Temporal aspects of binocular slant perception

Raymond van Ee and Casper J. Erkelens

Vision Research 36, 43-51, 1996.

Please send correspondence to the first author:
Utrecht University, Vakgroep Fysica v/d Mens,
Princetonplein 5, NL–3584CC, Utrecht, The Netherlands;
Email: r.vanee@phys.uu.nl;
Fax: +31 30 522664

Keywords: Binocular disparity, Binocular vision, Human slant perception, Stereopsis

Abstract

We investigate temporal aspects of binocular slant perception in the presence and absence of a visual reference. Subjects judge slant induced by large–field stereograms of which one half–image is either horisontally scaled or sheared relative to the other half–image. Each stimulus is presented for different observation periods ranging from 0.1 to 19.2 sec. We quantitatively corroborate earlier findings that perceived slant develops significantly faster and to higher levels with visual reference than without it. In daily life, when we are active, there will not be much time for slant to develop. We find that if observation periods are brief (a few seconds or less) slant is poorly perceived if there is no visual reference. We conclude that the visual system is relatively insensitive to large–field horisontal scale and shear.

Introduction

Experimental knowledge about disparity–based three–dimensional (3D) vision has been obtained mainly from studies which use stereograms produced by stereoscopes (first developed by Wheatstone, 1838) or Julesz random–dot patterns (first developed by Julesz, 1960).

From the time of Wheatstone it has been clear that perception of distance is induced by local spatial shifts between the pattern viewed by the right eye relative to the pattern viewed by the left eye.[1] Perceived orientations of planar surfaces are related to linear transformations between the half–images of stereograms, which include horizontal scale and horizontal shear (see figure 1 for an explanation of these transformations). Perception of non–planar surfaces depends on higher–order spatial differences between the half–images of stereograms. This study concentrates on perceived orientations of planar surfaces.

Gibson (1950) described surface orientation in human vision in terms of the amount and direction of slant. Stevens (1983) proposed a formal way for describing encodings of surface orientation and showed it to be consistent with various psychophysical phenomena. His proposal was that surface orientation can be uniquely quantified by slant and tilt. Slant is the angle between the surface–normal and the line of sight. Tilt denotes the slant direction which is the angle between the projection of the surface–normal on the fronto–parallel (frontal) plane and the horizontal in the frontal plane.

Figure 1 about here.

Horizontal scale of one half–image of a stereogram relative to the other half–image is perceived as slant about the vertical axis (e.g. Julesz, 1971; the tilt remains zero), whereas horizontal shear is perceived as slant about the horizontal axis (e.g. Wallach & Bacon, 1976; tilt is 90 deg). Slant estimation about oblique axes can be described, both theoretically and experimentally, solely by horizontal scale and horizontal shear between the images viewed by the left and the right eye (van Ee & Erkelens, 1993; 1995).

However, several studies report that horizontal scale and shear between the two half–images of a stereogram without a visual reference are poorly perceived (Shipley & Hyson,

[1]Since the retinae are spherical and observed stereograms are usually planar, linear transformations of the stereogram half–images relative to each other lead to non–linear transformations between their retinal images. For instance, planar half–patterns of a stereogram with zero shift relative to each other at a certain distance from the eyes nevertheless give rise to essentially non–linear retinal disparities (see also figure 4.4 in Foley, 1991). It can also be calculated that a horizontal scale (gradient) between the half–patterns of the stereogram is certainly not a gradient in the disparity domain. This study is concerned with transformed viewed (screen) images.

1972; Mitchison & Westheimer, 1984, 1990; Stevens & Brookes, 1987, 1988; Gillam, Chambers & Russo, 1988). In addition, horizontal scale between the two retinal images caused by an aniseikonic lens leads to perception of slant only after considerable latencies (Ames, 1946; Seagrim, 1967; Gillam, Flagg & Finlay, 1984).

During the conference/NATO-workshop on binocular stereopsis in Toronto in 1993 we reported on the fact that the visual system is poorly susceptible to horizontal scale and horizontal shear if these transformations comprise the entire visual field, which means that the transformations are presented without a visual frame of reference (Erkelens & van Ee, 1993). As a reason for this poor susceptibility we suggested that a whole–field disparity field caused by horizontal scale or horizontal shear between the half–images is an unreliable measure for the visual system. Such a whole–field disparity field could also be caused by a head rotation: Head rotation about the vertical axis causes a horizontal scale between the two retinal images; head rotation about the horizontal axis causes a horizontal shear between the two retinal images. During the same conference, Howard and Kaneko (1993) presented experimental results which showed that slant perception caused by whole–field horizontal shear was clearly perceived even when presented without visual reference. The most conspicuous differences between Howard and Kaneko's experiment and our experiment were the lengths of time during which the stimuli were observed. In the experiment of Howard and Kaneko the subjects were allowed to look at the stimulus for as long as they wished, which was on average for about 15 sec, whereas we presented the stimuli for less than 5 sec. Intrigued by the rather different experimental results we decided to investigate the temporal aspects of binocular slant perception.

The purpose of this study is primarily to compare slant estimation with and without the presence of a visual reference for a large range of presentation periods. We begin by investigating the perceived slant that can be expected on geometrical grounds from the horizontal scale and shear transformations.

The theoretical relationship between slant and horizontal scale and shear

We mentioned that much knowledge about binocular 3D vision has been obtained from studies with stereograms. Another method which has proved important for the understanding of binocular slant perception is based on the use of an aniseikonic lens which is positioned just in front of one eye (e.g. Ogle, 1950 or Gillam et al., 1984). Ogle derived a relationship between the horizontal magnification of the aniseikonic lens and the slant about the vertical axis:

$$\text{slant} = \arctan\left(\frac{M-1}{M} \cdot \frac{z_0}{I}\right) ,$$

where M is the horizontal magnification factor of the lens, I is the interocular distance and z_0 is the distance between the stimulus and the eyes. Several authors have used Ogle's relationship also as a measure of the perceived slant induced by horizontal scale between the half-images of a stereogram. Strictly speaking, however, Ogle's relationship was derived for aniseikonic lenses, not for stereograms.[2] Furthermore, as far as we know Ogle did not derive a relationship between horizontal shear and slant about the horizontal axis. We need these relationships for stereograms in order to produce stimuli which theoretically induce equal slants about the horizontal and vertical axes. Therefore we start by deriving relationships between slant and horizontal scale on the one hand and between slant and horizontal shear on the other hand. We will adopt Ogle's notation.

Suppose that there is a unique correspondence between both half-images of the stereogram and that fusion has been established. Consider the visual ray between a particular stimulus point in the left-eye half-image (x_0) and the nodal point in the left eye (N_l, figure 2). Consider also the visual ray between the corresponding stimulus point of the right-eye half-image (Mx_0) and the nodal point of the right eye (N_r). The point of intersection of the two visual rays defines the locus of a stimulus point (P) which can be positioned outside the plane in which the stereogram is presented. In figure 2, this point P is located behind the screen. The complete set of intersection points obtained by this method defines a three-dimensional object whose image has the same retinal positions in the two eyes as the individual half-images of the stereogram.

Figure 2 about here.

