THE ROYAL
SOCIETY

Increasing confidence in vergence as a cue to distance

James R. Tresilian[1], **Mark Mon-Williams**[1,2] **and Benjamin M. Kelly**[3]

[1]*Department of Human Movement Studies and* [3]*Department of Psychology, University of Queensland, St Lucia, Queensland 4072, Australia*
[2]*School of Psychology, University of St Andrews, St Andrews, Fife KY16 9JU, UK*

Multiple cues contribute to the visual perception of an object's distance from the observer. The manner in which the nervous system combines these various cues is of considerable interest. Although it is accepted that image cues play a significant role in distance perception, controversy exists regarding the use of kinaesthetic information about the eyes' state of convergence. We used a perturbation technique to explore the contribution of vergence to visually based distance estimates as a function of both fixation distance and the availability of retinal information. Our results show that the nervous system increases the weighting given to vergence as (i) fixation distance becomes closer; and (ii) the available retinal image cues decrease. We also identified the presence of a strong contraction bias when distance cues were studied in isolation, but we argue that such biases do not suggest that vergence provides an ineffectual signal for near-space perception.

Keywords: distance estimates; fixation distance; kinaesthetic information; perception; visual cues

1. INTRODUCTION

In normal circumstances there are multiple sources of stimulus information about aspects of the environment (such as the distance to an object). In such conditions it is typically found that the perceptual contribution of a cue depends upon the 'confidence' the nervous system attaches to the information: the greater the confidence, the larger a cue's relative contribution (von Holst 1973; Landy *et al.* 1995; Massaro 1988; Zacharias & Young 1977). In addition, the contribution of any given cue will tend to decrease as the number of other contributing cues increases (see, for example, Rogers & Bradshaw 1995). The manner in which the nervous system combines multiple cues has recently received theoretical attention resulting in various accounts of cue combination, such as the *modified weak fusion* scheme of Landy *et al.* (1995). We applied these theoretical ideas to study the role of vergence information in visual distance perception.

Information about the state of the extraocular muscles can, in principle, provide the central nervous system with an estimate of the angle of binocular convergence, from which the radial distance to the point of fixation can be determined. It has been established empirically that vergence information can contribute to distance perception in reduced-cue environments (Foley 1980; Rogers & Bradshaw 1995; Swenson 1932). Nevertheless, there is doubt about the usefulness of vergence-derived distance information in normal (full cue) viewing conditions. Experimental studies in conditions with very reduced cues (i.e. when vergence is the only cue or one of a very few cues) have shown distance perception to be inaccurate: for example, Gogel (1972) found that observers judged point lights at 6 and 3 m to be equidistant. In addition, perceived space is typically distorted, with near targets appearing further away than they actually are and far targets appearing closer, a phenomenon referred to as the *specific distance tendency* (Gogel & Tietz 1973). From these and related observations it has been concluded that vergence information is too imprecise to play a significant role in everyday distance perception (see, for example, Brenner & van Damme 1998; Turvey & Soloman 1984).

It would, however, be premature to conclude that vergence plays a negligible role in distance perception for at least three reasons. First, the geometry of binocular vision implies that the usefulness of vergence as a source of distance information is restricted to near space: it is unlikely to provide any useful information for fixation distances greater than about 3 m (figure 1*a*). Within a range from about 10 cm (limits of near convergence) to 2–3 m, vergence could in principle make a useful contribution to the perception of fixation distance: note that even within this range its accuracy as a cue falls off with increasing fixation distance (figure 1*a*). Second, it has been found that pointing to the perceived location of a target produces more reliable and accurate performance than the verbal report method traditionally used to assess distance perception (Bingham & Pagano 1998). Studies in which observers pointed to targets whose distance was defined only by vergence have generally reported reliable distance estimates within the pointing range (Mon-Williams & Tresilian 1998; Swenson 1932). Third, the distortion of space referred to as the specific distance tendency may be an example of the general tendency of observers to bias their responses towards the mean of a stimulus set, a tendency referred to as a *contraction bias* (Poulton 1981). Thus, the specific distance tendency may

