

# Temporal aspects of stereoscopic slant estimation: An evaluation and extension of Howard and Kaneko's theory

Raymond van Ee (1,2) & Casper J. Erkelens (2)

<sup>1</sup> Vision Science Program, University of California, Berkeley

<sup>2</sup> Helmholtz Institute, Princetonplein 5. 3584 CC, Utrecht, The Netherlands

email: [r.vanee@phys.uu.nl](mailto:r.vanee@phys.uu.nl)

**Abstract.** We investigated temporal aspects of stereoscopically perceived slant produced by the following transformations: horizontal scale, horizontal shear, vertical scale, vertical shear, divergence and rotation, between the half-images of a stereogram. Six subjects viewed large field stimuli (70 deg diameter) both in the presence and in the absence of a visual reference. The presentation duration was: 0.1, 0.4, 1.6, 6.4 or 25.6 s. Without reference we found the following: Rotation and divergence evoked considerable perceived slant in a number of subjects. This finding violates the recently published results of Howard and Kaneko. Slant evoked by vertical scale and shear was similar to slant evoked by horizontal scale and shear but was generally less. With reference we found the following: Vertical scale and vertical shear did not evoke slant. Slant due to rotation and divergence was similar to slant due to horizontal scale and shear but was generally less. According to the theory of Howard and Kaneko, perceived slant depends on the difference between horizontal and vertical scale and shear disparities. We made their theory more explicit by translating their proposals into linear mathematical expressions that contain weighting factors that allow for both slant evoked by rotation or divergence, subject-dependent underestimation of slant and other related phenomena reported in the literature. Our data for all stimulus durations and for all subjects is explained by this *unequal-weighting extension of Howard and Kaneko's theory*.

## 1 Introduction

In a recent series of papers, Howard and Kaneko proposed an interesting and attractive theory to explain how the visual system might utilize the differences in what they call *size* disparities and *shear* disparities to determine the perceived slant of stereoscopically-defined surfaces (Howard & Kaneko, 1994; Kaneko & Howard, 1996; 1997). More precisely, Kaneko and Howard (1996) reported that perceived slant about the vertical axis depends on the difference between horizontal *size* disparities derived locally and vertical *size* disparities derived regionally. In a subsequent paper (Kaneko & Howard, 1997), they reported that perceived slant about the horizontal axis depends on the difference between horizontal *shear* disparities derived locally and vertical *shear* disparities derived globally over the whole visual field<sup>1</sup>.

The transformations between the half-images of a stereogram that give rise to *size* disparities are horizontal scale and vertical scale. *Shear* disparities are caused by horizontal shear and vertical shear. In figure 1, the relevant transformations are defined using the nomenclature of Koenderink and van Doorn (1976)<sup>2</sup>. Horizontal scale plus the same magnitude of vertical scale of the same polarity is identical to divergence (Ogle, 1950; Koenderink & van Doorn, 1976). Horizontal shear plus the same magnitude of vertical shear is identical (infinitesimally) to rotation (Koenderink & van Doorn, 1976; Gillam & Rogers, 1991). Howard and Kaneko's theory is a development from the work of Ogle (1950), Koenderink and van Doorn (1976), Gillam and Rogers (1991), and van Ee and Erkelens (1993). However, none of these authors recognized the local/regional/global aspects as formulated by Howard and Kaneko.

---

<sup>1</sup> The literature is lacking in giving strict definitions of the terms global, regional and local. The reason for this lack is that these terms are not defined by a physical size. They should be defined operationally. Global (often called whole-field) might be referred to as the set of all visible stimuli; Regional as a significant part of the visible stimuli; Local as a non-significant part of the visible stimuli. Note that the consequence of defining them operationally is that in one experiment a local area can be larger than a regional area in another experiment.

<sup>2</sup> This nomenclature originates from vector field theories commonly used in physics. Where we use scale, Kaneko and Howard (and also Ogle, 1950) use magnification; where we use divergence, they use overall-size. They use inclination where we use 'slant about the horizontal axis' (Stevens, 1983). In the literature sometimes curl is used for rotation and def is (mis)used for scale and shear. def is an abbreviation of deformation.

*Slant about the vertical axis*--The magnitude of perceived slant about the vertical axis is related to the magnitude of horizontal scale. Ogle termed this the geometric effect because it is predicted by geometry. If one eye's image is vertically scaled, the subject perceives slant opposite in direction to that evoked by horizontal scale of the same image. Ogle termed this the induced effect. The magnitude of perceived slant evoked by whole-field divergence is usually small. Since divergence consists of equal magnitudes of horizontal and vertical scale disparities, Ogle suggested that slant perception about the vertical axis depends on the difference between horizontal scale and vertical scale disparities. Ogle's studies were focussed on global slant perception evoked by lenses. Koenderink and van Doorn decomposed the disparity field into the local spatial components of differential geometry. They showed that the divergence term does not contain information about slant. Their theory was developed for local estimates of slant. Evidence that the relevant measurements are not done just locally comes from the work of Stenton, Frisby and Mayhew (1984). They showed experimentally that vertical scale disparities are pooled globally. Rogers and Koenderink (1986) and Howard and Kaneko found that vertical scale disparities can also be processed regionally. Howard and Kaneko's theory is relevant for both local and non-local slant estimations.