We use an orthogonal coordinate system with the origin (O) at the centre between the eyes. The positive x-axis points to the right, the y-axis upwards and z-axis in the primary direction. With regard to a horizontal scale of magnification M (see figure 2) between both parts of the stereogram, the intersection point (P) of the visual rays of the stimulus point (x_0, y_0, z_0) can be obtained by a set of three equations which in their turn can be combined to form a single vector equation. The vectors we use may be thought of as arrows, one beginning at the nodal point of the left eye (N_l) and pointing to a stimulus point in the left-eye part of the stereogram (x_0), the other beginning at the nodal point of the right eye (N_r) and pointing to the corresponding, but transformed, stimulus point

[2]There are at least two principal differences between the two situations. 1) Retinal horizontal scale (caused by an aniseikonic lens) leads merely to modifications of the horizontal disparity. Horizontal scale of a stereogram also leads to modifications of vertical disparities because the left–most and right–most parts of the two half–images are at different distances, in particular for large–field stimuli. 2) The second reason is explained in footnote 1 and is based on the fact that stereograms are planar whereas the retinae are spherical. Because aniseikonic lenses are positioned directly in front of the eye, the retinal–disparity is proportional to the angle of the visual direction. In a stereogram the screen–disparity is proportional to the position on the screen.

in the right–eye part of the stereogram (Mx_0). The vector equation is:

$$\begin{pmatrix} -\frac{1}{2}I \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} \frac{1}{2}I + x_0 \\ y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}I \\ 0 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -\frac{1}{2}I + Mx_0 \\ y_0 \\ z_0 \end{pmatrix} ,$$

where λ and μ are scalars. Figure 2 illustrates in two dimensions how the x– and z–components of the vector–equation have been obtained. The coordinates (P_x, P_y, P_z) of the intersection point P are:

$$P = (-\frac{1}{2}I + \frac{I(\frac{1}{2}I + x_0)}{I + (1-M)x_0} , \frac{y_0 I}{I + (1-M)x_0} , \frac{z_0 I}{I + (1-M)x_0}) .$$

Not only the x– and z–coordinates of the intersection point but also the y–coordinate depend on the amount of horizontal scale. This means that slant due to horizontal scale is not exactly equivalent to rigid rotation of an object about the vertical axis. The dependence of the y–coordinate of the intersection point on the amount of horizontal scale explains why horizontally scaled rectangles are perceived as trapezoids of which the small vertical side is perceived to be nearer than the large vertical side. Figure 2 shows that slant equals $\arctan(\Delta z/\Delta x)$ and thus equals $\arctan((P_z - z_0)/P_x)$. As a result, the relationship between slant and horizontal scale is:

$$\text{slant} = \arctan(\frac{M-1}{M+1} \cdot \frac{2z_0}{I}) , \ (tilt = 0) .$$

Note that this relationship differs from the relationship that has been derived for aniseikonic lenses. (In practice the difference in predicted slant is small: with respect to our experimental set–up, the difference is 1 deg for a slant of 60 deg.) We repeat the procedure for a horizontal shear defined by angle β (see figure 1):

$$\begin{pmatrix} -\frac{1}{2}I \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} \frac{1}{2}I + x_0 \\ y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}I \\ 0 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -\frac{1}{2}I + x_0 + y_0 \tan\beta \\ y_0 \\ z_0 \end{pmatrix} ,$$

from which it follows that the intersection point is:

$$P = (-\frac{1}{2}I + \frac{I(\frac{1}{2}I + x_0)}{I - y_0 \tan\beta} , \frac{y_0 I}{I - y_0 \tan\beta} , \frac{z_0 I}{I - y_0 \tan\beta}) .$$

As before, not only the y– and z–coordinates, but also the x–coordinate of the intersection point depend on the amount of shear. This explains why horizontally sheared rectangles are perceived as trapezoids of which the small horizontal side is perceived to be nearer than the large horizontal side. In the case of horizontal shear, slant is equal to $\arctan(P_z - z_0/P_y)$. The relationship between slant and horizontal shear is:

$$\text{slant} = \arctan(\tan\beta \cdot \frac{z_0}{I}) \ , \ (tilt = 90\ deg) \ .$$

By using the two derived relationships between horizontal scale, horizontal shear and slant we are able to produce stimuli which should induce equal slants about the horizontal and vertical axes.

Figure 3 about here.

Methods

Figure 3 shows the experimental set–up. We used an anaglyph set–up for the generation of the stereograms. The stereograms were generated at a frequency of 70 Hz by an HP 750 graphics computer. Subsequently, the stimuli were back–projected on a frontal translucent screen by a projection TV (Barco Data 800). The subject was seated about 1.5 metres from the screen. The right–eye half–image was projected in green light and was observed through a green filter. Red filters were used to make the other image visible exclusively to the left eye. The transmission spectra of the filters (Schott Tiel, the Netherlands) were chosen such that they corresponded as well as possible to the emission spectra of the projection TV.

In order to compare our study closely with the Howard and Kaneko (1993) study we used similar stereograms. The stimulus was circular and contained randomly distributed circles (see figure 3). The advantage of such a stimulus is that perspective cues play only a minor role. The small circles had a diameter of 1.5 deg each and a density of about 10 %. The large circle had a diameter of 70 deg. A different, randomly chosen configuration of circles was presented during each trial.

The presented transformations of the green part relative to the red part of the stereogram were either horizontal scale (slant with a tilt of 0 deg) or horizontal shear (slant with a tilt of 90 deg). Horizontal scale varied between -9.0 % and 9.0 % in six steps (with a step–size of 3.0 %) and horizontal shear varied between -4.9 deg and 4.9 deg (again in six steps). The magnitudes of the scale and shear transformations were chosen such that they were identical to each other with respect to the amount of predicted slant. As can be inferred from the derived slant-equations, the magnitudes are related according to the equation $\beta = \arctan(2 \cdot \frac{M-1}{M+1})$, where β indicates the angle of shear and M the magnitude of scale. For instance, both a horizontal scale of 6.0 % and a horizontal shear of 3.3 deg theoretically induce a slant of 53 deg (for z_0 = 150 cm and I = 6.5 cm). The amounts of transformation presented comprise more or less the entire range of fusible disparities; care was taken to prevent fusion problems.

Figure 4 about here.

The task of the subject was to judge the perceived slant induced by the presented horizontal scale or shear transformations. After each presentation two lines (one fixed and one rotatable) appeared on the screen, as shown in figure 4. By changing the computer–mouse position, the subjects set the angle between the adjustable line and the fixed line; the angle represented the estimated slant. The fixed line represented the frontal plane, the adjustable line represented the slanted plane. The two lines were displayed without disparity. Therefore they were perceived in the plane of the screen. An advantage of this adjustment procedure was that during the adjustment, slant adaptation effects due to the former test stimulus were cancelled out. The adjustable stimulus served as a mask.

Experiments were of two types: transformations were presented either with or without a visual reference. In the situation without visual reference the stimuli were viewed in a completely dark room; the only thing visible in the experiments was the stimulus. This type of experiment was preceded by a dark–adaptation period of six minutes. Because the subjects were dark adapted, experiments could be done with low contrast and low brightness of the projection TV, without loss of visibility of either part of the stereogram. This means on the one hand that the screen did not (as far as possible) serve as an illuminated plane and on the other hand that we made optimum use of the transmission spectrum of the anaglyph glasses. The subjects did not experience any crosstalk between the right–eye and left–eye views of the stereogram. The brightness of the red part and the brightness of the green part of the stimuli were adjusted independently such that the two parts were perceived as equally bright.

During the series of trials in which we used a visual reference, a whole–field reference pattern was projected on the screen as shown in figure 5. The reference (width 70 deg and height 70 deg) consisted of a cross–hatched pattern. The cross–hatched pattern was made up of a field of adjacent squares with diagonals of 15 deg. To prevent aliasing effects (fixation on false depth planes) not every possible square was shown. The density of squares was about 60 %. The reference pattern was changed randomly every time a new stimulus appears. The room was also dimly lit (in the situation with reference) which prevented depth contrast effects that might have induced perceived slant of the reference pattern due to perceived slant of the circular test pattern.

Figure 5 about here.

The following observation periods were presented in random order: 100 msec, 200 msec, 400 msec, 800 msec, 1.6 sec, 3.2 sec, 6.4 sec, 12.8 sec and 19.2 sec. Each trial was repeated seven times. In all, each subject had to judge 1764 slants, namely 9 presentation times, 2 transformations (horizontal scale and shear), 2 conditions (with and without visual reference), 7 magnitudes of transformations (-9.0, -6.0, -3.0, 0.0, 3.0, 6.0, 9.0 % or -4.9, -3.3, -1.6, 0.0, 1.6, 3.3, 4.9 deg) and 7 repetitions. The subjects started with a series of trials without visual reference (882 trials). Then after a two–hour break the subjects repeated the same series of trials but with visual reference (again 882 trials). Each series lasted about 75 minutes.

Six subjects (5 males and 1 female, ages 23–29 years) took part in the experiment. Although each subject was familiar with the concept of mathematical angles, we checked before starting the experiment whether the subject was able to represent slant by our method in a consistent manner. Therefore, a series of trials with real and dichoptically projected slanted planes was conducted with each of our subjects. During the final experiment no feedback was given about the results. Except for the author (RE) the subjects were uninformed about the purposes of the experiment. Four subjects (FV, CG, OS, JZ) were inexperienced with respect to stereoscopic experiments. Four subjects (FV, JZ, RE and FO) showed refraction anomalies which were corrected by their own glasses. One subject (FO) showed a slight astigmatism (0.5 diopters, axis 15 deg) which was corrected by his own glasses.