Proc. R. Soc. Lond. B (1999) **266**, 39–44
Received 28 August 1998 *Accepted* 30 September 1998

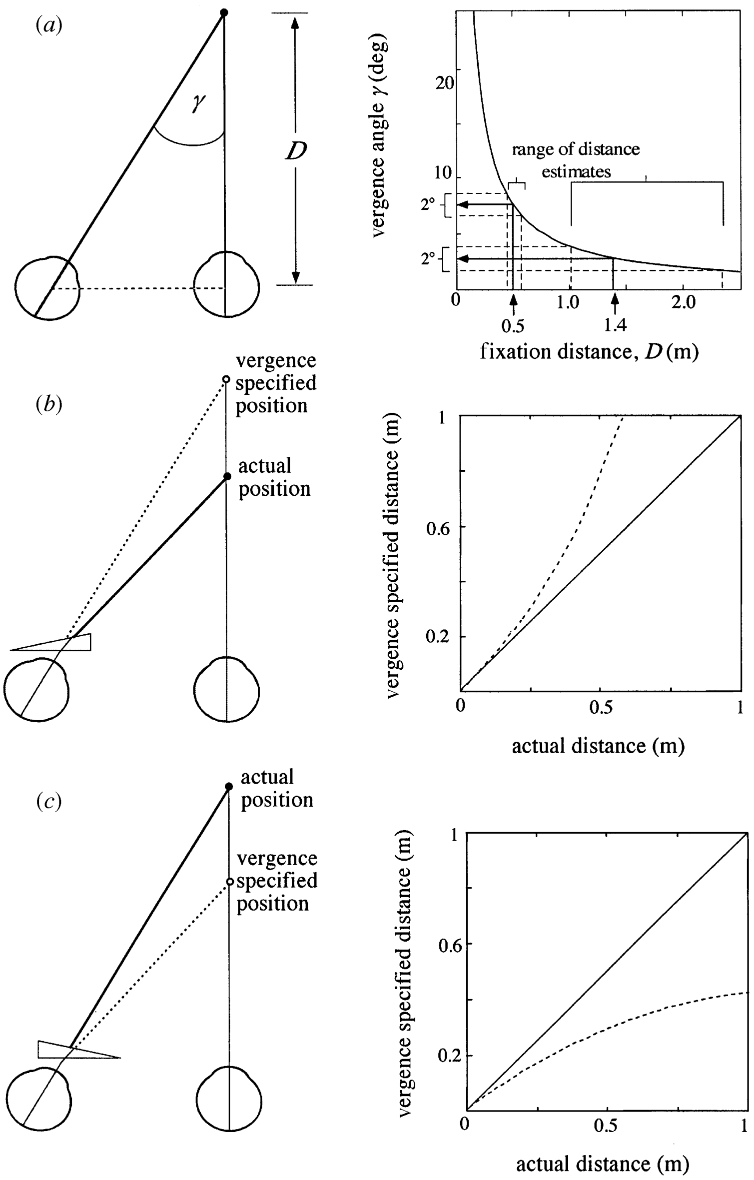

Figure 1. (*a*) Planar geometry of (asymmetric) binocular fixation when the target object (filled circle) is aligned with the axis of the right eye. Left: fixation of a target at distance D from the right eye requires a vergence angle γ. Right: the relation between vergence angle and fixation distance for the geometry shown on the left. Indicated on this graph is the fact that a given range in vergence gives rise to a range in distance which increases rapidly as fixation distance increases. Thus, for a given amount of uncertainty in the measurement of vergence angle, the uncertainty in the derived distance estimate increases with fixation distance. (*b*) Left: a prism base-in in front of the left eye makes the vergence angle required to fixate the target (filled circle) appropriate for a target further away (open circle). Right: the relation between the distance as specified by vergence angle and the actual target position (dashed line; the solid line is the case when no prism is present). (*c*): similar to (*b*) except that the prism is orientated base-out and the vergence is appropriate for a target closer (open circle) than its actual location (filled circle).

not reflect a unique and intrinsic inaccuracy of vergence information.