*Slant about the horizontal axis*--Perceived slant about the horizontal axis is related to the magnitude of horizontal shear and vertical shear (this slant is opposite in polarity than slant evoked by horizontal shear). Unlike slant evoked by vertical shear, slant evoked by horizontal shear can be predicted by geometrical principles (van Ee and Erkelens, 1996a). Usually, rotation of whole-field stimuli evokes only small magnitudes of perceived slant (Howard & Zacher, 1991). Because small rotations consist of equal magnitudes of horizontal and vertical shear, Howard and Kaneko suggested that slant perception about the horizontal axis depends on the difference between horizontal shear (derived locally) and vertical shear (derived globally over the whole binocular field). Gillam and Rogers (1991) (see also Cagenello and Rogers, 1990; Rogers, 1992) recognized that rotation transforms vertical shear into horizontal shear (figure 1). This is particularly interesting because, consequently, cyclovergence of the eyes transforms a vertically sheared half-image into a horizontally sheared half-image, and vice versa. However, Kaneko and Howard (1997) reported that the rotational state of the eyes hardly influences perceived slant evoked by vertical shear. From Koenderink and van Doorn's theory, it follows that slant depends on the difference between horizontal and vertical shear disparities; they showed mathematically that rotation does not contain information about slant. However, Koenderink and van Doorn's theory is only valid for infinitesimal small smooth surface patches.

*Slant in the presence of a visual reference*--The results of Gillam and Rogers (1991)<sup>3</sup> and van Ee and Erkelens (1993) showed that Koenderink and van Doorn's local theory fails to explain experimental results when a visual frame of reference is present<sup>4</sup>. The reason is that both vertical scale and shear do not evoke slant of a surface when an untransformed reference surface is present. Van Ee and Erkelens (1995) showed that slant about oblique axes in the presence of a visual reference can be described, both theoretically and experimentally, by a single linear combination of horizontal shear and horizontal scale. Howard and Kaneko's theory is valid in the presence of a reference because in their theory horizontal scale and shear are derived locally whereas vertical scale and shear are derived non-locally.

---

<sup>3</sup> A reviewer stated that Gillam and Rogers used whole-field stimuli. This is not correct. In fact their slant measurement device was visible during the presentation of the stimuli. This device consisted of an illuminated Meccano wheel (subtending half of their stimulus size) that was binocularly visible. There were also four markers ('for binocular alignment') visible in their experiment. In the discussion of their paper Gillam and Rogers explicitly stated that 'details in the room were visible and apparently provided a sufficient frame'. Whole-field or global stimuli are not limited by boundaries which consist of disparity steps. A whole-field analysis can not be applied to individual objects within a scene consisting of other objects (like a comparison stimulus).

<sup>4</sup> Gillam and Rogers did not attribute their results to the presence of a visual reference. Instead they stated that perceived slant was predicted from the orientation disparity at the vertical meridian. However, the concept of orientation disparity is unnecessary as has been shown by van Ee and Erkelens (1995). These latter authors found that positional disparities are sufficient to show the failure of Koenderink and van Doorn's theoretical predictions for perceived slant.