Results

Estimated slant as a function of predicted slant for a typical subject (CG) is shown in figure 6 for the transformations horizontal scale and shear and the conditions with and without visual reference. For the sake of clarity and in order to explain the method of data analysis, (for the moment) we show the results for only three observation periods, namely 0.1 sec, 1.6 sec and 19.2 sec. The results show, to a good approximation, a linear relationship between estimated and predicted slant. The settings for each transformation, condition and observation period are fitted by a linear function. The good fits to linear functions imply that slant estimation follows Weber's law. The slopes of the fitted linear functions represent the fraction of predicted slant that is estimated by the subject.

Figure 6 about here.

The complete data (estimated slant as a fraction of predicted slant for each observation period) for our six subjects are presented in figure 7. The fits to the raw data (such as presented in figure 6) are good in all cases. χ^2 (obtained by a least–squares method based on the experimental outcomes) is always larger than 0.92 and in most cases larger than 0.99. Figure 7 shows that estimated slant develops over time for each condition. Estimated slant in the presence of a visual reference develops faster and to a higher level than without visual reference. For brief observation periods (of the order of 1 second or less) slant is poorly estimated when no visual reference is present.

Figure 7 about here.

Like Gillam et al. (1984) and Mitchison and McKee (1990) we find that estimated slant as a fraction of predicted slant is smaller than unity in all cases, which means that slant is consistently underestimated. Slant judgments are rather similar for three subjects (FV, OS and JZ) with respect to both horizontal scale and shear. Two subjects (CG and FO) show the well–known anisotropy for slant perception in favour of shear (Wallach & Bacon, 1976; Rogers & Graham, 1983; Mitchison & McKee, 1990). One subject (RE), however, shows an anisotropy in favour of horizontal scale. For one subject (FV) the anisotropy in favour of shear exists only for observation periods of 100 and 200 msec. Subjects RE and FO are experienced in stereoscopic experiments. These subjects show smaller differences between the conditions with and without visual reference.

Discussion

We investigated temporal aspects of binocularly estimated slant induced by horizontal scale and horizontal shear. These transformations were presented by means of a stereogram. Our findings corroborate earlier findings that estimated slant develops over time. Slant in the presence of a visual reference develops significantly faster than without visual reference. For brief observation periods (of the order of 1 sec or less) slant is poorly perceived, particularly when no visual reference was present.

Significance of a visual reference

Many studies have shown that perception of slant induced by a certain stimulus depends on the presence of items surrounding the stimulus (Werner, 1938; Pastore, 1964; McKee, 1983; Mitchison & Westheimer, 1984, 1990; Gillam et al, 1984; Gillam et al. (1988); Fahle & Westheimer, 1988; Mitchison & McKee, 1990). Shipley and Hyson (1972), Gillam et al. (1984) and Gillam et al. (1988) showed that a second surface has a facilitating effect on slant perception of a test surface. Gillam et al. (1988) proposed that boundaries in the presented stereograms are effective in inducing slant (by a boundary they meant that steps in disparity were present between different parts of the stimulus). This finding confirmed the results of one of their earlier experiments which they performed with aniseikonic lenses (Gillam et al., 1984). Stevens and Brookes (1987, 1988) proposed that binocular 3D information is effectively integrated only where the surface exhibits curvature features or edge discontinuities. Another study on the significance of a visual reference was performed by Erkelens and Collewijn (1985a,b) and Regan, Erkelens and Collewijn (1986) who were dealing with oculomotor behaviour. They studied perception of motion in depth caused by whole–field spatial shifts between stereograms. Gillam et al. (1984), Erkelens and Collewijn (1985a,b), Regan et al. (1986), Gillam et al. (1988) concluded that absolute disparity is not a cue for depth perception.

Despite the above–mentioned findings, it is not yet generally acknowledged that linear transformations between the stereoscopic half–images elicit poor perception of slant. Many experiments have been done under different circumstances. A number of authors did not report explicitly whether, or to what extent, a visual reference was present. Several studies were conducted in a room which was not entirely dark, without the experimenters realising that a dimly lit room acts as a visual reference. Therefore, many experimental results are not unambiguously comparable with each other.

Howard and Kaneko (1993, 1994) (concerning horizontal shear) and very recently Kaneko and Howard (1994) (concerning horizontal scale) explicitly reported on the care they had taken to exclude all visual stimuli that could serve as a reference. They found that slant induced by horizontal scale and shear was clearly perceived without a visual reference. In their experiments subjects were allowed to take as long as they wished in order to estimate perceived slant (on average about 15 sec). One of the main reasons for our present experiments was the apparent discrepancy between the results of Howard and Kaneko (1993) and our previous results (Erkelens & van Ee, 1993) which showed poor slant perception. Our present findings explain why these results were different. The length of the observation period appears to be a decisive parameter.

Slant perception and latencies

Another study concerning temporal aspects of slant perception caused by horizontal scale and horizontal shear (and a number of other transformations) is the one by Gillam et al. (1988). They measured latency from stimulus onset to fusion and latency from fusion to stereoscopic resolution. By resolution they meant the recognition of one out of twelve configurations. The task of their subjects was to press a button when fusion was obtained and to release the button when the stimulus was recognised. Fusion was described by the authors as a state 'in which the subject could see a clear and single image without fuzziness or a feeling of busyness'. Fusion latency (which in fact is the time required for the subject to get the described feeling and to activate the motor system to press a button) was found to be about 1 sec. Recognition took roughly between 5 and 60 seconds.

In our study, we find much shorter latencies. However, the type of task of Gillam et al. (1988) differs fundamentally from our task where stereograms are presented for a limited period of time. Their experiment was designed to compare latencies of slant for different stimuli and not to measure latency per se. This means that from their study it is not possible to derive a precise relationship between the presentation period of the stimulus and the amount of perceived slant.

Uttal, Davis and Welke (1994) recently reported that a powerful and compelling stereoscopic experience can be elicited with very brief (≤ 1 *msec*) stimulus durations. In fact, what they measured was whether subjects could distinguish a flashed, stereoscopic surface from 7 other surfaces. Their results show that subjects' scores exceeded 70% correct (in 5 out of 8 surfaces) only when strongly curved surfaces are used. It would be interesting to know what the minimum stimulus duration is for subjects to be able to give a quantitative stimulus characterisation, for instance using the shape index (Koenderink, 1990). The results of de Vries, Kappers and Koenderink (1994) suggest that subjects need exposures much longer than 1 msec to do such quantitative tasks.

Slant perception and ego-motion

A possible explanation for the observed phenomena is that two different linear transformations (gradients) within the stereogram are required to obtain unambiguous slant perception; this explanation was suggested earlier by Shipley and Hyson (1972) and Gillam et al. (1988). As remarked by Mitchison and Westheimer (1990), it turns out that viewing in oblique directions introduces disparity gradients. They used this insight to give a possible explanation for the well-known horizontal/vertical anisotropy (Rogers & Graham, 1983). On the basis of their insight we suggest that the reason for the difference in perception of one gradient (poor perception) and two different gradients (better perception) is that one gradient can be the result of an ego-movement (for instance after a left-right rotation of the head). This gradient is therefore unreliable in the case of brief observation periods and is primarily ignored as a signal for perception of slant. Two different gradients present at the same time in the stereogram cannot be a result of an ego-movement and therefore form a more reliable stimulus for slant perception.

A possible explanation for the fact that subjects can perceive slant without visual reference, although only after long latencies, is that extra-retinal signals or information about perspective may be involved in the slant perception. Inspection of the stimulus by eye movements might also be involved in the enhancement of the slant perception with time because the subjects were not required to fixate a single point.

An explanation similar to the one which has been put forward for the poor perception of slant caused by linear transformations between the half-images has been proposed in the literature for the poor perception of whole-field shifts. Offsets between the half-images of a stereogram are not interpreted as changes in distance (Erkelens & Collewijn, 1985a; Regan et al., 1986). On the other hand, two different offsets are very effective. The analogy is that offsets are most probably due to errors in the horizontal vergence of the eyes. These offsets occur during binocular fixation of a target moving in depth, after saccadic eye movements, or during head movement and have magnitudes of as much as 1–2 deg in position (Erkelens & Collewijn 1985b; Collewijn & Erkelens, 1990) and 1 deg/s in velocity (Steinman & Collewijn, 1980) and are therefore best ignored as signals for depth.

