Systematic distortions of perceptual stability investigated using immersive virtual reality

Lili Tcheang*   Stuart J Gilson*   Andrew Glennerster*†

March 24, 2004

*University Laboratory of Physiology, Parks Road, Oxford, OX1 3PT

†To whom correspondence should be addressed: ag@physiol.ox.ac.uk

Abstract

Using an immersive virtual reality system, we measured the ability of observers to detect the rotation of an object when its movement was yoked to the observer's own translation. Subjects judged whether the object rotated 'with' or 'against' them as they moved (following Wallach et al., 1974). We found that adding a stable visual reference frame made the perception of the the object's rotation more accurate (biases were reduced) and more precise (thresholds improved). In the absence of a stable background, most subjects' biases were such that a static object appeared to rotate away from them as they moved. The biases are consistent with an under-estimation of the distance that the observer walks.

Key words: 3D perception, stability, allocentric, reference frame.

Introduction

The apparent stability of the visual world in the face of head and eye movements has been a long standing puzzle in vision research. Much of the discussion of possible mechanisms has focused on methods of compensating for rotations of the eye such as saccades (reviewed by Burr, 2004), including evidence of neurons that appear to shift their retinal receptive field to compensate for eye position with respect to the head (e.g. Duhamel et al., 1997); changes in perceived visual direction around the time of a saccade (e.g. Ross et al., 2001); and descriptions of a 'stable feature frame' that could describe the visual direction of features independent of eye rotations (Feldman, 1985; Bridgeman et al., 1994; Glennerster et al., 2001).

There have been fewer proposals about the type of representation that observers might build when the head translates in space. This is a more difficult computational problem than for the case of pure rotations of the eye. For one thing, the depth of objects must be known in order to 'compensate' for a head translation. The visual system must maintain some representation of a scene that is independent of observer translation but there are currently few detailed proposals about what form it might take. The representation could include the world-centred 3D coordinates of points, in which case there must be a coordinate transformation from the binocular retinal images into this frame. There are suggestions that such a transformation may not be required and that a 'piece-wise retinotopic' map could be sufficient for navigation (e.g. Franz and Mallot, 2000) or perception of depth (Glennerster et al., 2001). However, these ideas have not yet been developed into a detailed model.

There have also been fewer experimental studies addressing the consequences

of head movements than there have been for eye movements. This is due in part to the practical difficulties involved in psychophysical investigations using a moving observer. In studies where observers move their head, the focus has often been on the perception of surface structure or orientation (e.g. Rogers and Graham, 1982; Bradshaw and Rogers, 1996; Wexler et al., 2001b). A different question concerns the perception of stability: detecting whether an object moves relative to a world-based reference frame. In an early study addressing this issue, Wallach et al. (1974) yoked the movement of an object to the movement of the observer. Their experimental set-up was an ambitiously crafted mechanical apparatus that connected a helmet worn by the observer to the target object via a variable ratio gear mechanism, allowing the experimenter to vary the rotational gain of the target. Thus, with a gain of 1, the target object rotated so as to always present the same face to the observer, with a gain of -1 it rotated by an equal and opposite amount and with a gain of zero the ball remained stationary. They found that a gain of ± 0.25 was often tolerated before observers reported that the object had moved.

More recently, Wexler and colleagues have also used the technique of yoking an object's movement to the observer's movement to study the perception of stability. Wexler (2003) varied the gain with which an object translated as the observer moved towards it. Subjects judged whether the object moved in the same or opposite direction to their head movement. Wexler was primarily interested in the difference in perception produced by active or passive movement of the observer. Overall in these papers, Wexler and colleagues have shown that active movement alters observers' perceptions by resolving ambiguities that are inherent in the optic flow presented to them (Wexler et al., 2001a,b; van Boxtel et al., 2003; Wexler,

2003).

Rather than vary the proprioceptive information about observer movement, we have, like Wallach et al. (1974), examined the role of visual information in determining the perception of an object as static in the world. We have expanded their original experiment using an immersive virtual reality system. The advantages of our apparatus are that (i) the observer has more freedom to move as they would when exploring a scene naturally, (ii) we have greater flexibility to yoke movements of the target object to certain components of the observer's movement and not others and (iii) we can manipulate different aspects of the virtual environment. In this case we varied the number of other objects presented around the target and the viewing distance of the target. We find, unlike Wallach et al. (1974), that presenting the yoked target object within a static visual environment can have a profound effect on observer's perceptions of the target. We also show that the biases in observers' responses when the target is presented alone do not fit the pattern of biases found in other tasks where distance is misperceived. We discuss possible alternative explanations.

Methods

Subjects

All subjects had normal visual acuity without correction. With the exception of one (LT, author) subjects were naïve to the purposes of the experiment. Three subjects (JCM, JDS and PHF) had not taken part in psychophysical experiments before.

Equipment

The virtual reality system consisted of a head mounted display, a head tracker and computer to generate appropriate binocular images given the location and pose of the head. The Datavisor 80 (nVision Industries Inc, Gaithersburg, Maryland) head mounted display unit presents separate 1280 by 512 pixel images to each eye using CRT displays. In our experiments, each eye's image was $72°$ horizontally by $60°$ vertically with a binocular overlap of $44°$, giving a total horizontal field of view of $112°$ (horizontal pixel size 3.4 arc min). The DV80 has a see-through mode that allows the displayed image to be compared to markers in the real world using a half-silvered mirror. This permits calibration of the geometry of the display for each eye. In the experiment, the head mounted display was sealed, excluding light from the outside.