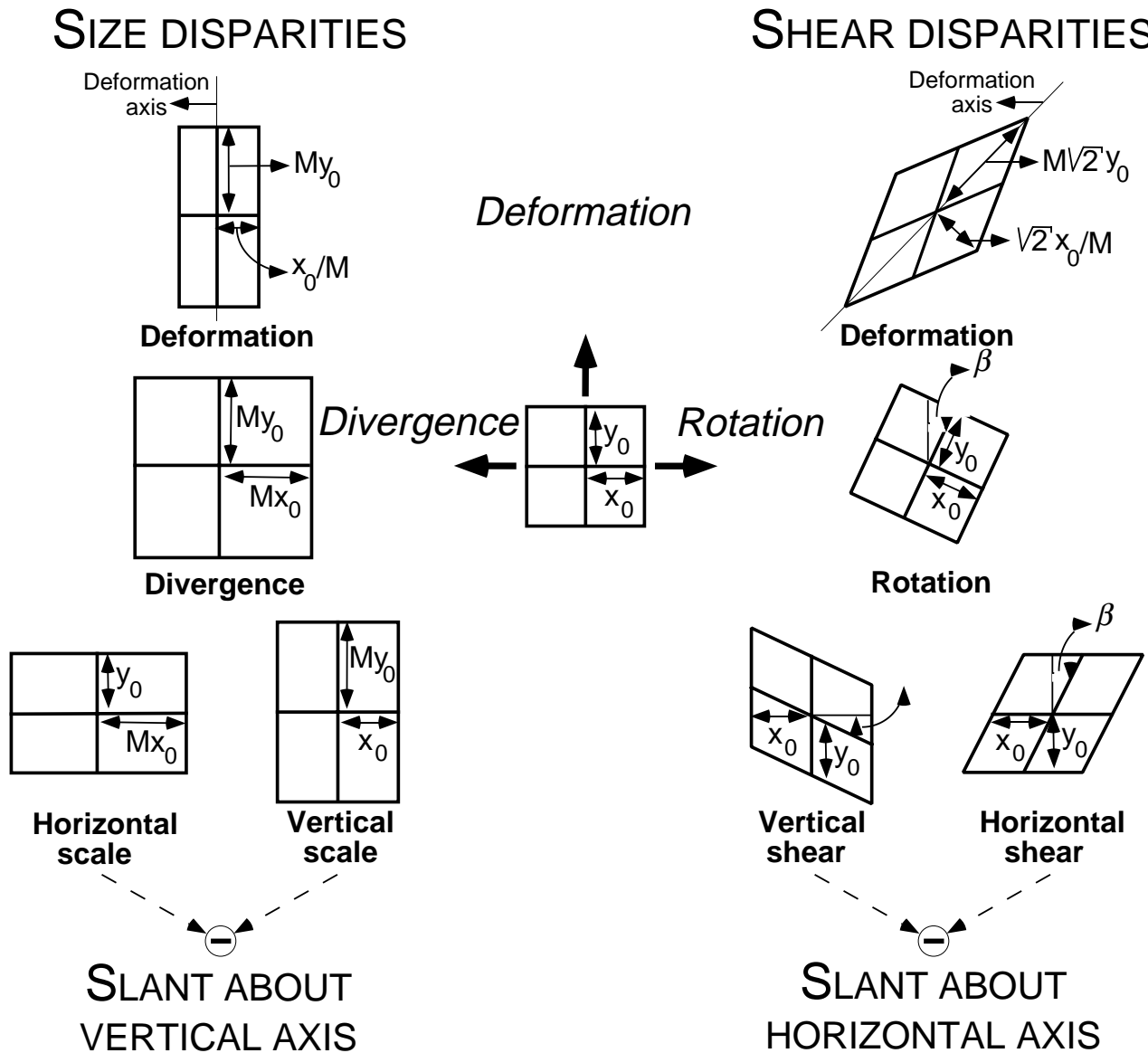


Figure 1: Definition of relevant linear transformations. The middle figure could be regarded as one of the half-images in a stereogram whereas the outer figures represent the transformed half-image. Provided that these half-images are fused, the size (shear) transformations in the left (right) part of the figure are associated with perceived slant about the vertical (horizontal) axis. The heart of the figure consists of divergence, deformation and rotation. Any linear transformation can be obtained by a combination of these three operations. In the present paper and in the work of Howard and Kaneko, the transformations horizontal scale, vertical scale and divergence as well as horizontal shear, vertical shear and rotation are used. Scale is a linear combination of deformation and divergence. Shear is a linear combination of (a different) deformation and rotation. Horizontal scale plus the same magnitude of vertical scale of the same polarity is identical to divergence. Infinitesimal small magnitudes of horizontal shear and vertical shear combine to rotation. Perceived slant about the vertical axis depends on the difference between horizontal and vertical scale disparities. Perceived slant about the horizontal axis depends on the difference between horizontal and vertical shear disparities. The amount of shear and rotation is expressed by the angle  $\beta$ . The amount of scale and divergence is expressed by the factor  $M$ . Because in many publications of the last decade deformation is an often (mis)used term we emphasize that none of the transformations used in the present paper (or in Howard and Kaneko's papers) should be confused with pure deformation. Deformation is historically defined as a linear combination of expansion and contraction in orthogonal directions with conservation of area. Two examples of deformation are shown in the top part of the figure. See the original papers of Koenderink and van Doorn for proper definitions or van Ee and Erkelens (1995) for a summary.