The study reported here extends previous investigations by attempting to quantify the contribution of vergence to the perception of fixation distance over an extended pointing range (up to 1 m distant) under seven visual cue conditions. The latter ranged from a condition in which no other cues except vergence were present through to a cue-rich environment. Vergence demand was systematically manipulated by placing a 5Δ (prism diopter; $1\Delta = \arctan 0.01$) ophthalmic prism over one eye. The prism could be orientated with its base either inwards towards the nose (base in) or outwards (base out). The prism perturbs the vergence cue while leaving other cues unchanged, and thus represents a variant of the perturbation methodology used by Young *et al.* (1993). The prism has an effect on vergence demand that is independent of target distance, although its effect on vergence-specified distance increases with fixation distance (figure 1*b*,*c*).

From 'weighted averaging' and related cue combination schemes, we predicted that the contribution of vergence should decrease with an increasing amount of information from other sources. We also hypothesized that the contribution of vergence to distance estimates should fall off with increasing target distance, for two reasons. First, it is an established principle that the noisier a cue is, then the lower the confidence placed in it (see, for example, Landy *et al.* 1995; Massaro 1988). Figure 1*a* shows that distance estimates from vergence will become noisier as fixation distance increases so that the contribution of vergence to distance estimates should decrease as fixation distance increases in multiple cue environments. Second, it has been proposed that if a single cue provides information that is at odds with the information provided by several other cues then the relative contribution of the discrepant cue will decrease as the discrepancy increases (Landy *et al.* 1995). Because the prism-induced discrepancy in distance increases with target distance (figure 1*b*,*c*), the contribution of vergence should decrease with increasing target distance. It is important to note (drawn to our attention by a reviewer) that this second reason assumes that the combined cues are expressed in common units of 'distance'. It is possible that the common units of the cue combination stage are not distance units but are represented, for example, as angles (the transformation to distance units may take place after cue combination). The unit of the cue combination stage has implications for the discrepancy introduced by the prism: in distance units the discrepancy increases with target distance but in angular units the discrepancy is constant.

2. METHODS

Forty-two undergraduate students (22 males and 20 females, age range 18–21 years) participated in the experiments for course credit. Participants viewed targets through an aperture (9 cm × 4 cm) in front of a rectangular viewing box (dimensions 130 cm long × 65 cm wide × 21 cm high). A moulded plastic restraint in front of the aperture minimized head movements, occluded peripheral vision and allowed the observers to correctly position themselves. The plastic constraint contained a pair of trial frames (diameter 3 cm) into which an ophthalmic prism could be placed. Eight apertures in the top of the box allowed targets to be aligned with the axis of the right eye (primary position) in 10 cm steps between 30 and 100 cm (± 0.5 cm) from the observer. The experimental task was to position the tip of an unseen stick 96 cm long (dowelling of 16 mm diameter tapered to a 2 mm tip) outside the box at the perceived distance of the target. Before running the experiment, participants were provided with some practise (*ca.* 5 min) pointing without seeing the stick (not in the actual experimental apparatus). The stick was held with the right hand and participants were free to grasp it wherever they chose. It was observed that participants held the stick closer to the pointing tip (30–40 cm) for nearest targets and close to the non-pointing tip for the most distant targets.

We explored the effect of manipulating vergence angle in seven different viewing conditions. In these conditions, target-size information was either present or absent and the viewing environment was either rich, reduced or absent (darkness). The different conditions were: A, target-size information in a cue-rich environment; B, no target-size information in a rich environment; C, target-size information in a reduced-cue environment; D, no target-size information in a reduced-cue environment; E, target-size information in darkness; F, no target-size information in darkness; G, only target size in darkness (monocular viewing through a 1.5 mm pinhole). Six participants were randomly allocated to one of seven groups with each group participating in just one viewing condition.