The location and pose of the head was tracked using an IS900 system (Intersense Inc, Burlington, Massachusetts). This system combines inertial signals from an accelerometer in the tracker with a position estimate obtained from the time of flight of ultrasound signals. Four ultrasound receivers are attached to the tracker, while more than 50 ultrasound emitters placed around the room send out a timed 40kHz pulse sequence. The data is combined by the Intersense software to provide a six degrees of freedom estimate of the tracker pose and location. This data is polled at 60Hz by the image generation program. Knowing the offset of the tracker from the optic centres of each eye, the position and pose of the head tracker allow the 3D location of the two optic centres to be computed. These are used to compute appropriate images for each eye. Binocular images were rendered using a Silicon Graphics Onyx 3200 at 60 Hz.

Stimulus and task

In the virtual scene, observers viewed a normal sized football (22cm diameter, see figure 1a) at a viewing distance of approximately 1.5m from the observer's starting position and 1.4m above the floor. Observers were prevented from approaching the target by a real table placed between them and the location of the target (but not visible in the virtual scene). The table was approximately 2m wide and guided their movement, $\pm$1m from side to side. Lateral movement beyond this range caused the target ball to disappear. Observers were permitted to walk back and forth as many times as they wished but, after the first few trials, observers normally did so only once before making their response. Subjects were instructed to fixate the ball as they walked.

The target rotated about a vertical axis as the observer moved (its centre point remained fixed). The amount of rotation was linked to the observer's movement by different gains on each trial. When the gain was +1 the ball rotated so as to always present the same face to the observer (see figure 2), i.e. the ball moves 'with' the observer. A gain of -1 would give rise to an equal and opposite angular rotation, i.e. the ball moves 'against' the observer. When the gain was zero the ball remained static. Any vertical movement of the observer had no effect on the ball's rotation. The rotation of the ball depended only on the component of the observer's movement in a lateral direction, as shown by the arrows in figure 2.

The subject's task was to judge whether the ball moved 'with' or 'against' them as they moved. No feedback was provided. After the subject indicated their response, by pressing one of two buttons on a hand-held 'wand', the target football disappeared. It reappeared after a delay of 500 ms. The surface of the ball had

no specular component since this would enable subjects to detect movement of the ball by judging the motion of features on the ball relative to the specular highlight.

[Figure 1 about here.]

Psychometric procedure

The first run or block of trials in each condition tested gains between -0.5 and 0.5 in increments of 0.1, i.e. 11 different gain values. A run consisted of 55 trials, with each gain value tested 5 times. The range of the next run was determined by the observer's bias on the previous runs, following a semi-adaptive procedure for deciding the range of gain values to be tested during a run (Andrews et al., 2001). At least eight runs were performed for each scene so that the minimum number of trials per psychometric function was 440. For each scene, the observer's responses were plotted against the rotation gain of the target. A cumulative Gaussian curve was fitted to the data using probit analysis (Finney, 1971) to obtain the bias (or point of subjective equality) and the threshold (standard deviation of the fitted Gaussian). Error bars shown on the psychometric functions (figures 2 and 4) show the standard deviation of the binomial distribution.

Figure 2 shows an example psychometric function. It illustrates the bias, i.e. the rotation gain at which subjects perceive the ball to move 'with' them on 50% of trials. This indicates the gain at which subjects perceive the ball as stationary. Error bars on the histograms of bias and thresholds show 95% confidence limits of these values from the probit fit.

[Figure 2 about here.]

Experiment 1: Detecting the movement of a yoked target

We examined the claim by Wallach et al. (1974) that the addition of a static environment around a yoked target did not affect subjects' perception of stability of the target. The static stimulus they used was a background of vertical stripes 40 cm behind the yoked target object. However, Wallach et al. (1974) did not use a forced choice procedure, measure psychometric functions or present data on individual subjects. We measured the bias and threshold of observers' responses when judging the direction of rotation of a yoked ball when the ball was presented alone or in the presence of a surrounding static scene. The static scene consisted of a virtual room with walls at least 1m from the ball and two other footballs close to the target ball (see 'rich cue' scene in figure 1d). For one observer, we measured performance for intermediate scenes, with static objects close to or distant from the yoked target (figure 1b and c).

[Figure 3 about here.]

Figure 3a shows the biases and thresholds for observers in the 'target alone' and 'rich cue' conditions. All four subjects show a large positive bias when the ball was presented alone That is, when subjects perceived the ball to be stationary it was in fact rotating to face them with a gain of 25-45%. Wallach et al. (1974) did not report this result, although their data is consistent with a small positive bias. Our data is also consistent with the direction and magnitude of bias found in a related experiment by Wexler (2003) (see discussion of experiment 2 and figure 8). Note that we did not find the large positive bias in all subjects. In a separate experiment manipulating viewing distance (experiment 2, figure 7) one subject had

biases close to zero.

It is clear from figure 3a that a static background can have a dramatic effect on subjects' perception of stability. For all four subjects, biases for the 'rich cue' condition are around 5%, far lower than when the target is presented alone. This is perhaps not surprising, given that subjects can use the relative motion between the target and static objects as a cue, but it is contrary to the conclusion of Wallach et al. (1974). Figure 3b shows the corresponding thresholds for these judgements. As in the case of biases, for all four subjects thresholds (shown by the histogram bars) are better in the 'rich cue' than the 'target alone' condition. However, the effect of the stable visual background on thresholds is considerably smaller than for biases.