There are reports in the literature that cannot be explained by Howard and Kaneko's theory. First, unlike estimates of slant evoked by horizontal scale, estimates of slant evoked by vertical scale do not vary with observation distance (Gillam, Chambers & Lawergren, 1988) for stimuli presented straight ahead (Backus & Banks, 1997). Howard and Kaneko's theory does not take into account changes in distance (and eccentricity) which means that their theory cannot be complete. Second, almost all subjects in the literature show a large underestimation of perceived slant. Howard and Kaneko also found strong underestimations of slant but gave no explanation. Third, several studies have reported large differences across subjects in whole-field slant estimations. Fourth, there are also large differences within individual subjects between slant evoked by horizontal and vertical scale (and shear) transformations (Gillam, Chambers & Lawergren, 1988). With regard to the geometric and the induced effect, Ogle (1950, page 195) stated: *Some subjects show great differences between the two effects, others very small ones*. Fifth, there is evidence that whole-field rotation does evoke perceived slant (Collewijn, van der Steen & van Rijn, 1991; Swash, Rogers, Bradshaw & Cagenello, 1995). We also found a number of subjects who perceived considerable slants evoked by whole-field rotation and divergence (van Ee & Erkelens 1996c). These observations violate one of the premises on which Howard and Kaneko's theory is based: Howard and Kaneko (1994) reported that rotation does not evoke perceived slant and Kaneko and Howard (1996) showed that even for large displays there is only a small perceived slant evoked by divergence which they attributed to a larger contribution of horizontal scale disparity than vertical scale disparity. Ogle (1950, page 195) also found an unequal weighting of horizontal and vertical scale disparities: *Subjects with whom the induced effect is found smaller than the geometric effect usually see a distortion of the leaf room when an overall magnification lens is placed before one eye*<sup>5</sup>. From the work of Howard and Kaneko it is not clear whether there were individual subjects who did perceive considerable slants from divergence and rotation because they presented means across three subjects only<sup>6</sup>.

Howard and Kaneko allowed subjects to take as long as they liked to make the slant estimation. In practice they took about 10 seconds. However, the perceived slant evoked by horizontal scale or horizontal shear disparities (Gillam, Chambers & Russo, 1988; van Ee & Erkelens, 1996a), vertical shear disparities (Cagenello & Rogers, 1990; Rogers, 1992) and vertical scale disparities develops over time (van Ee & Erkelens, 1996c). If subjects view a stimulus for more than, say, 10 seconds, slant estimation is far more veridical than if they view it for, say, one second. It is possible that subjects need much more time to perceive slant from vertical disparity than from horizontal disparity. In addition, in daily life, slants have often to be judged in shorter time scales. It is not clear whether Howard and Kaneko's theory holds for well-defined short (more realistic) presentation durations.

We therefore investigated the temporal aspects of estimation of slant evoked by scale, shear, rotation or divergence between the half-images of a stereogram. To allow for the unequal contribution of horizontal and vertical disparities, for slant evoked by rotation and divergence, for the underestimation of slant and for differences between subjects we present a model in which we incorporate weighting factors.

## 2 Theory

In the spirit of Howard and Kaneko's ideas we formulate three possible linear models. These models are shown in figure 2. An essential feature of the three models is the independent weighting of horizontal and vertical disparities. The difference between the three models becomes clear if they are expressed mathematically (see figure 2). It is not clear whether Howard and Kaneko had a linear model in mind but the linear model of figure 2a comes closest to their formulation. In this model, estimated slant 'is a function of' (in other words, 'depends on', as they formulate it) the difference between horizontal and vertical disparity. However, they subtracted (under)-estimated slants, which means that they implicitly used the model of figures 2b or 2c. The model of figure 2b will be used for interpreting our data. In the rest of this paper, we express estimated slant as the linear summation of weighted slants predicted from shear and scale disparities (model of figure 2b):

$$[\text{estimated slant}]_{\text{about hor axis}} = W_{\text{hor shear}}(t) [\text{predicted slant}]_{\text{hor shear}} - W_{\text{ver shear}}(t) [\text{predicted slant}]_{\text{ver shear}}$$

<sup>5</sup> The leaf room is a room with a minimum of empirical cues. The artificial vines stapled to the inside of the room provide many contours to stimulate stereopsis. However, with continued monocular observation the room appears to lose its shape entirely.

<sup>6</sup> Very recently, Howard (personal communication, February 1997) indicated that his laboratory has 'come across several subjects who show little effect of vertical shear disparity'

$$[\text{estimated slant}]_{\text{about ver axis}} = W_{\text{hor scale}}(t)[\text{predicted slant}]_{\text{hor scale}} - W_{\text{ver scale}}(t)[\text{predicted slant}]_{\text{ver scale}}$$

where hor and ver denote horizontal and vertical, respectively.  $W$  denotes the weighting factor associated with a particular transformation. The geometrically-derived relationship between the angle of horizontal shear disparity  $\beta$ ; the angle between corresponding vertical lines of the two half-images, see figure 1) and predicted slant ( $F$ , see fig. 2) about the horizontal axis is:

$$\text{predicted slant} = F(\beta) = \arctan(\tan \beta \cdot \frac{z_0}{I}),$$

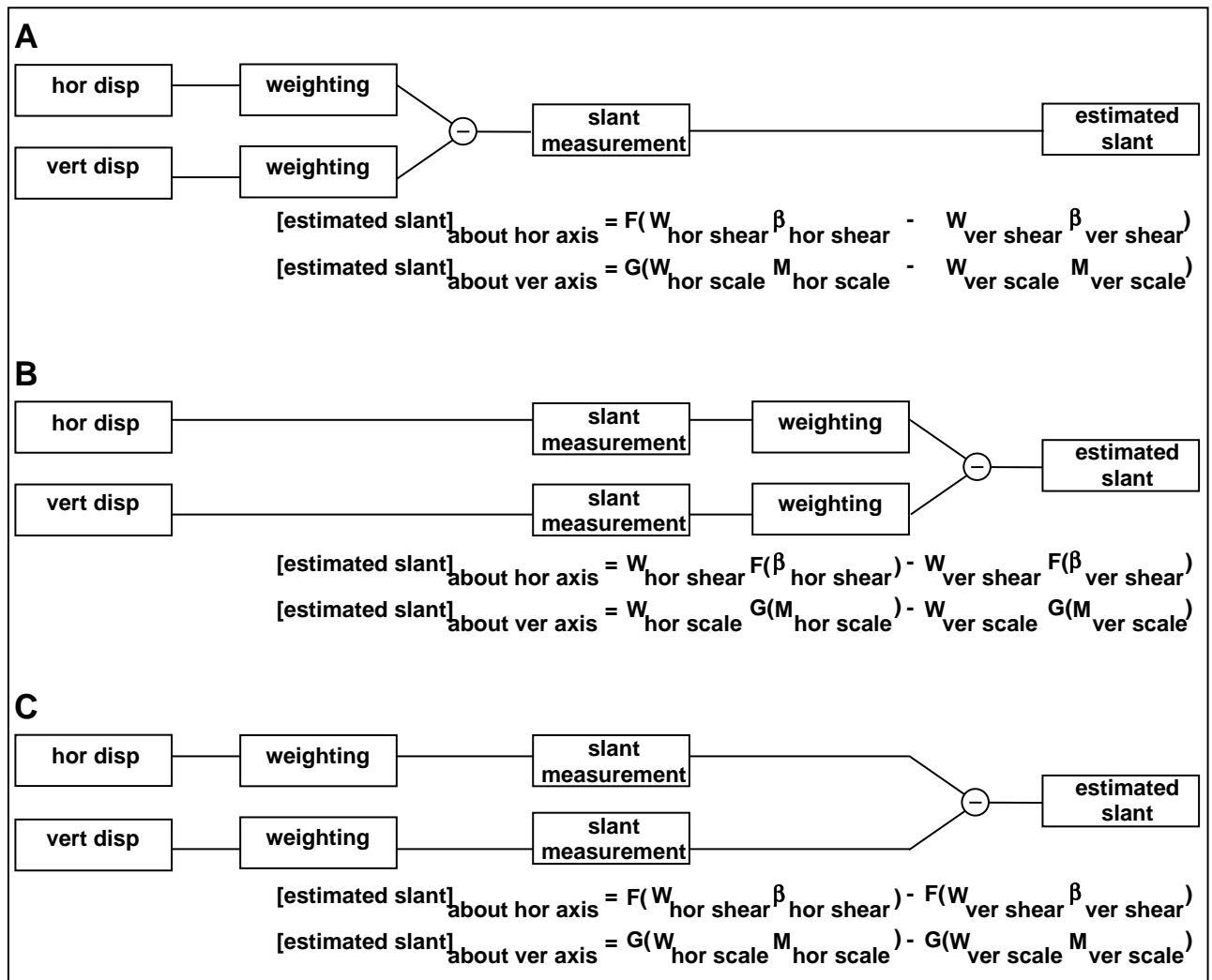


Figure 2: Three possible linear models and their mathematical expressions in the spirit of Howard and Kaneko's theory. In the 'slant measurement' block the slant is determined by the visual system as a function ( $F$ ,  $G$ ) of the (possibly weighted as in models a and c) magnitude of shear or scale. The functions  $F$  and  $G$  are given in the text.

where  $z_0$  denotes viewing distance and  $I$  denotes interocular distance. According to Howard and Kaneko's reasoning, the slant predicted by vertical shear disparities can be described by the same function  $F$ . In the case of vertical shear,  $\beta$  is the angle between corresponding horizontal lines of the two half-images (see fig. 1). The geometrically-derived relationship between the magnitude of horizontal scale disparity ( $M$ ; the horizontal size ratio of the two half images relative to each other, see fig. 1) and predicted slant ( $G$ , see fig. 2) about the vertical axis is:

$$\text{predicted slant} = G(M) = \arctan\left(\frac{M-1}{M+1} \cdot \frac{2z_0}{I}\right).$$

