., Graf, E. W., & Ernst, M. O. (2004). Experience can change the 'light-from-above' prior. *Nature Neuroscience, 7*(10), 1057–1058.

Collet, T. S., Schwarz, U., & Sobel, E. (1991). The interaction of oculomotor cues and stimulus size in stereoscopic depth constancy. *Perception*, *20*, 733–754.

Gillam, B. (1995). The perception of spatial layout from static optical information. In: S. R. W. Epstein (Ed.), *Perception of space and motion*. (pp. 23–67). London: Academic Press, Inc.

Knill, D. (2007). Learning Bayesian priors for depth perception. *Journal of Vision*, *7*(8):13, 1–20, http://www.journalofvision.org/content/7/8/13, doi:10.1167/7.8.13. [PubMed] [Article]

Körding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Letters to Nature, 427*(15), 224–247.

Lugtigheid, A., & Welchman, A. (2010). A surprising influence of retinal size on disparity-defined distance judgments. *Journal of Vision, 10*(7):63, http://www.journalofvision.org/content/10/7/63, doi:10.1167/10.7.63. [Abstract]

Mamassian, P., & Landy, M. S. (2001). Interaction of visual prior constrains. *Vision Research, 41*, 2653–2668.

Muller, C., Brenner, E., & Smeets, J. B. J. (2009). Maybe they are all circles: Clues and cues. *Journal of Vision*, *9*(9):10, 1–5, http://www.journalofvision.org/content/9/9/10, doi:10.1167/9.9.10. [PubMed] [Article]

Seydell, A., Knill, D., & Trommershäuser, J. (2010). Adapting internal statistical models for interpreting visual cues to depth. *Journal of Vision*, *10*(4):1, 1–27, http://www.journalofvision.org/content/10/4/1, doi:10.1167/10.4.1. [PubMed] [Article]

Sousa, R., Brenner, E., & Smeets, J. B. J. (2010). A new binocular cue for absolute distance: Disparity relative to the most distant structure. *Vision Research, 50*(18), 1786–1792.

Sousa, R., Brenner, E., & Smeets, J. B. J. (2011a). Judging an unfamiliar object's distance from its retinal image size. *Journal of Vision*, *11*(9):10, 11–16, http://www.journalofvision.org/content/11/9/10, doi:10.1167/11.9.10. [PubMed] [Article]

Sousa, R., Brenner, E., & Smeets, J. B. J. (2011b). Objects can be localized at positions that are inconsistent with the relative disparity between them. *Journal of Vision, 11*(2):18, 1–6, http://www.journalofvision.org/content/11/2/18, doi:10.1167/11.2.18. [PubMed] [Article]