A new binocular cue for absolute distance: Disparity relative to the most distant structure

When an object is presented binocularly at various distances in complete darkness the range of distances is usually underestimated. We found that adding a second object can reduce the extent to which the range of distances is underestimated. However this only happens if the second object is further than the one of which the distance is to be judged. We propose that the relative disparity between the two objects limits the possible distances of the nearer object because the lines connecting the further object to each of the two eyes must converge in the distance.

In the real world, people appear to be very good at judging reachable egocentric distances. It is easy to reach out and pick a cup of coffee when there is ample visual information about the distance of the cup and the hand. However, in the dark, with only a single visible object, people’s distance judgments can be biased (Tresilian et al, 1999; Mon-Williams & Tresilian, 1999; Morrison & Whiteside, 1984). There is a tendency to see objects near some default distance (Gogel, 1961; Gogel & Tietz, 1973; Foley, 1980): objects that are further away appear nearer and near ones appear further away.

This contraction bias might be caused by a general tendency to respond towards the mean (Poulton, 1981). However a more parsimonious explanation can be given in terms of cue combination (Landy et al, 1995; Hillis et al, 2004): in many studies in which a contraction of the range of distances is found, authors made sure that cues such as image size and height in the visual field are not informative about distance, for instance by keeping their values constant. They did so in order to isolate binocular distance cues (e.g. Johnston, 1991; Philbeck & Lomis, 1997; Viguier et al, 2001). However subjects may rely on such cues to some extent, even if their values do not vary, which would lead to a contraction bias. In that case, if we add informative cues, the relative weight given to the non-informative cues will be reduced, resulting in less contraction bias.

We investigated whether adding the simplest possible visual surrounding, a single reference, helps judge the distance of a target in the dark, and specifically whether it reduces the contraction bias. Previous studies have found that adding a complex visual surrounding makes distance judgments more accurate (Brenner & van Damme, 1999; Foley, 1977; Coello & Magne, 2000) but it is not known why it does so. Adding a reference introduces information from relative disparity between the two structures, which makes it possible to calculate target distance not only relative to oneself but also relative to the reference. But how could this help?

One way in which adding a second object could help is by providing a stable reference. Glennerster et al (2006) showed that subjects assume that certain aspects of the world are stable, and base their judgments about the surrounding environment on this assumption, even if there is direct evidence to the contrary. If a reference is constantly visible at a fixed position, subjects may use its position as a reference. If they do so consistently there will be less contraction bias, even if the estimate of that position is not correct. We will refer to this as the *stable reference* explanation.
Another way in which a second object could help is if it provides more precise depth information so that its distance can be judged more precisely than the first object. The subjects can then first judge the distance to the reference, and then judge the target’s distance relative to the reference using relative disparity. According to this better precision explanation, subjects will benefit from a reference if its distance can be judged better than can the distance of the object of interest.

A third suggestion is that the range of disparities present in the image provides information about distance. Glennerster et al (1998) pointed out that the disparities between objects with a given distance between them are larger when the objects are nearby than when they are far away (for geometrical reasons), so if the disparities between objects are small, the objects are likely to be far and if the disparities are large then objects are likely to be near (see also Harris, 2004; Hibbard, 2007). If the range of disparities is considered, even when there are only two objects, then there may be a tendency to judge each object to be nearer when the reference is at a very different distance than the target. We will refer to this hypothesis as the peak disparity explanation.

We propose a fourth option: relative disparity as a limiting factor. Relative disparity does not provide direct information about distance: a given magnitude of relative disparity between two objects is compatible with many viewing distances and separations in depth between the two objects (figure 1a and b). We will refer to the angle between the lines from the object to the two eyes as object vergence (irrespective of the orientation of the eyes). An object vergence of zero corresponds with the object being extremely far away. Relative disparity is the difference between two objects’ vergences, and no object can have an object vergence below zero, so relative disparity can constrain the possible distances of the nearer of the two objects. Thus, according to the limiting factor explanation, a second more distant object influences the perceived distance to the original object, because it reduces the range of possible positions. Figure 1 illustrates that considering the disparity with the reference the near dot can be at the positions shown in a, b and c, but no further than the distance shown in c. For the distance in d the reference would have an impossible object vergence. A reference that is nearer than the object of interest does not limit the possible viewing geometries (unless it is very near the nose), so it should not influence distance judgments. The results of previous experiments provide some support for this proposal in that judgments of a far target’s
distance appear to be influenced less by adding a nearer target than are judgments of a nearby target by adding a target further away (Bank, 1958; Foley, 1985; Gogel, 1972).