Predicted slant evoked by vertical scale disparities can be expressed by the same function  $G$ , according to Howard and Kaneko<sup>7</sup>. In the case of vertical scale,  $M$  is the vertical size ratio of the two half-images relative to each other (see fig. 1)<sup>8</sup>. The weighting factors  $W$  are a function of time.  $W_{\text{hor shear}}(t)$  and  $W_{\text{hor scale}}(t)$  are derived locally (Kaneko & Howard, 1996; 1997). When the horizontal disparity information is used effectively for slant perception without underestimation,  $W_{\text{hor shear}}(t)$  and  $W_{\text{hor scale}}(t)$  are unity. In the case of the transformations horizontal scale and shear there would be flat lines at unity in our data figures (3 and 4) if an observer utilized the geometrically present disparity information perfectly without underestimation.  $W_{\text{ver scale}}(t)$  is derived regionally (Kaneko & Howard, 1996),  $W_{\text{ver shear}}(t)$  is derived globally (Kaneko & Howard, 1997). When the vertical disparity information is used effectively for slant perception without underestimation,  $W_{\text{ver shear}}(t)$  and  $W_{\text{ver scale}}(t)$  are unity. According to reasoning given in the Introduction we do not expect slant from rotation (divergence) whenever horizontal shear (scale) and vertical shear (scale) are both equally large and equally weighted. Whenever there is unequal weighting we expect the contribution of rotation (divergence) to the estimated slant to be equal to the difference between the contributions of horizontal shear (scale) and vertical shear (scale):

$$[\text{estimated slant}]_{\text{rotation}} = [W_{\text{hor shear}}(t) - W_{\text{ver shear}}(t)] \cdot [\text{predicted slant}]_{\text{hor shear}} \quad ,$$

$$[\text{estimated slant}]_{\text{divergence}} = [W_{\text{hor scale}}(t) - W_{\text{ver scale}}(t)] \cdot [\text{predicted slant}]_{\text{hor scale}} \quad .$$

In terms of weighting factors solely:

$$W_{\text{hor shear}}(t) - W_{\text{ver shear}}(t) - W_{\text{rotation}}(t) = 0 \quad ,$$

$$W_{\text{hor scale}}(t) - W_{\text{ver scale}}(t) - W_{\text{divergence}}(t) = 0 \quad .$$

### 3 Methods

This experiment extends the experiment described in the paper by van Ee and Erkelens (1996a). We used exactly the same hardware, stimulus generation, task for the subject and data analysis. The only differences were the transformations presented and the stimulus durations. The part of the methods which was identical to the earlier experiment is described only briefly.

#### 3.1 Apparatus and procedure

Stereograms were presented using red/green anaglyphs on a 70 deg wide flat screen at a distance of 1.5 m in front of the subject. Head movements were restricted by a chin-rest and a skull-rest. Care was taken to ensure that the interocular axis was parallel to the frontal screen. The subjects were free to make eye movements. Randomly distributed circles appeared on the screen in circular configurations (70 deg diameter). The small circles had diameters of 1.5 deg. The density of the small circles was such that they covered about 10 % of the stereogram. The task of the subject was to estimate the perceived slant evoked by the transformations presented. After each presentation, two lines (one fixed and one rotatable) appeared on the screen. By changing the computer-mouse position, subjects set the angle between the rotatable line and the fixed line; the angle represented the estimated slant. 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; only the stimulus was visible. In the series of trials with a visual reference, a large-field reference pattern, which covered the whole screen, was projected on the screen and the room was dimly lit to prevent depth contrast effects. The reference

<sup>7</sup> Note there is no direct geometrical relationship between predicted slant and vertical scale or shear under the given oculomotor conditions.

<sup>8</sup> Note that the relationship between horizontal scale and slant is slightly different from the one Ogle derived because he derived his relationship for the case that the horizontal scale was evoked by a lens. [See van Ee and Erkelens (1996a) for a derivation of the given relationships].

consisted of a transparent cross-hatched pattern. The cross-hatched pattern was made up of a field of adjacent squares with diagonals of 15 deg and a density of about 60 %.

Specific for this experiment are the transformations presented and the presentation durations. The transformations of the green part relative to the red part of the stereogram were horizontal scale, horizontal shear, vertical scale, vertical shear, divergence or rotation. Horizontal scale, vertical scale and divergence varied between -6.0 % and 6.0 % in four steps (with a step-size of 3.0 %) and horizontal shear, vertical shear and rotation varied between -3.3 deg and 3.3 deg (again in four equal steps)<sup>9</sup>. The magnitudes of the scale and shear transformations were chosen such that they were identical to each other with regard to the magnitude of predicted slant.

The durations of the presentations were in random order: 0.1, 0.4, 1.6, 6.4 or 25.6 s. Each trial was repeated seven times. In all, each subject completed 2100 slant judgments: 5 presentation durations, 6 transformations (horizontal scale and shear, vertical scale and shear, divergence and rotation), 2 conditions (with and without visual reference), 5 magnitudes of transformations (-6.0, -3.0, 0.0, 3.0, 6.0 % or -3.3, -1.6, 0.0, 1.6, 3.3 deg) and 7 repetitions. The tests began with a series of randomly intermixed trials without visual reference (1050 trials). The following day, the subjects repeated the same series of trials but with visual reference. Each series lasted about two hours.