In conditions A–E the participants viewed the targets as follows: (i) with no prism in place; (ii) through a 5 Δ prism placed base-in in front of the left eye; (iii) through a 5 Δ prism placed base-out in front of the left eye. The observers pointed to each target four times for each condition, with the order of presentation randomized for each participant. The mean of the four points was used within the analyses but we also recorded the mean variable error (standard deviation) for individual participants at each target distance.

In the rich viewing environment the box was well illuminated (*ca.* 500 lux) by an internal light bulb, was fully carpeted on the walls and floor (proving a rich texture gradient), had its far end open (providing a vertical disparity gradient over *ca.* 30°) and contained many familiar objects (providing potential size and ordinal cues to distance). It should be noted that, although this condition was relatively 'rich' with visual cues, the available information was still somewhat reduced compared with normal viewing (for example, motion parallax was minimized and the vertical disparity gradient was relatively small). In the reduced-cue condition the internal surfaces were smooth and painted matt black, the illumination was reduced (*ca.* 250 lux) and no objects apart from the target were visible. In conditions E, F and G the box was light-sealed and the room lights were switched off to ensure that nothing was visible apart from the target. The room lights came on between trials to ensure that participants did not dark-adapt.

Targets were long and narrow, reaching from the top to the bottom of the box (the top and bottom of the target could not be seen by the participants). Target width at each distance was either set so that its horizontal angular subtense was 1.15° at the observer's eye (there was some variation—not correlated with distance—equal to less than 2 arcmin in the target's angular size) or varied systematically with distance (i.e. the same target of width 0.7 cm was used in each position) so that target width could potentially provide information on distance. In conditions A–D, the targets consisted of a single solid black rectangular piece of card; in conditions E and G the target was a luminous piece of tubing (diameter 0.7 cm). In condition F, the target was a pinpoint light source placed at 50 cm and different prism powers were used to create the appropriate vergence-specified distance. Conditions F and G specified the same distances used in conditions A–E but included the additional target distance of 25 cm. The mean positional pointing accuracy was measured for 0.5 s at a sampling rate of 60 Hz by means of an Optotrak 3D optoelectronic movement recording system (accurate to within 0.2 mm).

3. RESULTS

We carried out linear regressions on the group data to plot the relation between pointing response and target position: these results are presented in figure 2. The correlation coefficient was above 0.98 for all regression lines and all passed an ANOVA test for linearity. It may be seen from figure 2 that the effect of the prism was to cause an overshoot when orientated with its base inwards and an undershoot when orientated with its base outwards. We used repeated-measures ANOVA to explore the effect of the prism on the pointing response in conditions A–E. The ANOVA revealed that the prism had a reliable effect on pointing response ($F_{2,8}=132.46$, $p<0.0001$) and that there was a reliable interaction between condition and prism ($F_{2,70}=47.32$, $p<0.0001$). It should be noted that the vergence specified distance in condition F is a function of the interpupillary distance and this was taken into account in the analyses (although mean values are presented in the figure for ease of interpretation). The mean variable error (averaged across participants) is shown plotted as a function of target distance for conditions A–F in figure 3.

A measure of the contribution of vergence to the response is provided by the ratio of the observed pointing error in the presence of the prism to the expected error due to the prism (assuming that vergence completely determines the response). This ratio, termed the 'prism bias ratio' (Landy *et al.* (1995) refer to this as the weight α, where $\alpha=\Delta\text{depth}/\Delta\text{cue}$) gives the observed error as a proportion of the expected error: the larger the ratio (the closer it is to unity) the greater the contribution of vergence. It was calculated separately for base-in and base-out prisms by dividing the change in the pointing response created by the prism (determined by subtracting the baseline pointing response from the prism pointing response)