Different measures of threshold

The histogram bars and the diamonds in figure 3b show thresholds for the same data calculated by different methods, as follows. The first method is to fit a cumulative Gaussian to the entire data set for one condition (440 trials, see figure 4a). These thresholds are shown as bars in figure 3). The second method is to fit a cumulative Gaussian to the data for each individual run of 55 trials and calculate the average (root mean square) of the thresholds for all 8 runs. These thresholds are shown as diamonds. If there is a significant variation in the bias for different runs then the threshold according to the first method can be substantially larger than the second. This was the case in the 'target alone' condition for more than one subject, most dramatically subject JDS. Figure 4a illustrates the point for subject LT. As the inset to shows, there was a systematic drift in the bias to progressively larger

values across runs causing the averaged data to have a shallower slope than any of the individual runs. Note that this systematic drift in bias was due in part to the fact that the range of cues presented was varied according to the subject's responses on previous runs (see Methods). For all four subjects, the root mean square thresholds (diamonds) are lower than the thresholds obtained from the combined data (bars) in the 'target alone' condition, consistent with drifting biases in each case. There is less difference between the different threshold measures for the 'rich cue' condition, presumably because subjects used the relative motion between target and static background to help make their judgements.

[Figure 4 about here.]

Intermediate conditions

[Figure 5 about here.]

Figure 5 shows data for one subject for the conditions shown in figure 1b and 1c which provide more information than the 'target alone' but less than the 'rich cue' condition. When the reference football is adjacent to the target the bias is close to zero. This suggests that the most important components of the static environment shown in the 'rich cue' condition are likely to be the objects close to the target, as would be expected if relative motion is an important cue.

Limits on performance in the task

Wallach et al. (1974) described the range over which subjects perceived no rotation as 'large'. It is not clear quite what this means. It is possible to ask whether

observers' thresholds in this task are congruent with thresholds that would be measured for an equivalent visual task in which the observer does not move and is not asked to make a judgement about the allocentric movement of the object. The task in the walking experiment relies on observers making a speed discrimination judgement. We measured thresholds when this was the only element of the task.

Specifically, subjects were seated while wearing the head mounted display and saw a football presented alone at a viewing distance of 1.5m, as in condition figure 1a. On each trial they saw the images corresponding to the view of an observer moving along a circular path centred on the football and always facing towards it. The simulated observer's speed varied according to a cosine function, slowing down at either extremity of the path. The amplitude of the trajectory was $\pm 45°$ around the circle. In fact, both subjects perceived the ball to rotate about a vertical axis rather than perceiving themselves to be moving around the ball. The simulated observer's location was static for 1s at the beginning of the trial, it moved through a single oscillation lasting 3s and then the ball disappeared until the subject made their response, triggering the next trial. The gain of the football's rotation varied, as in the walking experiment, from -0.5 to 0.5 initially and was, as in the walking experiment, changed in accordance to the subject's bias in the preceding runs. Consequently, for subject LT, the gain varied from -0.8 to 0.5, corresponding to a maximum retinal speed of 4.3°/s and 1.2°/s respectively. The subject's task was to determine whether the rotation speed of the ball was greater or less than the mean rotation speed across trials (e.g. McKee et al., 1990). In a run of 130 trials the responses from the first 20 trials were discarded. Judgements of the mean cue across trials have been shown to lead to thresholds that are very similar to those when the

standard is shown in every trial (Morgan et al., 2000)).

The root mean square threshold rotation gain for subject LT was 0.162. As can be seen from figure 5 (asterisk), this is higher than her threshold for the 'target alone' condition in the walking experiment. A second subject, AG, had a slightly better threshold rotation gain of 0.125, which is within the range of root mean square thresholds of all four subjects tested in the walking experiment (figure 3).

This control condition demonstrates that the thresholds for 'constancy of object orientation' (Wallach's phrase) are not radically higher than one might expect purely from the low level visual demands of the task. In fact, the one subject who did both experiments showed lower thresholds in the walking condition. A possible explanation for this improvement is that the allocentric judgement of 'with' or 'against' rotations provided a more stable perceptual criterion for this subject than the judgement of mean speed. Another is that this subject had more practice at the walking task than the mean speed judgement. A more thorough study would be required to make any firm statement beyond the observation that thresholds in the walking and static conditions are broadly similar.

Experiment 2: Mis-estimation of viewing distance or distance walked?

In experiment 1, when the target football was presented alone, all observers perceived it to be stationary when it was actually rotating with them. In this experiment, we explored possible reasons for the bias, namely that subjects perceive the ball to be further away than it really is or under-estimate the distance that they have walked. Figure 6 illustrates how each of these mis-estimates could give rise to a

positive bias.

[Figure 6 about here.]

If the cause of the bias is a mis-estimation of viewing distance, then one would expect the size of the bias to vary with viewing distance in a way that is consistent with previous experiments (e.g. Ogle, 1950; Foley, 1980; Johnston, 1991; Cumming et al., 1991). These studies have generally found that close distances are over-estimated and far distances are under-estimated with a distance between the two, sometimes called the 'abathic distance', at which viewing distances are estimated correctly. The method is often indirect, so that the judgement the observer makes is one of shape rather than distance. Ogle (1950) found that the shape of an apparently fronto-parallel plane was convex at distances closer than about 5m and was concave at distances greater than this, corresponding to an overestimate of near and an underestimate of far distances. Johnston (1991) found a similar result using a different shape judgement, although in this case the 'abathic distance' was approximately 1m. There is other evidence that the absolute value of the abathic distance varies with the subject's task (see review by Foley, 1980). However, all these cases can be interpreted as showing a 'compression of visual space', i.e. an overestimate of near and an underestimate of far distances. We repeated the 'target alone' condition of the previous experiment at different distances to see whether the same type of distortion of space could explain our results.