### 3.2 Subjects

Six subjects (4 males and 2 females, ages 23-44 yr) took part in the experiment. No feedback was given about the results. Except for the author (RE) the subjects were not aware of the purpose of the experiment. Three subjects (KY, BJ, IH) were inexperienced with respect to stereoscopic experiments in general. Subjects RE and JZ were very experienced in stereoscopic slant estimation experiments. Four subjects (EC, KY, JZ, and RE) showed refraction anomalies which were corrected by their own glasses or contact lenses.

### 3.3 Data analysis

Previous work has shown that the function relating estimated and predicted slant is linear. The slope of this function was used to characterize performance in each condition. The values of the slopes are presented in the figures. Because the transformations, vertical scale, vertical shear, divergence and rotation as used in the experiment do not mimic objects in the real world there is no geometrical relationship between the magnitudes of these transformations and slant predicted from these magnitudes. Therefore, our figures show the *normalized* estimated slant as a fraction of the predicted slant. Normalized means that the estimated slant is divided by the predicted slant of horizontal scale or horizontal shear. As an example: We present 3 % vertical scale for a presentation duration of 25 seconds. Say the estimated slant is -15 deg. The geometrically predicted slant evoked by 3 % horizontal scale is 34 (for an observation distance of 150 cm and an interocular distance of 6.5 cm)<sup>10</sup>. The estimated slant divided by the predicted slant is  $\$-15/34 = -0.44\$$ . Furthermore, in order to be able to compare in one figure the results of vertical shear and vertical scale with the results of horizontal shear, horizontal scale, rotation and divergence we determined the absolute value of this fraction for vertical shear and scale (but not in the case of rotation and divergence). Thus, in our example the normalized estimated slant is 0.44.

---

<sup>9</sup> This time, the range of transformations is not as large (-9 % to 9 %; -4.9 deg to 4.9 deg) as used in the experiment of van Ee and Erkelens (1996a) where only horizontal scale and shear were presented. The reason is that fusional ranges in the vertical direction are not as large as they are in the horizontal direction; care was taken to prevent fusion problems.

<sup>10</sup> In general, a positive magnitude of horizontal scale or shear of the right eye's half-image relative to the left eye's evokes a positive angle of perceived slant. In contrast, a positive magnitude of vertical scale or shear of the right eye's half-image relative to the left eye's evokes a negative angle of perceived slant.