Analoguous reasoning, but now concerning cyclodisparity, has been used by Howard, Ohmi and Sun (1993). They based their arguments on the results of Howard and Zacher (1991) who showed that even when cyclovergence (disjunctive ocular torsion) occurs slant perception is not altered. Howard et al. (1993) concluded that overall cyclodisparities could signify that the eyes are misaligned and are therefore ignored by the perceptual system for the purpose of judging slant. Cyclodisparities are of considerable amplitude and occur frequently during natural behaviour, as was found by van Rijn, van der Steen and Collewijn (1994). Whereas in the previous paragraphs two gradients were required for unambiguious perception, here two different cyclodisparities are required for reliable slant perception: Collewijn, van der Steen and van Rijn (1991) reported that thresholds for perception of slant due to cyclodisparity increased by a factor of 7 when the visual reference was removed.

Cyclovergence and slant perception

The occurrence of cyclovergence could, in principle, alter the perception of slant (Rogers, 1992). Rogers hypothesized that the role of cyclovergence is to nullify the transverse positional disparities along the horizontal (interocular) axis. He suggested therefore that the transformation horizontal shear does not generate cyclovergence. This was indeed one of his experimental results (Rogers, 1992). By the same reasoning the transformation horizontal scale will not induce cyclovergence: on the one hand, cyclovergence cannot minimize the disparities caused by the horizontal magnification, on the other hand no transverse positional disparities along the horizontal axis are present to drive cyclovergence. Finally, Howard and Zacher (1991) and Howard et al. (1993) showed that perception of slant remains stable even when cyclovergence changes. On the basis of these reports we decided that monitoring cyclovergence in our experiment would not produce new information.

Underestimation of slant

Perception of depth based on stereopsis alone is not veridical (Gogel, 1960; Foley, 1980; Johnston, 1991; Parker, Johnston, Mansfield & Yang, 1991; Johnston, Cumming & Parker, 1993; Johnston, Cumming & Landy, 1994). At short distance depth is overestimated and at long distance it is underestimated. There is an intermediate distance where depth perception is veridical which varies between subjects, averaging 80 cm (Johnston, 1991). Like Gillam et al. (1984) and Mitchison and McKee (1990) we find that perceived slant is consistently underestimated. The fact that in our experiment the distance between observer and screen was much larger than 80 cm (namely 150 cm) can therefore explain the underestimations.

Another possible explanation for the underestimation of slant in our experiment is that subjects cannot make precise estimations of angles. In the study of Jastrow (1893) subjects had to view a drawn angle as long as was needed to fix it in mind. Immediately afterwards, the subject had to draw, from memory, another angle equal to the first. Jastrow found that angles larger than 25 deg and smaller than 75 deg were consistently underestimated. The order of underestimation, however, was only about 10 % which is too small to explain our underestimations.

Conclusion

We investigated the relationship between estimated slant (caused by the transformations horizontal scale and shear between the two half–images of a stereogram) and the observation period of the stimulus. In previous studies dealing with binocular slant perception, stimuli were presented for long observation periods in order to allow perceived depth to develop over time. We quantitatively corroborate earlier findings that binocular slant perception develops over time. In daily life, when we are active, there will not be enough time for slant to develop. We conclude that the visual system is relatively insensitive to large–field horizontal scale and shear.

References

Ames, A. (1946). Binocular vision as affected by relations between uniocular stimulus–patterns in commonplace environments. *American Journal of Psychology*, 59, 333-357.

Anstis, S. M., Howard, I. P. & Rogers, B. (1978). A Craik–O'Brien–Cornsweet illusion for visual depth. *Vision Research*, 18, 213-217.

Collewijn, H. & Erkelens, C. J. (1990). Binocular eye movements and perception of depth. In Kowler, E. (Ed.), *Eye movements and their role in visual and cognitive processes* (pp 213–261). Amsterdam: Elsevier Science Publishers B.V.

Collewijn, H., van der Steen, J. & van Rijn, L. J. (1991). Binocular eye movements and depth perception. In Gorea, A. (Ed.), *Representations of vision, trends and tacit assumptions in vision research* (pp 165–183). Cambridge: Cambridge University Press.

van Ee, R. & Erkelens, C. J. (1993). Slant perception in relation to the geometry of binocular vision: Effects of local and global disparity fields. *Poster presentation at the International Conference and Nato Advanced Workshop on Binocular Stereopsis and Optic Flow*, Toronto, Canada.

van Ee, R. & Erkelens, C. J. (1995). Binocular perception of slant about oblique axes relative to a visual frame of reference. *Perception, in press.*

Erkelens, C. J. & Collewijn, H. (1985a). Motion perception during dichoptic viewing of moving random-dot stereograms. *Vision Research*, 25, 583–588.

Erkelens, C. J. & Collewijn, H. (1985b). Eye movements and stereopsis during dichoptic viewing of moving random-dot stereograms. *Vision Research*, 25, 1689–1700.

Erkelens, C. J. & van Ee, R. (1993). Disparity processing for depth perception under dynamic viewing conditions. *Oral presentation at the International Conference and Nato Advanced Workshop on Binocular Stereopsis and Optic Flow*, Toronto, Canada.

Fahle, M. & Westheimer, G. (1988). Local and global factors in disparity detection of rows of points. *Vision Research*, 28, 171-178.

Foley, J. M. (1980). Binocular distance perception. *Psychological Review*, 87, 411–434.

Foley, J. M. (1991). Binocular space perception. In Regan, D. (Ed.), *Binocular vision* (pp 75–92). London: Macmillan Press.

Gibson, J. J. (1950). *The perception of the visual world*. Boston: Houghton Mifflin.

Gillam, B., Chambers, D. & Russo, B. (1988). Postfusional latency in stereoscopic slant perception and the primitives of stereopsis. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 163-175.

Gillam, B., Flagg, T. & Finlay, D. (1984). Evidence for disparity change as the primary stimulus for stereoscopic processing. *Perception & Psychophysics*, 36, 559-564.

Gogel, W. C. (1960). The perception of shape from binocular disparity cues. *Journal of Psychology*, 50, 179-192.

Howard, I. P. & Kaneko, H. (1993). Shear disparity, cyclovergence and the perception of surface inclination. *Oral presentation at the International Conference and Nato Advanced*

Workshop on Binocular Stereopsis and Optic Flow, Toronto, Canada.

Howard, I. P. & Kaneko, H. (1994). Relative shear disparities and the perception of surface inclination. *Vision Research*, 34, 2505-2517.

Howard, I. P., Ohmi, M. & Sun, L. (1993). Cyclovergence: a comparison of objective and psychophysical measurements. *Experimental Brain Research*, 97, 349–355.

Howard, I. P. & Zacher, J. E. (1991). Human cyclovergence as a function of stimulus frequency and amplitude. *Experimental Brain Research*, 85, 445–450.

Jastrow, J. (1893). On the judgment of angles and position of lines. *American Journal of Psychology*, 5, 217-248.

Johnston, E. B. (1991). Systematic distorsions of shape from stereopsis. *Vision Research*, 31, 1351-1360.

Johnston, E. B., Cumming, B. G. & Landy, M. S. (1994). Integration of stereopsis and motion shape cues. *Vision Research*, 34, 2259-2275.

Johnston, E. B., Cumming, B. G. & Parker, A. J. (1993). Integration of depth modules: Stereopsis and texture. *Vision Research*, 33, 813-826.

Julesz, B. (1960). Binocular depth peception of computer–generated patterns. *Bell Systems Technical Journal*, 39, 1125-1162.

Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago: University of Chicago Press.

Kaneko, H. & Howard, I. P. (1994). Compression—magnification disparities and the perception of surface slant. *Perception*, 23, p.33.

Koenderink, J. J. (1990). *Solid shape*. Cambridge, Mass: MIT Press.

McKee, S. P. (1983). The spatial requirements for fine stereoacuity. *Vision Research*, 23, 191–198.

Mitchison, G. J. & McKee, S. P. (1990). Mechanisms underlying the anisotropy of stereoscopic tilt perception. *Vision Research*, 30, 1781-1791.

Mitchison, G. J. & Westheimer, G. (1984). The perception of depth in simple figures. *Vision Research*, 24, 1063-1073.