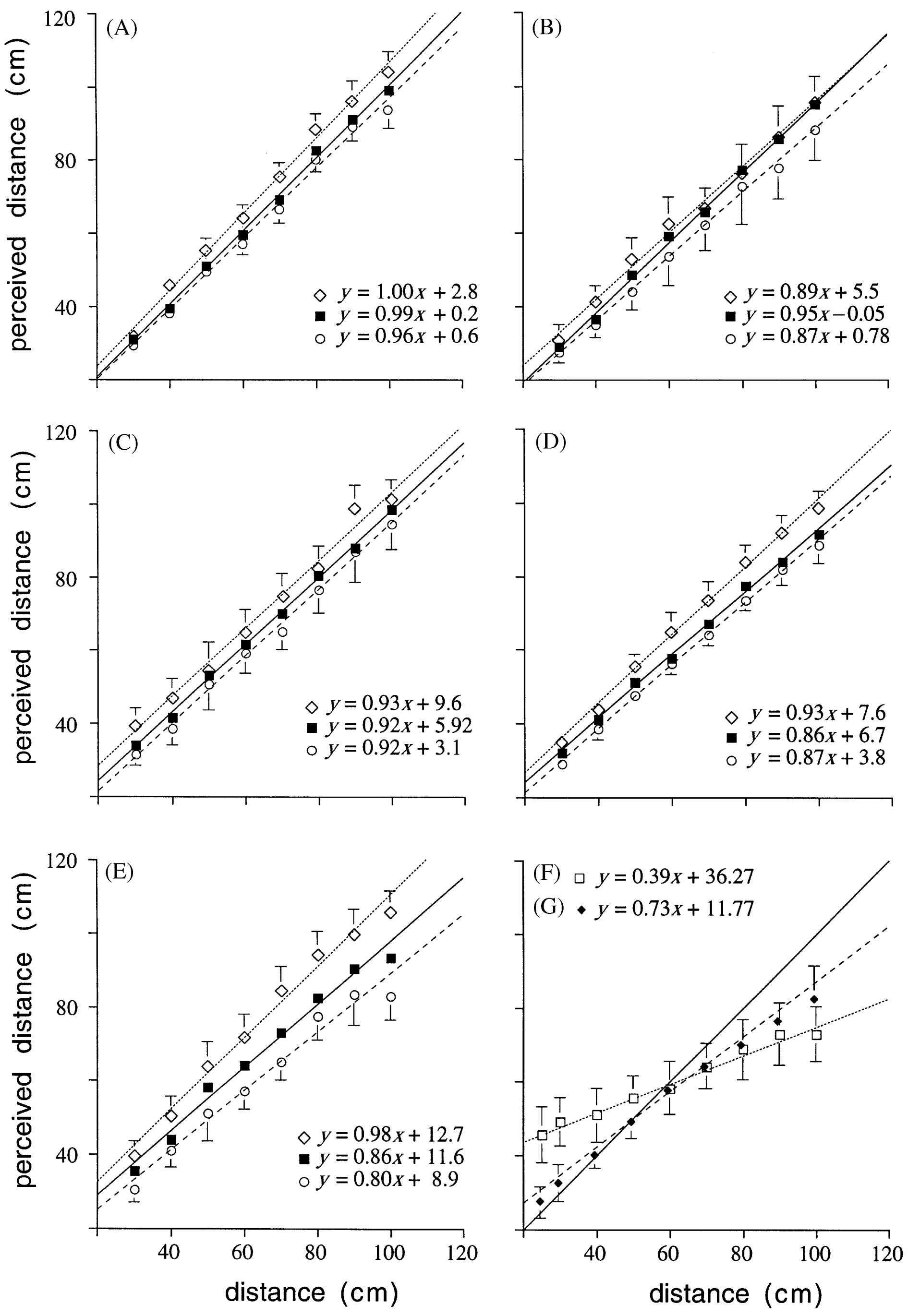

Figure 2. Perceived distance (pointing response) averaged across participants, plotted as a function of target distance from stimulus conditions A–G. Panel labels correspond to stimulus conditions; condition G (size only) is plotted in panel F (filled diamond). The different conditions were as given in the text. Error bars show the standard deviation between subjects and so indicate group variability. Symbols indicate the conditions: open diamond, prism base-in; filled square, no prism; open circle, prism base-out; filled diamond, size only, no prism. Straight lines are least-squares fits to the data by means of linear regression analysis. The equations for each fit are shown (y = perceived distance, x = target distance). The dotted line in panel F is the perceived distance = target distance line, included for comparison.

by the difference between the vergence-specified distance (calculated according to an individual's interpupillary distance; see figure 1*b*,*c*) and the targets' physical location (determined by subtracting the physical position from the geometrically calculated vergence-specified distance). Figure 4 shows the prism bias ratio plotted against vergence angle in radians for conditions A–E with the results of the linear regression analyses shown in the figure.