Figure 7a shows the biases for viewing distances of 0.75, 1.5 and 3m. The definition of a gain of 1 remained that the ball rotated to face the observer. Thus, for a given distance walked, a gain of 1 corresponds to a smaller rotation of the ball when the viewing distance is large. The biases for subjects LT and JCM are

large and positive at all three viewing distances, as in experiment 1. (Their data

for 1.5m is re-plotted from figure 3a, the 'target alone' condition in experiment

1.) By contrast, the biases for subject AJMF are close to zero. (Note that of the

five subjects to do the 'target alone' experiment, he is the only one to show small

biases: figures 3a and 7a.) Despite this variability between subjects, figure 8 shows

how all the data can be used to assess the hypotheses described above. It shows

the biases in figure 7a converted into estimated viewing distance and estimated

distance walked, calculated as described below.

[Figure 7 about here.]

Converting biases into distance estimates

The biases in figure 7a can be converted to an estimated viewing distance if one

assumes that the observer estimates correctly the distance that they walk. Estimated

viewing distance, D', is given by:

$$D' = \frac{D \tan \theta}{\tan((1 - \beta)\theta)} \tag{1}$$

where D is the true viewing distance and β is the bias. If x is the lateral distance

walked from the starting position (1m in our experiment), then $\theta = \arctan(x/D)$

(see Appendix for details). Estimated viewing distance, D', is plotted against real

viewing distance, D, in figure 8a. It is possible that observers make their judge-

ments on the basis of a small head movement rather than the whole ± 1m excursion

in which case small angle approximations apply and the estimated viewing dis-

tance is given by $D' \approx D/(1 - \beta)$. As figure 8a shows (square symbols), these

values are only slightly different. The estimated viewing distances for AJMF are

close to veridical, corresponding to the small biases shown in figure 7a. The estimated viewing distances for observers LT and JCM, on the other hand, are greater than the true viewing distance and increase with increasing viewing distance. This trend is the reverse of that expected from previous experiments (Ogle, 1950; Foley, 1980; Johnston, 1991).

[Figure 8 about here.]

The second possible explanation of biases in the 'target alone' conditions is that subjects mis-estimate the distance they have walked. The ratio of estimated lateral distance walked, x', to real distance walked, x, is given by:

$$\frac{x'}{x} = \frac{\tan((1-\beta)\theta)}{\tan\theta} \qquad (2)$$

as shown in the Appendix. As before, if subjects use information from only a short head movement then the equation can be simplified. Here, $x'/x \approx (1-\beta)$, shown by the squares in figure 8b.

Figure 8b shows estimated distance walked, x'/x, for the three conditions tested in experiment 2. Although there are differences between subjects, for each subject the extent to which distance walked is mis-estimated is almost constant across viewing distance. This makes mis-estimation of distance walked a plausible explanation of the biases. It has the advantage over the viewing distance hypothesis that it does not contradict results of earlier experiments.

In fact, the conclusion that subjects mis-estimate distance walked is consistent with data from a related experiment (Wexler, 2003). For the purposes of comparison, we have plotted the estimate of distance walked derived from Wexler's experiment in figure 8b (arrow). In Wexler's experiment, as in ours, observers judged

whether an object moved 'with' or 'against' them as they moved, although in their case the target object translated rather than rotated. The target was presented alone and, as in our experiment, subjects displayed a large bias. Wexler (2003) found that for the conditions in which observers moved their head the underestimation of distance moved was about 40%, very close to the values found for subject LT and JCM.

Discussion

Using an immersive virtual reality system, we measured the ability of observers to detect that an object rotated when its movement was yoked to the observer's own translation. The experiment extends observations made by Wallach et al. (1974), who arranged for their observers to be linked to the rotating object by a physical mechanism.

Unlike Wallach et al. (1974), we found that the visual context of the yoked object can have a significant effect on the accuracy (bias) of observers' judgements. Four out of the five subjects we tested had large biases when the yoked object was presented alone (see figures 3 and 7) such that the object was perceived to be static when it in fact rotated to follow the observer as they walked. Presenting the object in the context of a static scene dramatically reduced the bias. The effect of the background was particularly strong when some of the static objects were presented close to the yoked object (figure 4), as one would expect if relative motion between different objects is an important cue.

We have considered two possible causes of the large biases that occur when the object is presented alone. One is an over-estimation of the distance of the

target. The other is an under-estimation of the distance walked (see figure 6). Our results do not fit with previous results on a mis-estimation of distance which have shown, using a variety of tasks, that subjects tend to over-estimate near distances and under-estimate far distances, with a cross-over point at some 'abathic' distance (e.g. Foley, 1980; Johnston, 1991; Glennerster et al., 1996). In order to explain the biases in our experiment within the same framework, one would have to postulate a quite different pattern of estimated viewing distances: an expansion of visual space rather than a compression around an abathic distance (see figure 8a).

On the other hand, the explanation that subjects mis-estimate the distance that they walk fits both our data and that from a previous experiment (Wexler, 2003). For each subject in our experiment, the extent to which distance walked was mis-estimated remained constant across viewing distances (figure 7b). For two subjects, the mis-estimation was a factor of about 40%, very similar to the mean value found for subjects in Wexler's study.