Mitchison, G. J. & Westheimer, G. (1990). Viewing geometry and gradients of horizontal disparity. In Blakemore, C. (Ed.), *Vision: Coding and Efficiency* (pp 302–309). Cambridge: Cambridge University Press.

Ogle, K. N. (1950). *Researches in binocular vision*. Philadelphia: Saunders.

Parker, A. J., Johnston, E. B., Mansfield, J. S. & Yang, Y. (1991). Stereo, surfaces and shape. In Landy, M. S. & Movshon, J. A. (Eds.), *Computational models of visual processing* (pp 359–381). Cambridge, Mass.: MIT Press.

Pastore, N. (1964). Induction of stereoscopic depth effect. *Science*, 144, p. 888.

Regan, H., Erkelens, C. J. & Collewijn, H. (1986). Necessary conditions for the perception of motion in depth. *Investigative Ophthalmology and Visual Science*, 27, 584–597.

van Rijn, L. J., van der Steen, J. & Collewijn, H. (1994). Instability of ocular torsion during fixation: cyclovergence is more stable than cycloversion. *Vision Research*, 34, 1077–1087.

Rogers, B. J. (1992). The perception and representation of depth and slant in stereoscopic surfaces. In Orban, G. & Nagel, H. (Eds.), *Artificial and Biological Vision Systems* (pp 241–266). Berlin: Springer.

Rogers, B. J. & Graham, M. E. (1983). Anisotropies in the perception of three-dimensional surfaces. *Science*, 221, 1409–1411.

Seagrim, G. N. (1967). Stereoscopic vision and aniseikonic lenses I. *British Journal of Psychology*, 58, 337–350.

Shipley, T. & Hyson, M. (1972). The stereoscopic sense of order—a classification of stereograms. *American Journal of Optometry and Archives of American Academy of Optometry*, 49, 83-95.

Steinman, R. M. & Collewijn, H. (1980). Binocular retinal image motion during active head rotation. *Vision Research*, 20, 415–429.

Stevens, K. A. (1983). Slant–Tilt: The visual encoding of surface orientation. *Biological*

Cybernetics, 46, 183-195.

Stevens, K. A. & Brookes, A. (1987). Depth reconstruction in stereopsis. *Proceedings of the first IEEE international conference on computer vision London*, 682-686.

Stevens, K. A. & Brookes, A. (1988). Integrating stereopsis with monocular interpretations of planar surfaces. *Vision Research*, 28, 371-386.

Uttal, W. R., Davis, S. D. & Welke, C. (1994). Stereoscopic perception with brief exposures. *Perception & Psychophysics*, 56, 599-604.

de Vries, S. C., Kappers, A. M. L. & Koenderink, J. J. (1994). Influence of surface attitude and curvature scaling on discrimination of binocularly presented curved surfaces. *Vision Research*, 34, 2409-2423.

Wallach, H. & Bacon, J. (1976). Two forms of retinal disparity. *Perception & Psychophysics*, 19, 375-382.

Werner, H. (1938). Binocular depth contrast and the conditions of the binocular field. *American Journal of Psychology*, 51, 489-497.

Wheatstone, C. (1838). On some remarkable and hitherto unobserved phenomena of binocular vision. *Philosophical Transactions of the Royal Society*, 128, 371–394.

Acknowledgements—We are grateful to P. Schiphorst for developing the stimuli. We thank S.M. McNab for linguistic advice. The work was supported by the Netherlands Organisation for Scientific Research (NWO), grant no. 810-404-006/1.

Figure captions

Fig. 1: The left–eye part of the stereogram is drawn in thin lines and has dimensions x_0 in the horizontal (interocular) direction and y_0 in the vertical direction. The right–eye part of the stereogram is transformed by a horizontal scale (upper figure) or horizontal shear (lower figure) and drawn in thick lines. Horizontal scale is usually expressed as a percentage. For instance, a horizontal scale of 3 % corresponds to a magnification factor M of 1.03. Horizontal shear is expressed in angles (β).

Fig. 2: The geometry of slant induced by horizontal scale (top view). To compute the slant which is induced by horizontal scale we find the set of intersection points (P) which belongs to fused (corresponding) stimulus points of both half–images of the stereogram. Each point of intersection is located at the position where the visual rays (arrows) of both eyes meet. The visual rays are expressed as vectors which start at the nodal point of each eye (N_l and N_r) and point to the stimulus locus. I is the interocular distance and M the horizontal magnification factor.

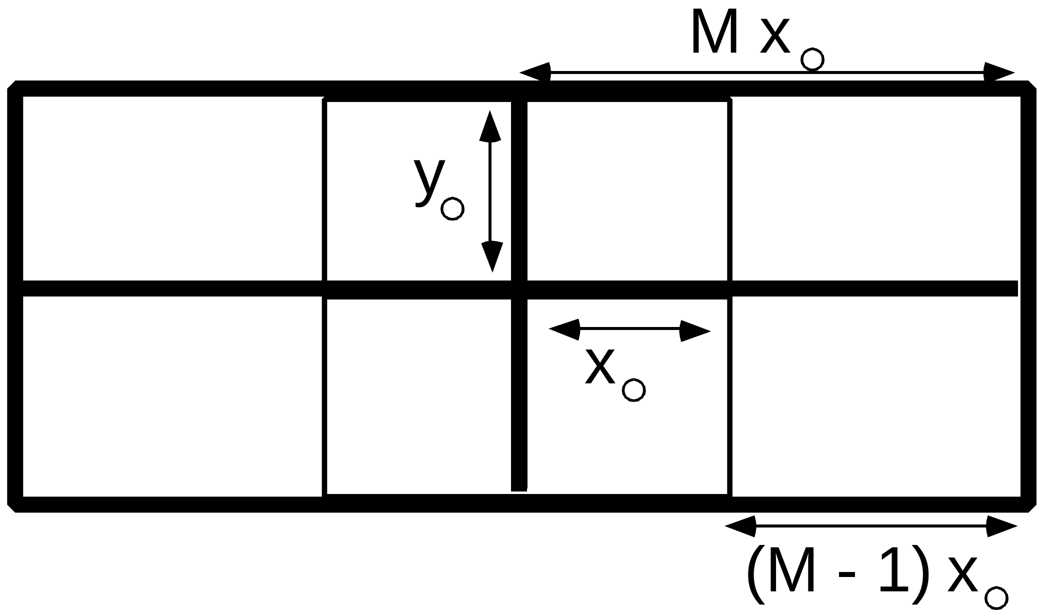
Fig. 3: The experimental set–up.

Fig. 4: The subject estimates the angle of perceived slant by manipulating the computer–mouse. In the case of pre–set horizontal scale (which means slant about the vertical axis) the left panel is presented to the subject. This panel corresponds to a top view of the experimental set–up. Horizontally sheared stimuli (slant about the horizontal axis) are followed by the screen image shown in the right panel (which corresponds to a side view).

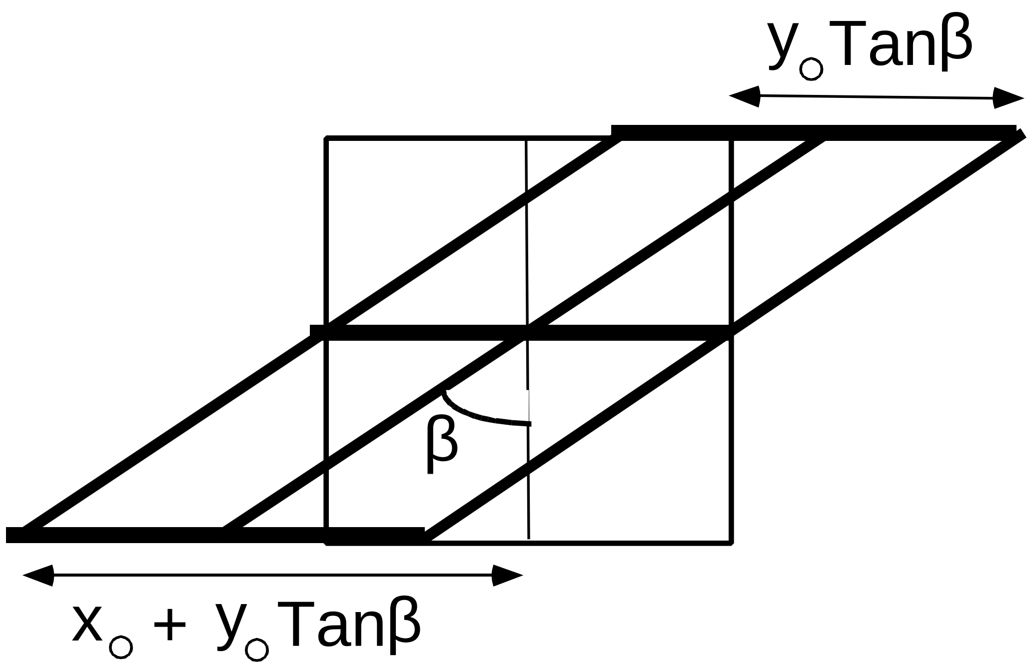
Fig. 5: The cross–hatched pattern serves as a visual reference and is presented in the plane of the screen. The diagonals of the individual squares are 15 deg. The density of the squares is about 60 %, the density of the circles is about 10 %.