**Figure 1.** Relative disparity as a limiting factor. The same distance between retinal points (retinal disparities, red arcs) can correspond to different positions and separations of two objects (a, b). If the far (blue) object is infinitely far away, the near (green) object will be at some distance that depends on its disparity relative to the far object (c). The near object cannot be further than this distance, because a further distance for the near object would lead to an impossible object vergence for the far object (d).

In this study we tested the four described ways in which adding a reference object can help judge target distance. We examined whether stability in the reference position across trials matters, whether the reference has to provide more reliable information than the target for it to have an effect, whether the distance in depth between the reference and the target is important and whether it matters which object is further away. To do so, we asked subjects to point at small virtual targets that were presented with or without a reference objects.

To assess the **stable reference** explanation, we fixed the reference at a certain position in one condition and we positioned it at a different position on each trial in another condition. **According to the stable reference** explanation, there should be less contraction bias in the condition in which the reference is always at the same position compared with the condition in which the reference is at a different position on each trial.

To evaluate the **better precision** explanation, subjects were asked to point at a sphere with a cube as a reference as well as at a cube with a sphere as a reference. The cube had a constant (simulated) size, so its
distance could be judged from retinal image size as well as binocular cues. The cube also provides more reliable binocular information than an untextured sphere because the changes in the retinal images with distance are larger. If subjects only use the reference when it provides better information, then performance should not differ between pointing at the cube when it is presented alone and pointing at the cube when there is a reference sphere present, but there should be a difference for pointing at the sphere with or without a reference cube.

The distance between the target and the reference varied across trials. According to the peak disparity difference explanation, if the distance (in depth) between the objects is small subjects should judge them to be further away than if the distance between the objects is large. The limiting factor explanation makes a similar prediction, but only if the reference is further than the target. Otherwise the judged distance should be the same as when there is no reference. To evaluate these two explanations we compared judgments with and without a reference in relation to the relative disparity between the target and the reference.

**Methods**

**Subjects**

Seven subjects participated in the experiment. All of them were right-handed, naive about the purpose of the experiment, and had normal binocular vision.

**Apparatus**

To create three-dimensional virtual visual stimuli, we used a set-up with mirrors that reflect the images from two monitors (1096 x 686 pixels, 47.3 x 30.0 cm) to the two eyes to produce binocular simulations of the objects. The mirrors are half silvered, so that when occluding panels behind them are detached one can see though them. This was used for testing the calibration (see below). The computers that generated the images were two Apple G5s. New images for each eye were created with the frequency of the refresh rates of the two CRT monitors (160 Hz). The 3D positions of the subject’s head and right index finger were recorded at 250 Hz using Infra-red 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. The task was to bring this finger to the target position. Three IREDs were attached to a bite-board. The positions of the subject’s eyes relative to the bite-board were determined in advance (see calibration section). The bite-board was held in the mouth, but it was not attached to anything else, so subjects could move their head freely during the experiments. This allowed them to move their hand naturally. Knowing the time-varying positions of the IREDs attached to the bite-board, and the fixed positions of the eyes relative to the IREDs, allowed us to adapt the images to changes in the eyes’ positions.

Calibration

To determine the eyes’ positions relative to the bite-board IREDs (and therefore also each subject’s inter-ocular separation), we constructed a long tube with IREDs at one end and two pairs of intersecting threads inside (one pair at the end and one in the middle; figure 2).

The subjects were asked to look into the tube with one eye and align the intersections of the threads. By doing so, the two intersections were aligned with the optical center of the eye. Subjects performed this alignment 20 times with the tube in different orientations relative to the eye, yielding 20 lines that connected the two intersections with the eye. The point in space (relative to the bite-board) at which the median distance to the 20 lines was smallest was considered to be the position of the eye. This calibration procedure was done for each eye separately.

Figure 2. Determining the location of the eye. When the intersections of the two pairs of threads (thin dashed lines) are visually aligned, the calibration tube is aligned with the line of sight. Doing this with the calibration tube in various orientations relative to the head allows one to determine the location of the eye (intersection of the dotted lines).

To calibrate the virtual space we attached a frame between the monitors, at about the same distance from the mirrors as the monitors, in the space where the images were seen (figure 3a). This frame had IREDs attached to it. We matched reference lines on the monitors, as reflected by the half silvered mirrors, with lines on the frame as seen though the mirrors (with
the occluding panels removed). To check that our calibration was successful, we displayed a target at the calculated position of the IRED on the finger and compared this target’s position with the IRED’s actual position as seen through the half silvered mirror. The systematic errors were no more than 2 mm (within the range of target positions used in the experiment).