Finally, we have shown that thresholds for observers making a judgement about rotation relative to an allocentric (world-based) frame when walking round an object were of a similar magnitude to those measured for static observers judging the mean speed of a rotating object. Apart from the fact that the observer walked in one experiment and remained stationary in the other, the experiments were similar. The result does not support the characterisation by Wallach et al. (1974) that the visual system has only a 'crude' mechanism of for perceiving object orientation as constant. Of course, thresholds for detecting that the ball moved are very much greater than if the observer had remained static. However, using an object-centred measure of thresholds rather than a retina-centred measure would be mis-leading,

as has been pointed out in other contexts (e.g. Eagle and Blake, 1995).

Conclusion

The ability of observers to detect that an object is rotating when its movements are correlated with the observer's movements is improved, particularly in the accuracy of judgements, by adding a stable visual reference. The most likely explanation for the bias in observers' performance when there is no stable visual reference is that they under-estimate the distance they have walked.

Acknowledgements

Appendix

Here we give the derivations of equations 1 and 2 for 'estimated viewing distance' and 'estimated distance walked' plotted in figure 8.

The estimated viewing distance of the yoked object can be calculated from the bias (the rotation gain at which the ball was perceived to be stationary) if it is assumed that the subject estimates correctly the distance they have walked. As can be seen from figure A1a, the distance walked, x, is given by

$$x = D \tan \theta \tag{A1}$$

where D is the real distance of the target ball at the start of the trial and θ is the

angle between the line of sight to the ball at the start of the trial and the line of sight after walking laterally by distance x. When the ball rotates with a certain rotation gain, g, the angle through which the line of sight moves *relative to the ball* is $\theta(1-g)$. For example, for a gain of 1 the view does not change; for a gain of 0 the line of sight moves through an angle θ. Let β be the bias, i.e. the gain at which observers perceive the ball to be stationary. As figure A1a illustrates, if observers perceive the ball to be at distance D' and they perceive the distance they walk correctly as x, then:

$$x = D' \tan((1-\beta)\theta). \tag{A2}$$

Equation 1 follows from equations A1 and A2.

Similarly, the estimated distance walked can be calculated from the bias if it is assumed that the subject estimates the viewing distance of the target correctly. As can be seen from figure A1b, the distance of the object D, is given by

$$D = \frac{x}{\tan\theta} = \frac{x'}{\tan((1-\beta)\theta)} \tag{A3}$$

from which equation 2 follows.

[Figure A1 about here.]

References

Andrews, T. J., Glennerster, A., and Parker, A. J. (2001). Stereoacuity thresholds in the presence of a reference surface. *Vision Research*, 41:3051–3061.

Bradshaw, M. F. and Rogers, B. J. (1996). The interaction of binocular disparity and motion parallax in the computation of depth. *Vision Research*, 36:3457–3768.

Bridgeman, B., der Heijden, A. H. C. V., and Velichkovsky, B. M. (1994). A theory of visual stability across saccadic eye movements. *Behavioural and Brain Sciences*, 17:247–292.

Burr, D. (2004). Eye movements: Keeping vision stable. *Current Biology*, 14:R195–R197.

Cumming, B. G., Johnston, E. B., and Parker, A. J. (1991). Vertical disparities and perception of 3-dimensional shape. *Nature*, 349:411–413.

Duhamel, J. R., Bremmer, F., BenHamed, S., and Graf, W. (1997). Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature*, 389:845–848.

Eagle, R. A. and Blake, A. (1995). Two-dimensional constraints on three-dimensional structure from motion tasks. *Vision Research*, 35:2927–2941.

Feldman, J. A. (1985). Four frames suffice: A provisional model of vision and space. *Behavioural and Brain Sciences*, 8:265–313.

Finney, D. (1971). *Probit Analysis*. Cambridge University Press, Cambridge, 3rd edition.

Foley, J. (1980). Binocular distance perception. *Psychological Review*, 87:411–433.

Franz, M. O. and Mallot, H. A. (2000). Biomimetic robot navigation. *Robotics and Autonomous Systems*, 30:133–153.

Glennerster, A., Hansard, M. E., and Fitzgibbon, A. W. (2001). Fixation could simplify, not complicate, the interpretation of retinal flow. *Vision Research*, 41:815–834.

Glennerster, A., Rogers, B. J., and Bradshaw, M. F. (1996). Stereoscopic depth constancy depends on the subject's task. *Vision Research*, 36:3441–3456.

Johnston, E. (1991). Sytematic distortions of shape from stereopsis. *Vision Research*, 31:1351–1360.

McKee, S., Levi, D., and Bowne, S. (1990). The imprecision of stereopsis. *Vision Research*, 30:1763–1779.

Morgan, M. J., Watamaniuk, S. N., and McKee, S. P. (2000). The use of an implicit standard for measuring discrimination thresholds. *Vision Research*, 40:2341–2349.

Ogle, K. (1950). *Researches in binocular vision*. Philadephia and London: W.B. Saunders Comp., 1st edition.

Rogers, B. J. and Graham, M. (1982). Similarities between motion parallax and stereopsis in human depth perception. *Vision Research*, 22:261–270.

Ross, J., Morrone, M. C., Goldberg, M. E., and Burr, D. C. (2001). Changes in visual perception at the time of saccades. *Trends in Neuroscience*, 24:113–121.

van Boxtel, J. A., Wexler, M., and Droulez, J. (2003). Perception of plane orientation from self-generated and passively observed optic flow. *Journal of Vision*, 3:318–332.