Fig. 6: Estimated slant (and standard deviations) as a function of predicted slant of (a typical) subject CG, for three observation periods. Each data point is based on seven slant judgments. For each observation period the data are fitted by a linear function.

Fig. 7: Estimated slant as a function of the observation period. Each data point is based on 49 slant judgments. Differences in slant estimated with and without visual reference are large for the inexperienced subjects and smaller but still significant for the experienced subjects (RE and FO), especially when observation periods are shorter than one second.



Horizontal scale



Horizontal shear

Fig 1, van Ee & erkelens

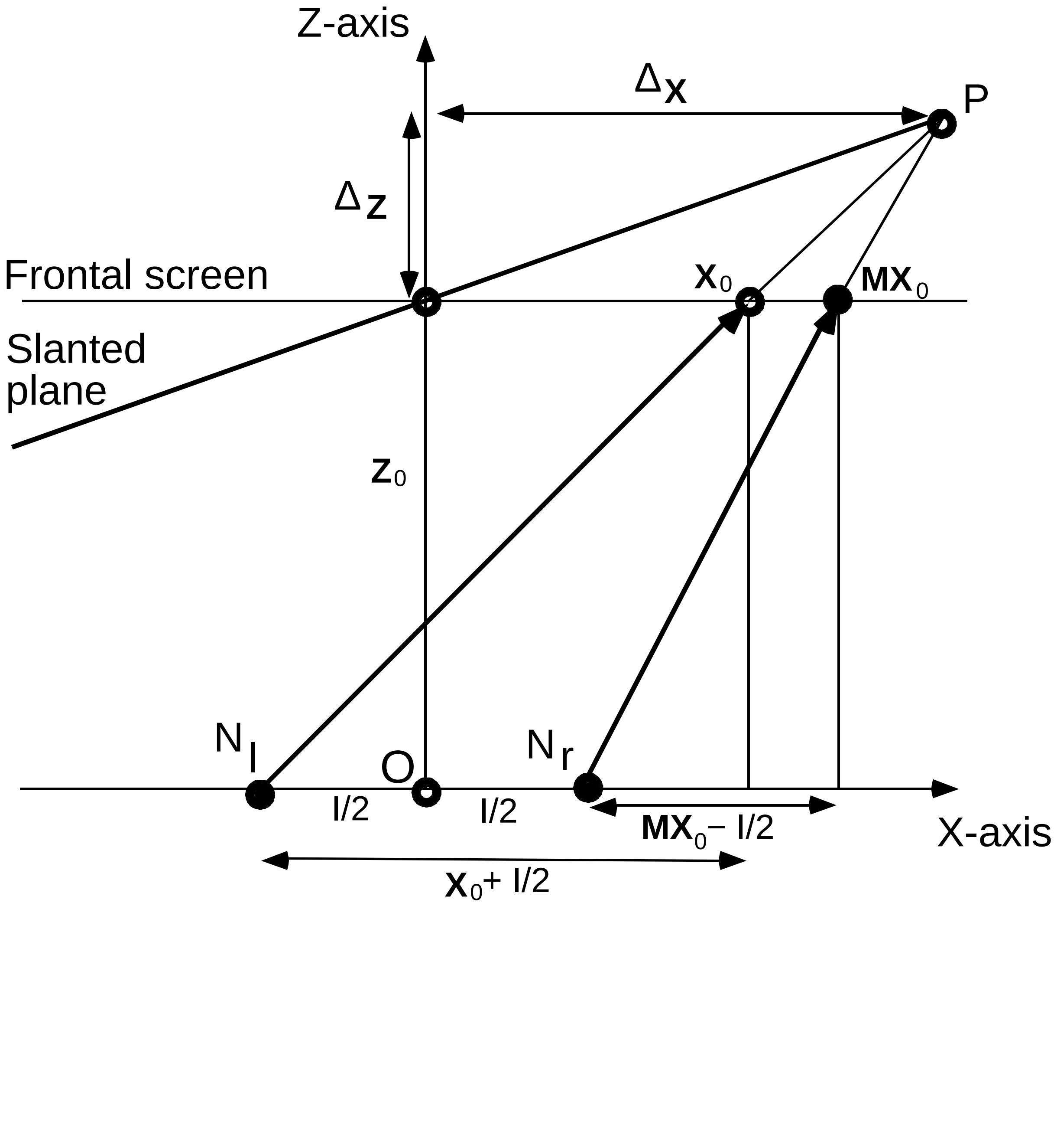


Fig 2, van Ee & erkelens

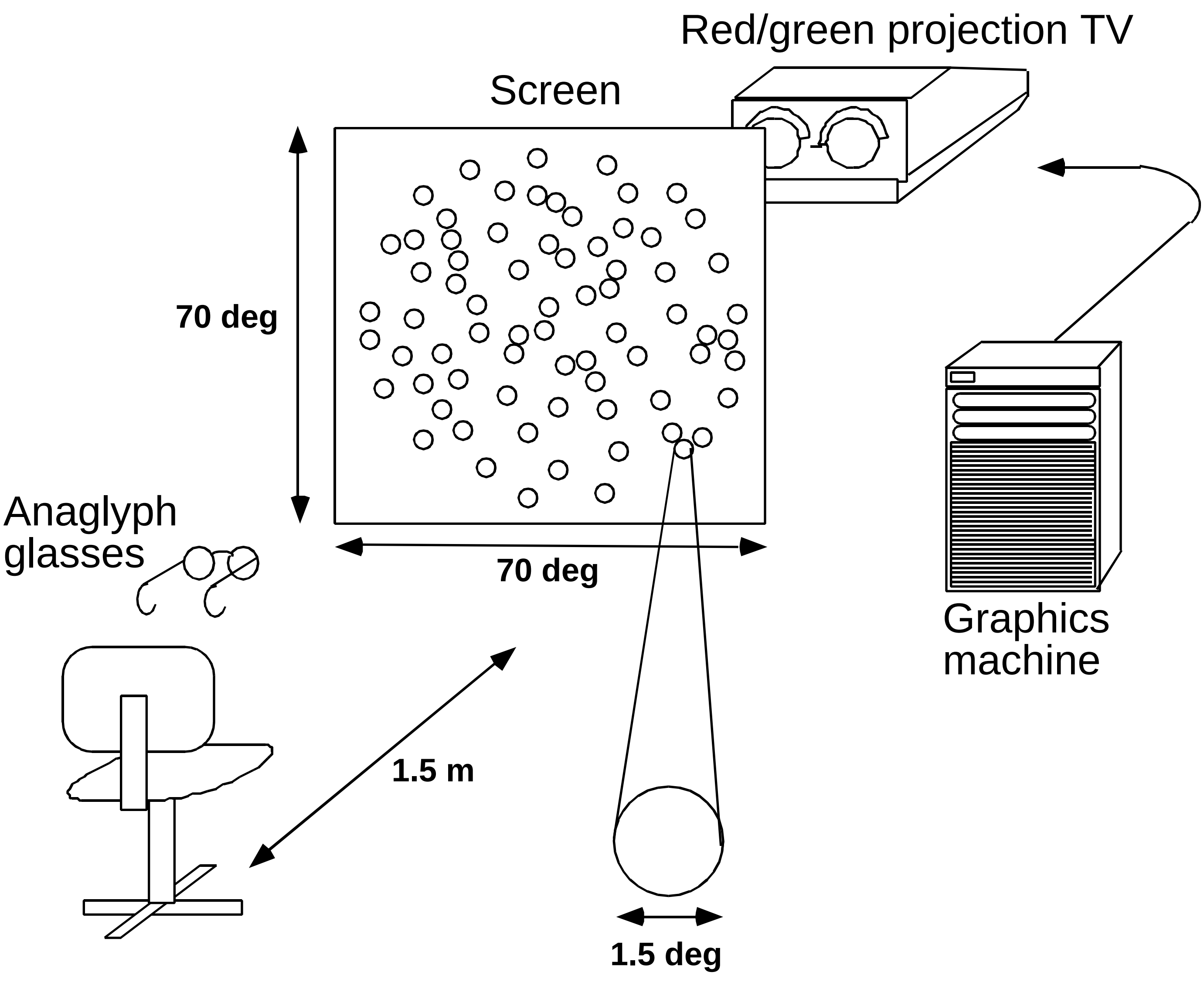