4. DISCUSSION

A strong contraction bias (specific distance tendency) was observed (figure 2*f*) when vergence was the only distance cue available, as previously reported (Foley 1980; Gogel & Tietz 1973; Mon-Williams & Tresilian 1998). A similar, but slightly less pronounced, contraction bias was observed when size was the only cue (condition G, figure 2*f*, filled diamonds). It may be noted that a similar pattern of results is found when binocular disparity is the only source of information for determining the physical depth of an object: deeper objects appear more contracted than they actually are and shallower objects appear more extended (Johnston 1991; Johnston did not interpret her findings as the result of a contraction bias). A contraction bias appears to be a general feature of distance estimates made on the basis of any cue in isolation. As has been noted previously, cues tend to be inaccurate and ambiguous in isolation but when combined together can provide an accurate perceptual representation (see, for example, Grossberg & Mingolla 1985).

The contraction bias was progressively eliminated as the amount of available information was increased. This is demonstrated by the increasing slope and decreasing intercept of the regression lines from stimulus condition F

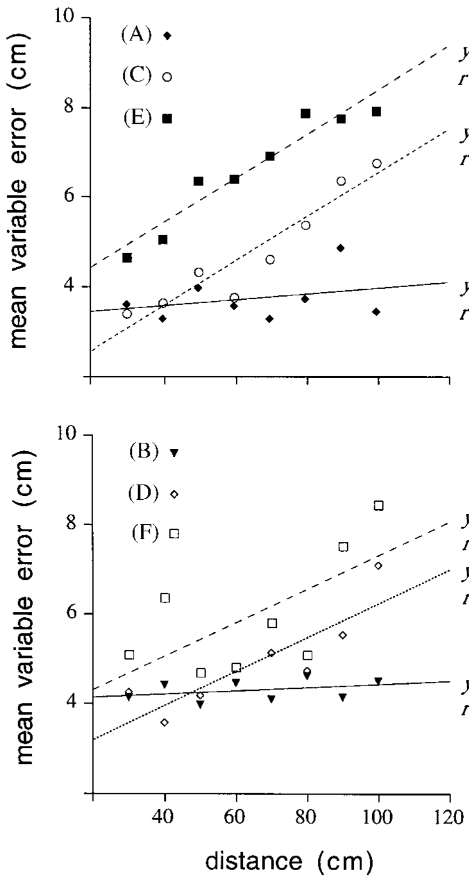

Figure 3. Standard deviation of pointing responses to each target (variable error) averaged across participants, plotted as a function of distance for stimulus conditions A–F (top panel, conditions A, C, E; bottom panel, B, D, F). Symbols denote stimulus conditions as indicated in the key. Straight lines are fits to the data by means of linear regression analysis. The equations for the fits (y = variable error, x = distance) and their r^2 values are shown.