**Figure 3.** The setup. In this schematic representation the subject is behind the mirrors, looking at the images on the monitors via the mirrors. (a) The calibration frame (dashed lines) with a square in the middle, the sides of which were matched with lines presented on the monitor (red lines). (b) The 8x8x20cm space within which the objects were presented (dashed lines).

**Stimuli**

We presented either one or two objects in total darkness. The objects could be spheres or cubes. If there were two objects, they were always different. One of the objects was the target of the pointing movements. The simulated objects appeared at random locations within a 8 x 8 x 20cm volume of space of which the long midline was more or less aligned with the gaze direction when looking straight ahead (laterally) and downwards by about 30° (fig 3b). The position and orientation of the above-mentioned volume of possible positions was fixed in space. Its orientation was based on where subjects held their heads in pilot experiments. This space was therefore only approximately aligned with the above-mentioned gaze direction for individual subjects, because subjects were free to move their heads. However this freedom was limited by having to look into the mirror, so the deviation from the intended alignment was never very large. On average, the center of the space was 44.1 cm from the subjects’ eyes, so the objects were at distances of between about 34 and 54 cm from the subject.
When the simulated object was a sphere, it was red and its visual extent (diameter) varied randomly between 0.15° and 0.59°. Its angular size varied independently of its distance, so its simulated size was on average smaller when it was nearer. When the simulated object was a cube, it was red and had a constant simulated size (1cm sides), so its angular size varied systematically with its distance.

Procedure

The subjects received instructions about the pointing movements that they had to make. They received no instructions about where to look, and were therefore free to direct their gaze wherever they wanted. They started each pointing movement with their right hand near their body. When the target appeared, they had to move their unseen index finger to where they saw the target, and hold the finger steady until the trial ended with the target disappearing. The trial ended if the hand was within 30 cm of the center of the possible range, and had not moved more than 1 mm in 300 ms. At that moment the finger position was recorded. After the target disappeared, the subjects had to bring the hand back near to their body and wait until a new target appeared at another location.

Conditions

There were eight conditions (figure 4). The target was either a sphere or a cube. When the target was a sphere, it was either presented alone, or else it was presented together with a reference cube. The reference could either be at a fixed position or at a different position on each trial. When it was at a fixed position, at a distance of about 37cm or about 51cm, the reference remained visible between the trials. When the reference was at a different position on each trial, it disappeared at the same time as the target and was absent until a new target appeared, so nothing was visible between trials. When the target was a cube, it was either presented alone or together with a reference sphere at a random position. All types of references could be further or nearer than the target.
Since trials in which the reference was nearer and further than the target were interleaved, the eight conditions were presented in five groups (each cell of figure 4 represents a group). Each of the five groups of conditions was presented three times using three different sets of 60 target (and reference) positions. The same sets of target positions were used in all five groups. The 37cm reference was used for two of the sets of trials in the group with a fixed reference position. The same sets of reference positions were used in the two groups with random reference positions.

The sets were presented in a semi random order. The nine sets in which subjects pointed at the sphere (three for each of the three blocks) were performed before the six sets in which they pointed at the cube. Some days before the experiment, the subjects practiced the task with three different sets.

### Figure 4

The eight conditions. The five cells in the table show how the conditions were combined into five groups.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Pointing at the sphere</th>
<th>Pointing at the cube</th>
</tr>
</thead>
<tbody>
<tr>
<td>No reference</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
</tr>
<tr>
<td>Reference further</td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
<tr>
<td>Reference nearer</td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
<tr>
<td>Reference at a fixed position</td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
</tr>
</tbody>
</table>

**Analysis**

We determined the distance of the target and the pointing finger from the head (a position halfway between the two eyes) at the end of each trial. A smooth curve representing the relationship between perceived and presented distance was constructed by averaging the pointing distances as a function of target distance with weights determined by a moving Gaussian window ($\sigma = 8$ cm). The smoothed line was only determined if there were at least 12 data points within ± 4 cm of the target distance in question.
For the spherical targets we conducted a multiple regression with simulated distance, angular size and height in the visual field as the independent variables, and pointing distance as the dependent variable, to examine whether the random values that we assigned to the cues influenced where participants pointed. For the cubes, only simulated distance and height in the visual field were used in the multiple regression, because angular size co-varied with simulated distance. Separate regressions were conducted for each subject and condition. Paired student’s t-tests were used to test for consistency across subjects of differences between the slopes for conditions within the same columns in figure 4.