Wallach, H., Stanton, L., and Becker, D. (1974). The compensation for movement-produced changes of object orientation. *Perception and Psychophysics*, 15:339–343.

Wexler, M. (2003). Voluntary head movement and allocentric perception of space. *Psychological Science*, 14:340–346.

Wexler, M., Lamouret, I., and Droulez, J. (2001a). The stationarity hypothesis: an allocentric criterion in visual perception. *Vision Research*, 41:3023–3037.

Wexler, M., Panerai, F., Lamouret, I., and Droulez, J. (2001b). Self-motion and the perception of stationary objects. *Nature*, 409:85–88.
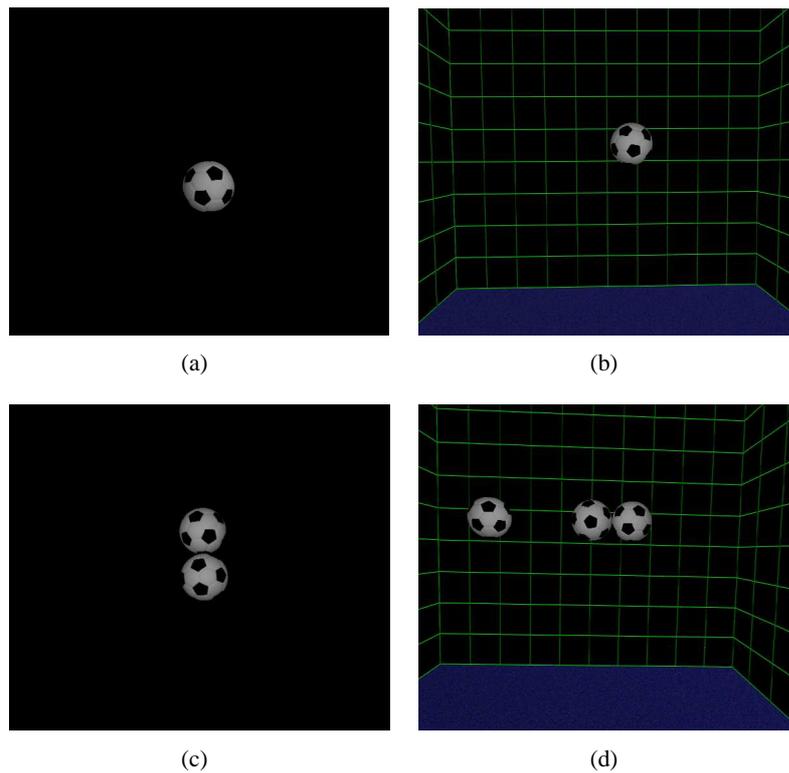
List of Figures