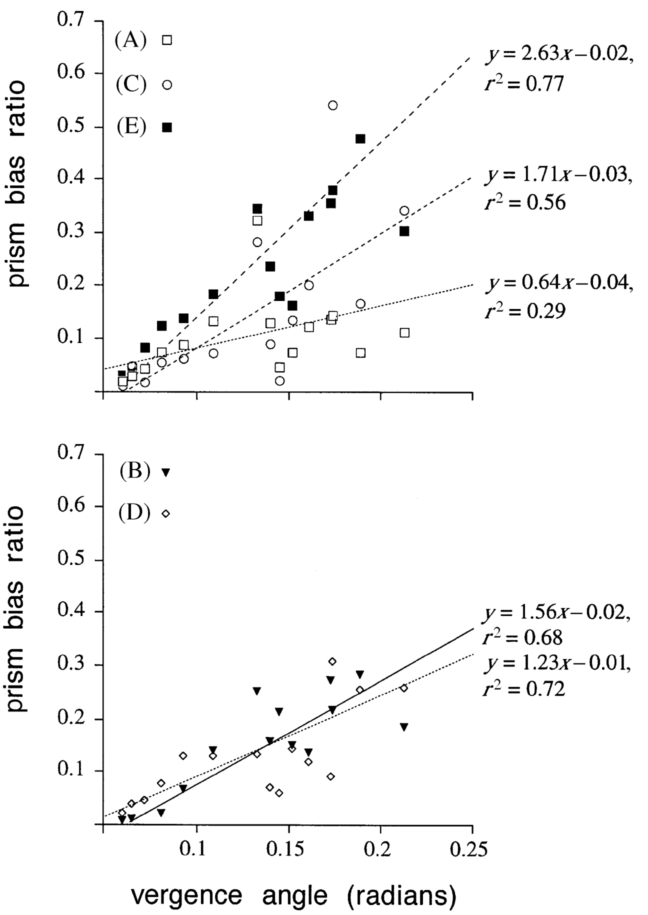

Figure 4. Mean results from the prism conditions (both base-in and base-out) expressed as the prism bias ratio (see text) and plotted as a function of target vergence. Symbols denote the stimulus conditions as shown in the key. Straight lines are fits to the data from each condition by means of linear regression analysis. Equations of the fits (y = prism bias ratio, x = vergence angle) and their r^2 values are shown.

to condition A for the no-prism trials (figure 2, filled square). A slope of unity and a bias of zero corresponds to perfect performance (perceived and actual distance correspond) and this was closely approximated in cue-rich conditions (figure 2*a*). This pattern of results indicates that the contraction bias represents a response of the nervous system to conditions of uncertainty. The fewer the cues that are available, the more uncertainty is likely to be associated with any distance estimate, especially if the available information is perceived to be unreliable. A sensible strategy for minimizing the average error in such circumstances is to bias distance estimates to the perceived centre of the range of possible estimates.

Figure 4 shows that the data were consistent with the two predictions stated in the introduction: the contribution of vergence to the pointing response was smaller for decreased vergence demand (smaller vergence angles) and was generally smaller as the amount of retinal information increased. The prism bias ratio provides an indication of the weight given to vergence in determining the response. The decrease in the weighting attached to vergence with increasing distance is possibly a function of two factors: (i) the decreasing reliability of vergence as a cue; and (ii) the increasing discrepancy between vergence-specified distance (with a prism in place) and the distance as signalled by other available cues. The latter factor assumes that the nervous system conducts cue-combination in distance units. If the combination were conducted in angular units the prism-induced discrepancy would be constant. The present results do not allow us to determine the extent to which either factor was responsible for the observed effects. It should also be noted that there is a problem with using the prism bias ratio as an absolute measure of the contribution of vergence, especially for individual participants. The measure is far more sensitive to response variability at near distances, because then the effect of the prism becomes relatively small (see figure 1*b*,*c*).

The mean variable error data (figure 3) lend some support to the hypothesis that the decrease in the reliability of vergence with increasing target distance contributes to the pattern of results shown in figure 4. Inspection of figure 3 shows that the variability of pointing increases with distance for those conditions in which vergence was predicted to make the larger contribution to distance percepts. As discussed in the introduction, the variability in vergence-based distance estimates will become larger as fixation distance increases (see figure 1*a*). In multiple-cue conditions (A and B) the variability of responding did not change with increasing fixation distance (figure 3). The variable error was largest in those conditions in which vergence was the only or the dominant cue (E and F). Comparison of figures 3 and 4 shows that the pattern of results was similar for the variable error and the prism bias ratio. Of course, it might be expected that response variability would tend to increase as the amount of available information decreases. The important feature of the data presented in figure 3 is that the variability in the different stimulus conditions is much the same (about 4 cm) for near fixation distances. The response variability only becomes larger as fixation distance increases in those conditions in which vergence is making a significant contribution to perceived distance.