Results

Figure 5a shows a single subject’s data when pointing at spheres with the reference either absent or at a different position on each trial. When there is a reference further than the target, the pointing distance changes more with changes in the distance to the sphere (steeper slope) than when pointing at the sphere on its own. The subject also overestimated the distance less: data closer to the line that indicates veridical judgments. Having a reference nearer than the target was like having none at all.

There were systematic differences between where subjects pointed (some overestimated more than others) but curves summarizing the data of all the subjects illustrate that all subjects show the same difference between the conditions (fig 5b).

Whether the reference was continually visible at a fixed position or disappeared after each trial and was presented at a different position on the next trial made no difference to the pattern of results (compare figures 5b and c). When pointing at the cube we found overall steeper slopes but a similar effect of the second object (Figure 5d). The difference between pointing at the cube when it is on its own or with a nearby reference, and pointing at it when there is a reference further away, is not as conspicuous as the differences when pointing at the sphere.
Figure 5. Target positions and pointing positions. (a) Data for one subject when pointing at a sphere (with variable simulated size) with or without a reference cube (of constant simulated size) that was at a different position on each trial. Each dot shows the distances for one trial. The curves are smoothed representations of the data. (b) Average of the 7 subjects data for these conditions (c). Average results for the conditions with a reference cube at a fixed position. The data for the condition in which the sphere was presented on its own are the same as in panel b, and are shown here again to ease the comparison. (d) Average results for the three conditions when pointing at the cube.

That the differences are consistent across subjects is confirmed by examining the slopes from the regression analyses (see figure 6). For the sphere, the average contraction bias was smaller (larger slopes) when there was a reference further than the target than when there was none. When the reference was nearer than the target, the slopes did not differ significantly from when there was no reference. Slopes when pointing at the
cube with a reference nearer than the target, differ significantly from the ones when pointing at the cube with a reference further than the target.

Our regression analysis included image size and height in the visual field as factors, although neither was correlated with the simulated target distance in the conditions in which they were included as factors in the regression. Subjects nevertheless clearly considered the size of the target’s image when judging its distance (figure 6b). We found no significant contribution of height in the visual field.

Figure 6. Averages of subjects’ slopes from the multiple regressions. (a) Influence of simulated distance on pointing distance (with 95% confidence intervals). A slope of 1 means that subjects judged differences in distance veridically. Paired t-tests were used to compare the slopes across conditions (*p<0.05; ** p<0.01; ***p<0.001). (b) Influence of image size on pointing distance. A negative slope means that objects with smaller retinal images are judged to be further away (t-tests across subjects were used to determine whether the slopes were reliably different from zero: # p<0.05; ## p<0.01).

In figures 5b and d, the data when there are two objects is treated separately depending on which object was nearer, but the magnitude of the relative disparity is not considered. In order to evaluate the influence of the magnitude of relative disparity, we determined the difference between the pointing distance for individual targets with and without a reference, and plotted this as a function of the relative disparity when there were two objects. This comparison was possible since we had matched target positions across the different groups. When comparing matched trials we ignore differences in head position (for all the other analyses we
consider head position for each trial by taking the exact value of the head position at the moment of pointing). We felt justified in ignoring variability in head position for this comparison because the standard deviation in the position of the head during a session was only about 5mm. We plotted the effect that the reference had on the pointing distance as a function of the relative disparity between target and reference (Figure 7a). When relative disparity was positive (reference further) subjects systematically pointed nearer for larger relative disparities.

Figure 7. Relative disparity as a limiting factor. (a) Difference in pointing distance when there is no reference and when there is one, for matched target positions, as a function of the relative disparity when there is a reference. Positive values of relative disparity mean that the reference is further away (uncrossed disparity relative to the target). (b) A simple model. The dark blue curve shows the estimate of object vergence (the likelihood in Bayesian terms). It is presented here at the mean target position of about 44.1cm. Each green curve shows how likely different values of object vergence are, given a certain relative disparity with the more distant reference (this can be considered as a prior in Bayesian terms). These curves are based on it being considered to be unlikely that any of the objects would be further than 56 cm (object vergence below 6.12°). The light green curve is for a near-zero relative disparity. The dark green curve is for a relative disparity of 3°; values below 9.12° are considered unlikely because that value corresponds with a value of 6.12° for the reference. The green areas are obtained by combining (multiplying) the blue and green curves. They represent the perceived distance (the posterior distribution; peaks indicated by blue arrows). Distances were calculated for an inter-pupil distance of 60.7mm.