Figure 1: **Different backgrounds.** The central, target football rotated around a vertical axis as the observer moved. The gain of this yoked movement varied from trial to trial, as described in the text. Observers judged the direction of rotation, as 'with' or 'against' them, as illustrated in figure 2. The walls of the room and the other footballs shown in (b), (c) and (d) were static throughout. Results for the 'target alone' condition, (a), and the 'rich cue' scene, (d), are shown in figure 3. Results when the target was presented just with a background, (b), and when a static reference was presented adjacent to the target, (c), are shown in figure 5.
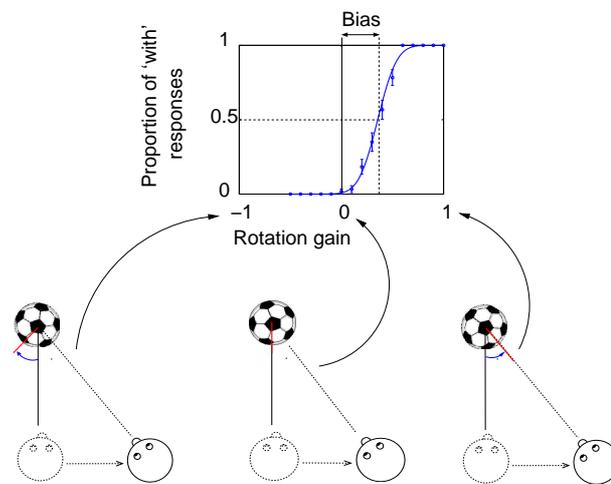
Figure 2: **An example psychometric function.** The proportion of trials on which the subject responded that the target ball had moved 'with' them as they moved is plotted against the rotation gain of the ball. As shown in the diagrams below, when the gain is 1 the ball rotates so as to always face the observer, when it is -1 the ball has an equal and opposite rotation and when the gain is 0 the ball remains static. The red line is included in order to illustrate the rotation of the ball. The bias, or shift in the 50% point, indicates the rotation gain at which the subject perceived the ball to be stationary as they moved. The threshold is the standard deviation of the fitted cumulative Gaussian.
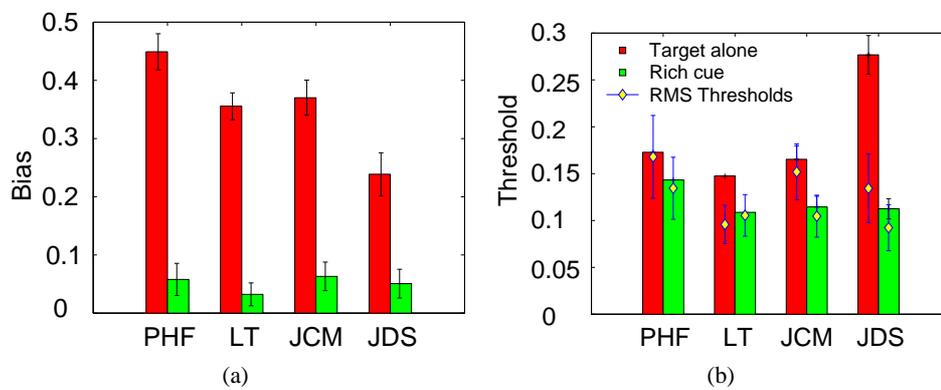
Figure 3: **Results from experiment 1.** Biases (a) and thresholds (b) are shown for four observers when the yoked target football was presented alone (red) or in the 'rich cue' environment (green) which consisted of a static background and adjacent static objects (see figure 1a and d). Biases and thresholds are given as rotation gains (see figure 2) and hence have no units. The white diamonds show the average (root mean square) threshold for individual runs (see text and figure 4 for explanation).
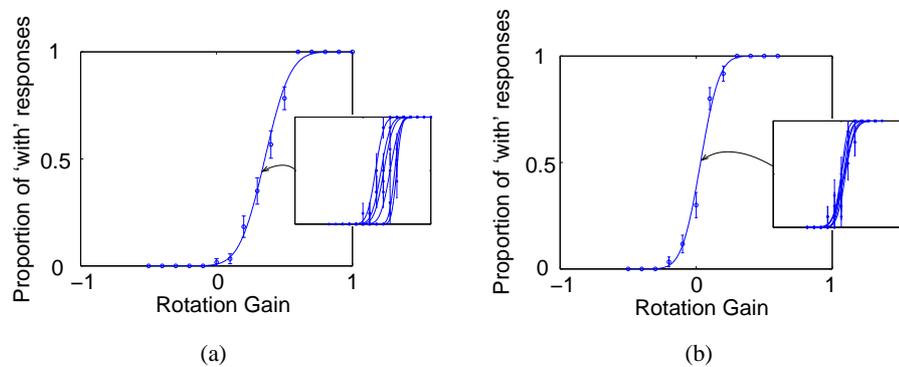
Figure 4: **Thresholds raised by a drifting bias.** Psychometric functions for the data of subject LT shown in figure 3b for (a) the 'target alone' condition and (b) the 'rich cue' condition. These illustrate how thresholds for individual runs of 55 trials were similar across runs in both conditions (see insets). The values of the root mean square thresholds for the two conditions are plotted in figure 3b (subject LT). However, because the bias has drifted between runs in the 'target alone' condition, the psychometric function for the combined data in (a) has a shallower slope, i.e. a higher threshold, than in (b).
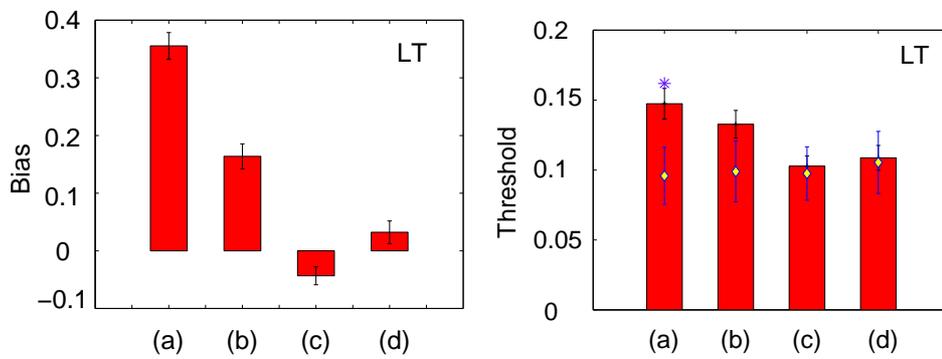