The following conclusions concerning the role of vergence in visual perception of distance can be drawn. First, contrary to previous arguments, it does not follow from the specific distance tendency that vergence is too inaccurate a cue to contribute to percepts of visual fixation distance in full-cue environments. The most likely explanation of the specific distance tendency is that it represents a general response of the nervous system to conditions of uncertainty (contraction bias) and would be observed for any distance cue studied in isolation, such as size (condition G) or disparity (Johnston 1991). Second, the contribution of vergence in the prism-perturbed conditions dropped off with increasing target distance as the result of two influences on its weighting: decreasing reliability as fixation distance increases, and increasing discrepancy between vergence-specified distance and distance specified by other cues (assuming distance units are used in the combination process). Third, inaccuracy of vergence-based estimates of fixation distance for fixations beyond 2–3 m is irrelevant for understanding the contribution of vergence to this type of distance perception. Vergence is an unreliable cue at large fixation distances and is subject to a strong contraction bias when studied in isolation, but these facts do not imply a minor role for vergence in near-space perception.

Production of this manuscript was supported by grants from the Australian Research Council; the National Health & Medical Research Council (Australia); the Engineering and Physical Research Council (UK); and Action Research (UK).

REFERENCES

Bingham, G. & Pagano, C. 1998 The necessity of a perception–action approach to definite distance perception: monocular distance perception to guide reaching. *J. Exp. Psych: Hum. Perc. Perf.* **24**, 145–168.

Brenner, E. & van Damme, W. J. M. 1998 Judging distance from ocular convergence. *Vision Res.* **38**, 493–498.

Foley, J. M. 1980 Binocular distance perception. *Psychol. Rev.* **87**, 411–434.

Gogel, W. C. 1972 Space perception with binocular cues to distance. *Am. J. Psychol.* **85**, 477–498.

Gogel, W. C. & Tietz, J. D. 1973 Absolute motion parallax and the specific distance tendency. *Percept. Psychophys*, **13**, 284–292.

Grossberg, S. & Mingolla, E. 1985 Neural dynamics of form perception: boundary completion, illusory figures and neon color spreading. *Psychol. Rev.* **92**, 173–412.

Johnston, E. B. 1991 Systematic distortions of shape from stereopsis. *Vision Res.* **31**, 1351–1360.

Landy, M. S., Maloney, L. T., Johnston, E. T. & Young, M. 1995 Measurement and modeling of depth cue combination: in defence of weak fusion. *Vision Res.* **35**, 389–412.

Massaro, D. W. 1988 Ambiguity in perception and experimentation. *J. Exp. Psychol.: Gen.* **117**, 417–421.

Mon-Williams, M. & Tresilian, J. R. 1998 Some recent studies on the extraocular contribution to distance perception. *Perception.* (In the press.)

Poulton, E. C. 1981 Human manual control. In *Handbook of physiology.* 2. *The nervous system* (ed. V. B. Brooks), pp. 1337–1389. Bethesda, MD: American Physiological Association.

Rogers, B. J. & Bradshaw, M. F. 1995 Disparity scaling and the perception of frontoparallel surfaces. *Perception* **24**, 155–179.

Swenson, H. A. 1932 The relative influence of accommodation and convergence in the judgement of distance. *J. Gen. Psychol.* **7**, 360–380.

Turvey, M. T. & Soloman, J. 1984 Visually perceiving distance. *J. Exp. Psychol.: Hum. Perc. Perf.* **10**, 449–454.

Von Holst, E. 1973 *The behavioural physiology of animals and man.* London: Butler and Tanner.

Young, M. J., Landy, M. S. & Maloney, L. T. 1993 A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Res.* **33**, 2685–2696.

Zacharias, G. & Young, L. 1977 Manual control of yaw motion with combined visual and vestibular cues. In *13th Annual conference on manual control,* pp. 389–402. Cambridge, MA: MIT Press.