This is what both the limiting factor and the peak disparity difference explanation predict. When relative disparity was negative (reference nearer condition) there was no effect of the reference on pointing. This is in agreement with the limiting factor explanation, but not with the peak disparity difference explanation.
In summary, adding a reference only influences pointing at an object if the reference is further away. This is qualitatively consistent with relative disparity being used as a *limiting factor* in the manner illustrated in figure 1.

**Discussion**

We find that when judging the location of an object in the dark, adding a second object improves the judged distance of the first object if the second object is further away. Our evidence for this is the reduction of the contraction bias (steeper slopes) when there is a reference further than the target (Figs 5 and 6; a similar result can be found in Foley (1985)).

We did not find support for the *stable reference* explanation: there is no difference between the pointing distances when the reference was stable (figure 5c) and when it was at a different position on each trial (figure 5b). Glennerster et al (2006) found that in a virtual environment, subjects ignore cues from binocular vision and even distance walked, rather than acknowledge that the size of the scene changed. Although we found that reference stability across trials made no difference, we did observe that subjects used the size of the sphere as a cue for judging its distance. Thus they assumed that certain sphere sizes are more likely than others. This assumption can be inferred from the significant slopes for object size in the multiple regressions.

The finding that subjects considered image size for judging distance supports our suggestion that part of the observed contraction bias in previous experiments is caused by subjects relying on sources of information that do not vary with distance in the experiments in question (as a result of attempting to isolate the influence of certain sources of distance information), although they normally would vary with distance. The tendency to point too far (that has been reported before for a similar range of target distances; Foley, 1975) could have many causes, one being that subjects assume that the simulated objects are larger than the objects that we simulated.

The *better precision* explanation was not corroborated either. Even when the reference could be localized less precisely than the target (pointing at the cube with a sphere as a reference; figure 5d), the slopes were steeper when a more distant reference was present than when there was only the target or a nearer reference present. The effect of the reference when pointing at the cube was smaller than it was when pointing at the sphere (figure 6). This was expected because the cube can be localized more precisely, so
the slopes are already quite high without a reference and therefore the possible improvement is smaller. Despite the additional and more reliable distance cues for the cube, the extra information from the far reference is still given some weight.

The *peak disparity* explanation predicted that the range of disparities would influence the judged distance. This was not fully supported because having a reference nearer than the target makes no difference. The slopes are only steeper when the reference is further than the target. Thus only the *limiting factor* explanation is consistent with all the data.

We developed a simple model to evaluate the *limiting factor* explanation quantitatively. Assuming that object vergence is primarily judged from extra-retinal signals about the orientation of the eyes when fixating the object, and that uncertainty about the orientation of the eyes is normally distributed, we can describe the likely positions of the target by a Gaussian on an object vergence scale (figure 7b). If there is a reference further than the target, there is a lower limit to the possible values of object vergence for the reference: it cannot have an object vergence lower than zero. Thus the likely positions of the target (on the object vergence scale) are constrained by the disparity relative to the furthest object. Assuming that there is also some uncertainty about the magnitude of the relative disparity, there will be a smooth transition between possible and impossible values, rather than this being a step function. We modeled this by multiplying the above-mentioned Gaussian by a cumulative distribution with its inflexion point at the object vergence that corresponds with the disparity relative to the furthest possible position (infinity). This operation results in a new distribution with a shifted peak. In our model we consider that subjects will point at this shifted peak.

It was immediately evident when developing this model that considering infinity (zero object vergence; see figure 1) as our limit could not explain our data. However it is not unreasonable to assume that the most distant structure is nearer than infinitely far away, especially when one is in an enclosed space and is looking downwards. For a target at the average target distance (44.1cm), using the distance of the furthest reference as the limit for our space and 2° for the standard deviation, both for the target vergence and for the cumulative distribution (figure 7b), our model predicts an influence that is reasonably similar to the data (green dotted line in figure 7a).

Of course several of the assumptions are questionable, but the model indicates that the general reasoning is also quantitatively plausible. Moreover, it explains why the most distant target is seen at a more or less fixed distance, irrespective of its true distance, when it is far away (well beyond reachable distances; Blank, 1958):
the oculomotor estimate of distance is poor at such distances (Brenner & Smeets, 2000) so the distance limit dominates the judgment.

In sum, we demonstrated that disparity relative to the furthest object is used as a cue for distance. If there are large uncrossed disparities relative to the target, the target is judged to be nearer. This is a new distance cue that could explain why performance for isolated targets is much poorer than performance in a full scene.