Fig 3, van Ee & erkelens

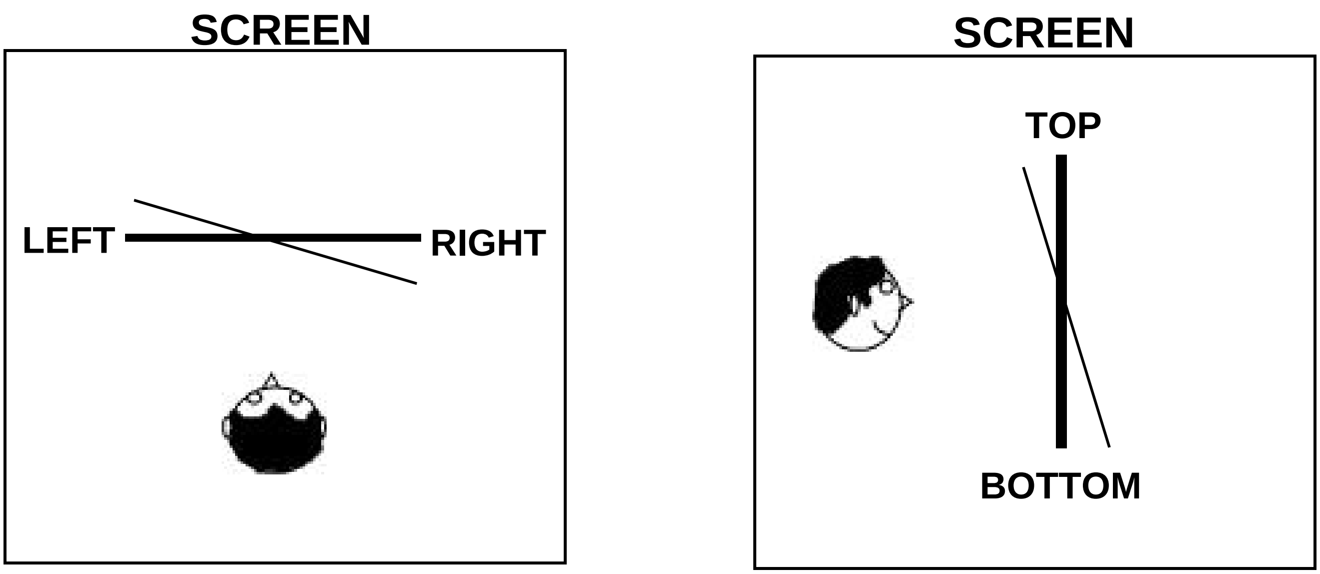


Fig 4, van Ee & erkelens

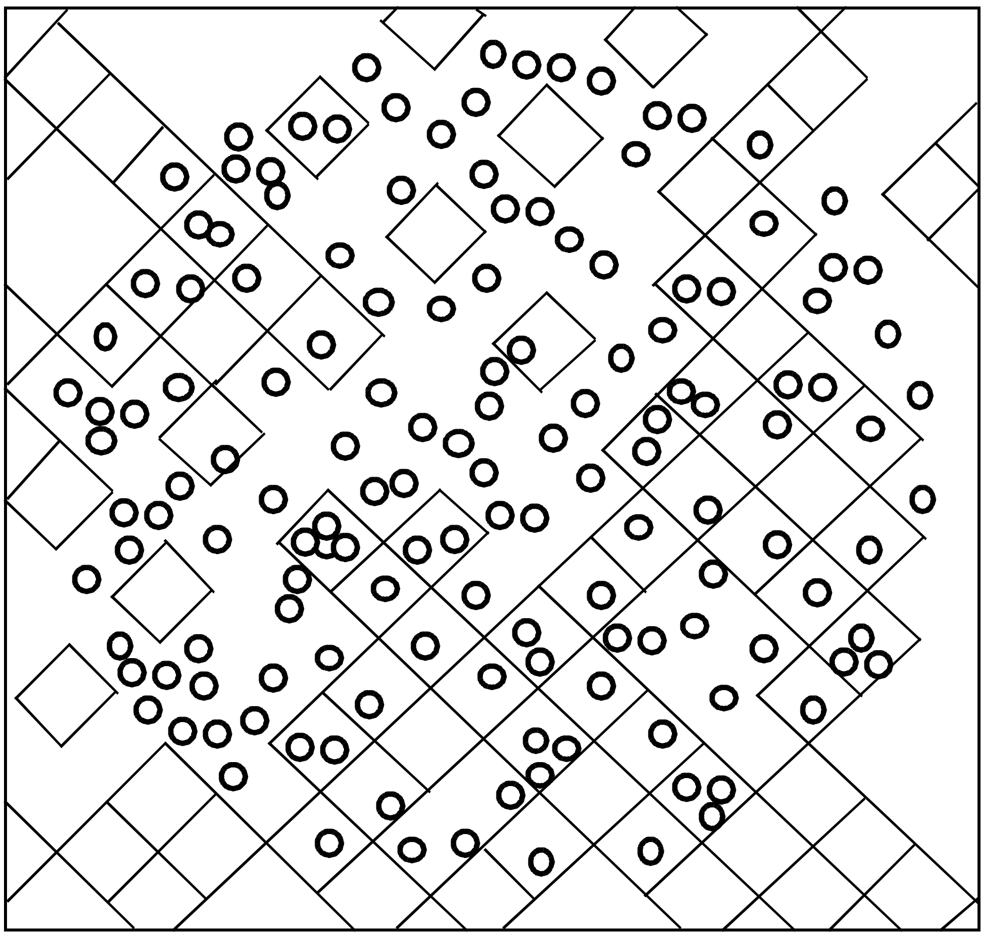


Fig 5, van Ee & erkelens

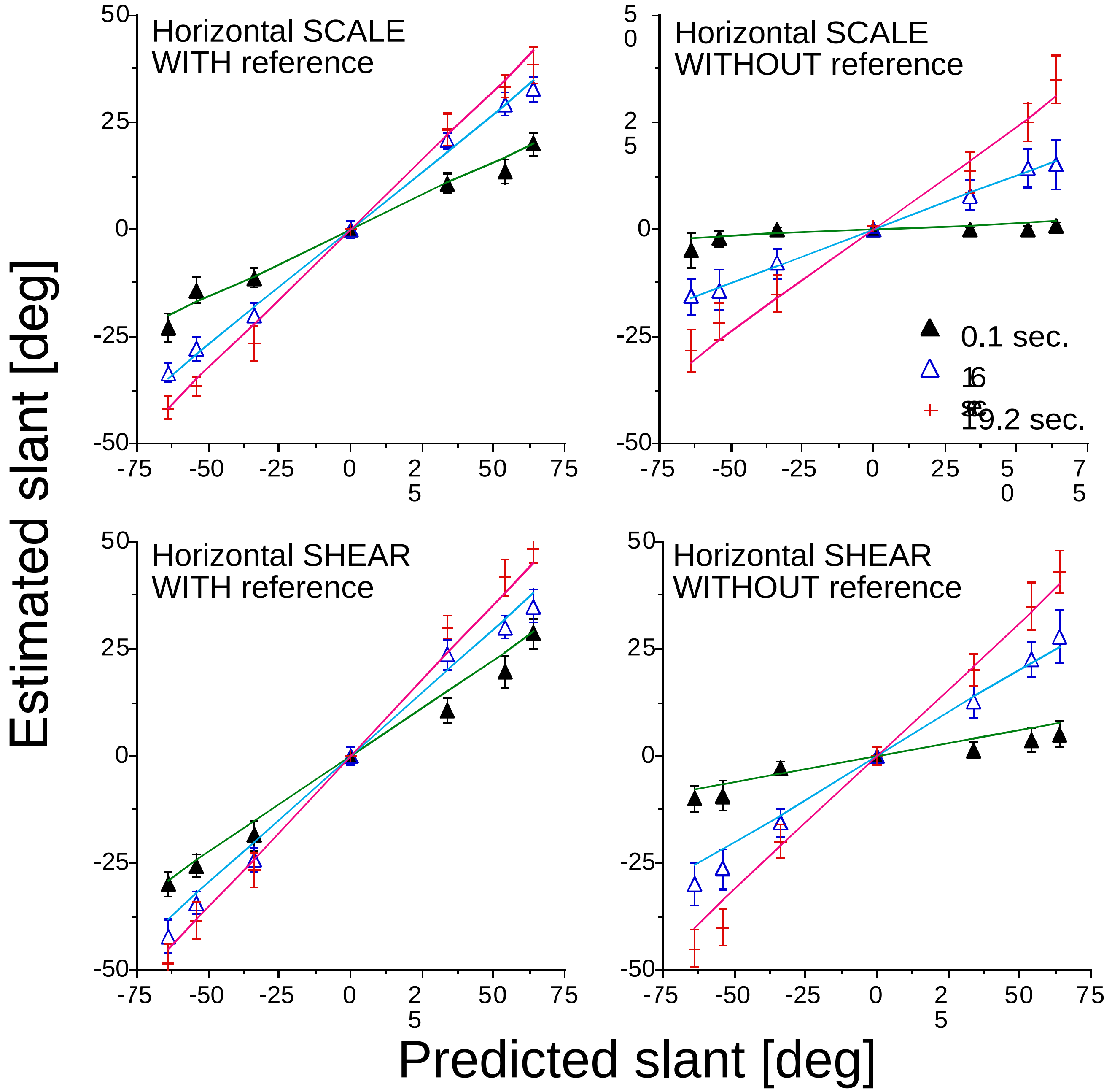


Fig 6, van Ee & erkelens

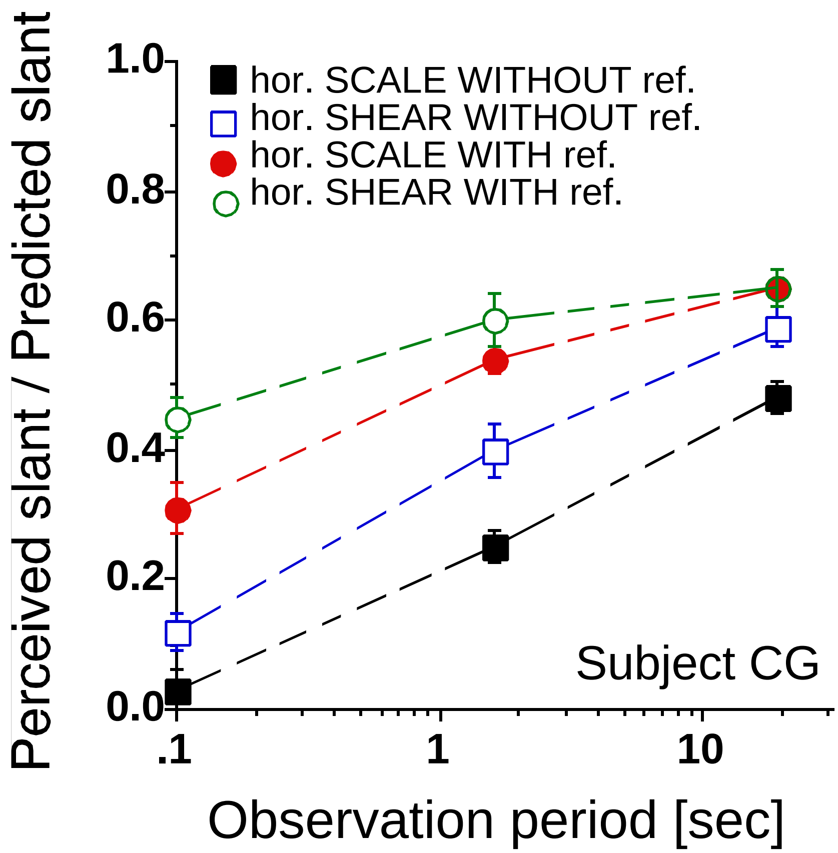


Fig 7, van Ee & erkelens

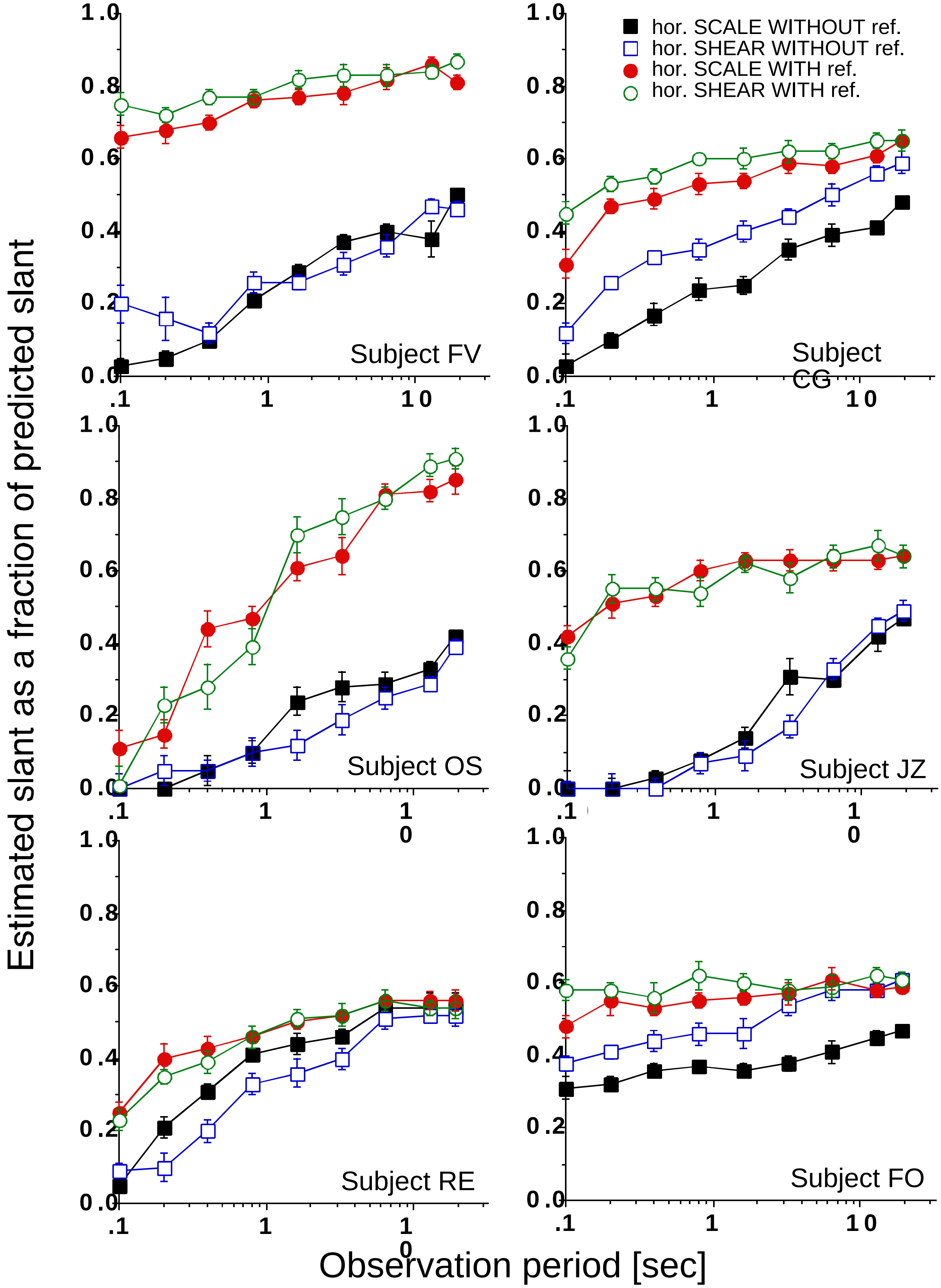


Fig 8, van Ee & erkelens