Figure 5: **Other backgrounds.** Bias and thresholds in experiment 1 for one subject using all four types of background shown in figure 1. The labels (a) to (d) correspond to the labels of the conditions illustrated in figure 1. Data for the 'target alone', (a), and 'rich cue' conditions, (d), are re-plotted from figure 3. (b) shows data for the target football with a static room. (c) shows data for the target with an adjacent static football. The white diamonds show average thresholds across individual runs, as in figure 3. The asterisk (column (a), thresholds) shows data for a control condition in which the subject remained static, as described in the text.
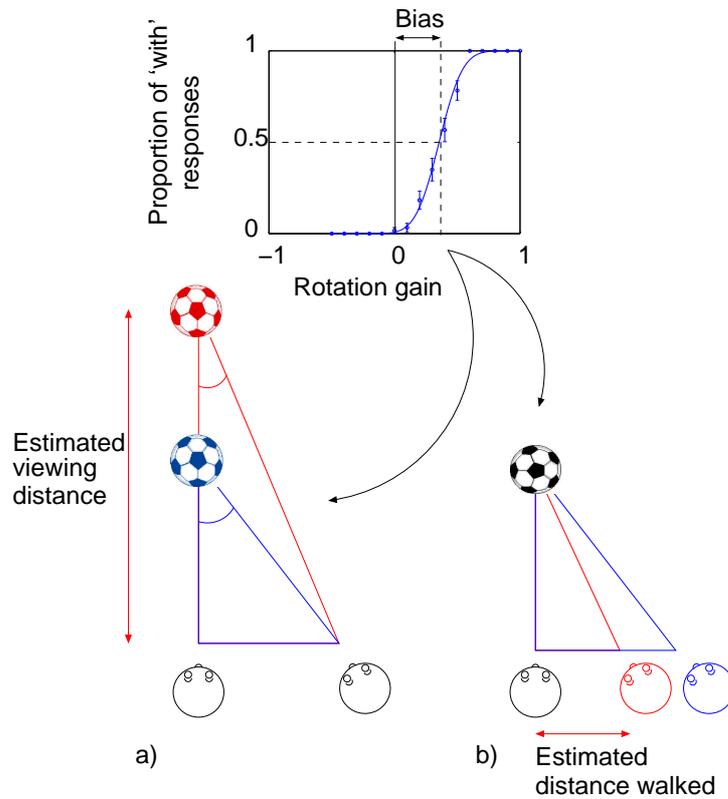
Figure 6: **Possible causes of bias in the 'target alone' condition.** A positive bias can be attributed to either a) an overestimate of distance to the target or b) an underestimate of the distance walked. In either case, the subject expects their view of the ball to change by a smaller amount than would be the case if their estimate were correct. Hence, the ball they see as stationary is one that rotates 'with' them.
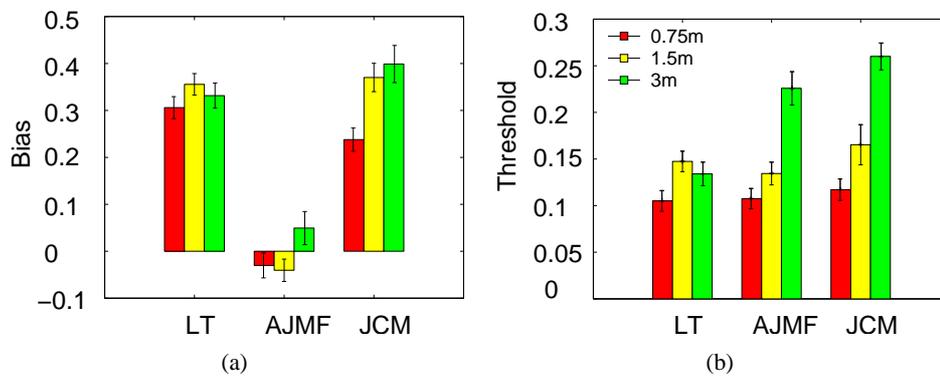
Figure 7: **Results of experiment 2.** The 'target alone' condition from experiment 1 (1.5m) repeated at 0.75 and 3m viewing distances. For subjects LT and JCM, the bias and thresholds for 1.5m viewing distance is re-plotted from figure 3. Figure 8a shows how the bias data here can be related to the two hypotheses illustrated in figure 6.
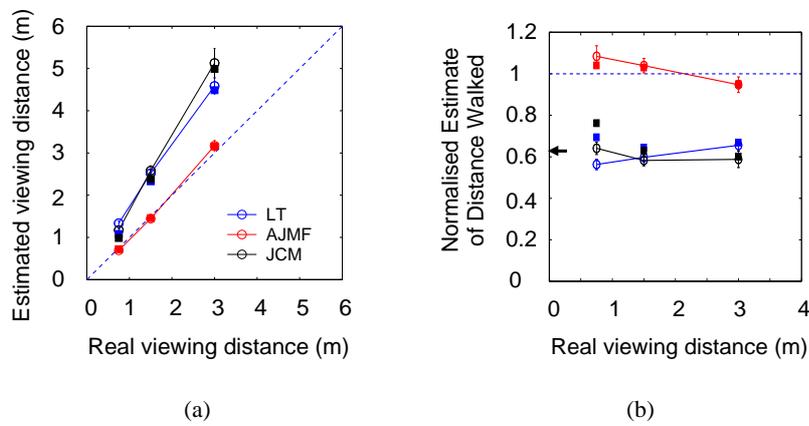
Figure 8: **Tests of the hypotheses in figure 6.** (a) Data from figure 7 re-plotted to show the estimated viewing distance that would account for the rotation perceived as static if other parameters were judged correctly (equation 1). (b) The same data re-plotted but now assuming that distance walked is mis-estimated. The ordinate shows estimated distance walked normalised by the true distance walked (x'/x, equation 2). The squares show, in corresponding colours for each subject, (a) estimated viewing distance and (b) estimated distance walked assuming that subjects make their judgement on the basis of a short translation rather than the ± 1m maximum excursion (see text). The arrow in (b) shows the estimated distance walked derived from a related experiment by Wexler (2003).
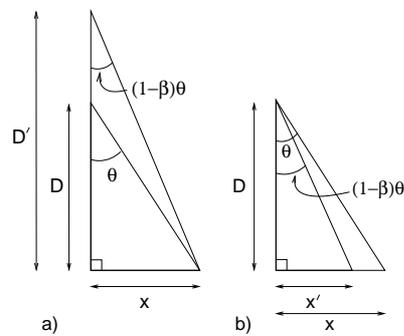
Figure A1: **Assumptions used in calculating estimated viewing distance and distance walked.** D is the real distance of the target ball at the start of the trial, θ is the angle between the line of sight to the ball at the start of the trial and the line of sight after walking laterally by distance x and β is the subject's bias. In (a), D' is the estimated viewing distance of the target assuming x is judged correctly. In (b), x' is the estimated distance walked assuming D is judged correctly. D' and x' are plotted in figure 8 for the conditions tested in experiment 2.