### In the absence of the visual reference

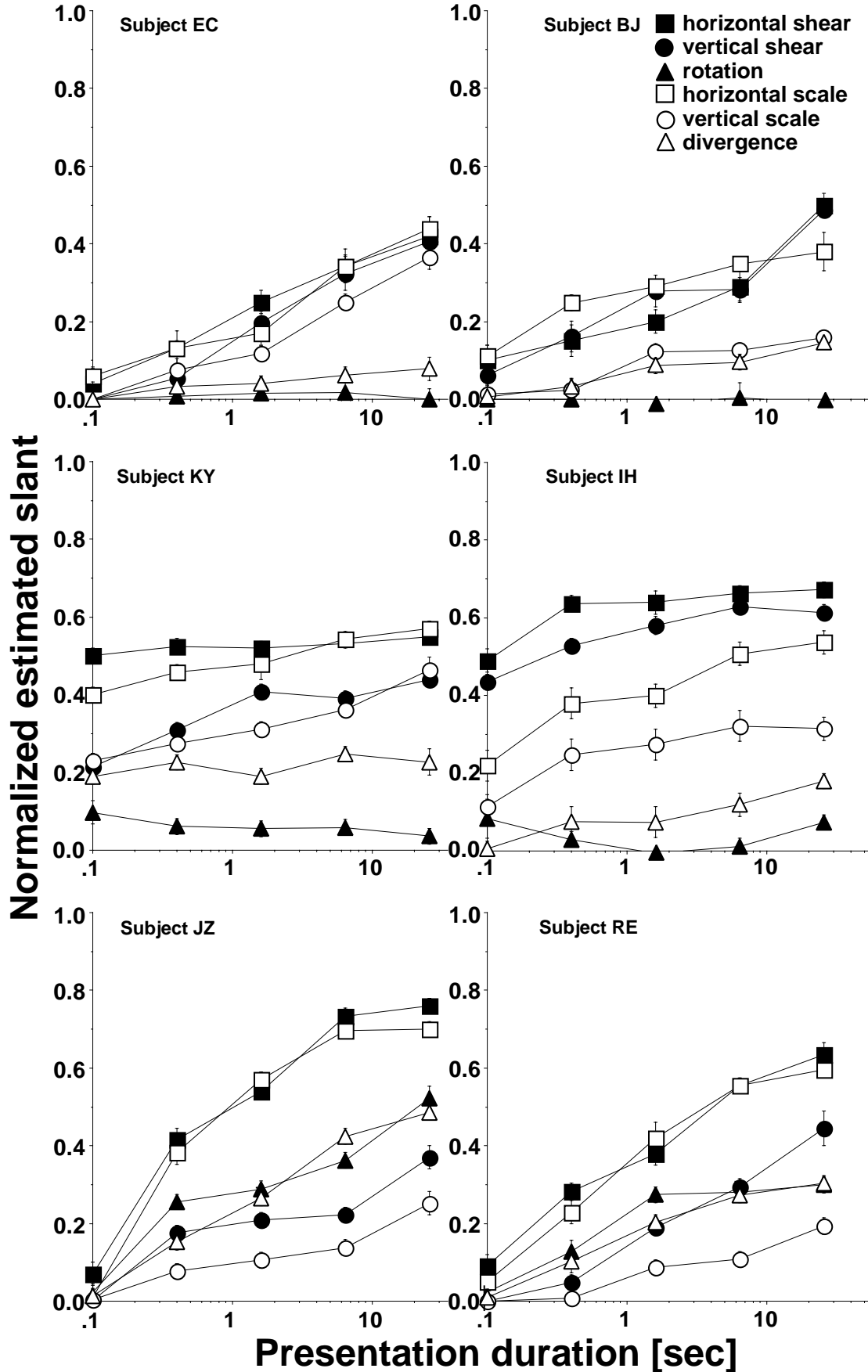




Figure 3: Normalized estimated slant as a function of the presentation duration when the stimuli were viewed in the *absence* of the visual reference. The error bars represent the standard deviation over five repetitions.

#### 4 Results

The normalized estimated slant for each presentation duration in the absence of a visual reference is presented in figure 3. Slant estimates developed over time for each transformation but in different ways. Rotation and divergence evoked significant magnitudes of estimated slant in a number of subjects but not in others. Especially in subjects RE and JZ slant evoked by rotation and divergence was estimated to be even larger than slant evoked by vertical shear and vertical scale for most of the presentation durations. As mentioned above, both RE and JZ were very experienced in slant estimation experiments. Estimated slant due to vertical scale and shear was similar to slant evoked by horizontal scale and shear but was generally smaller. For brief presentation durations (especially of the order of 1 second or less) slant was strongly underestimated. Figure 4 shows the results of the same experiment, but now in the presence of a visual reference. Vertical scale and vertical shear led to estimated slants of zero deg. Slant due to rotation and divergence was similar to the slant estimated from horizontal scale and shear but was generally less.

The transformation pairs (horizontal scale, horizontal shear), (vertical scale, vertical shear) and (rotation, divergence) generally evoked similar slant development in individual subjects over time. Perception of slant in the presence of a visual reference develops faster and to a higher level than without visual reference. Subjects' responses differed far more from each other when stimuli were viewed without a reference than when they were viewed with a reference. In figures 3 and 4 we showed the normalized estimated slant, which means that the sign of the perceived slant is not shown. Rotation and divergence, apart from two expectations (see in figure 3 subjects BJ and IH for a presentation duration of 1.6 s) always evoked an estimated slant with the same sign as estimated slant evoked by horizontal shear and horizontal scale. Thus, horizontal disparity was almost always weighted more than vertical disparity.

The results of subjects EC (for both scale and shear) and BJ and IH (only for shear) clearly fit Howard and Kaneko's theory to a large extent for all tested presentation durations. They perceived little or no slant from whole-field divergence (EC) and rotation (EC, BJ, IH) and an almost equally large slant from horizontal and vertical scale (EC) and shear (EC, BJ, IH). We calculated the weighting factors in order to fit the results of all subjects with Howard and Kaneko's theory for all measured presentation durations. In fact, the data given in figures 3 and 4 are the above-defined weighting factors, because normalized slant is estimated slant divided by predicted slant. Figure 5 shows the results of subtracting the weighting factors of vertical shear and rotation from horizontal shear, for slant settings in the absence (fig. 5a) and presence (fig. 5b) of a visual reference. Within the experimental uncertainties (standard deviations of about 0.15) the results of the subtractions are fairly close to zero for almost all subjects. Apparently, the weighting factor of rotation is almost the same as the difference between the weighting factors of horizontal shear and vertical shear for the entire range of presentation durations. Figure 5 also shows the subtraction of the weighting factors of vertical scale and divergence from horizontal scale, for slant settings in the absence (fig. 5c) and presence (fig. 5d) of a visual reference for each individual. The weighting factor of divergence is almost the same as the difference between the weighting factors of horizontal scale and vertical scale for the entire range of presentation durations. All of these results are consistent with the 'unequal weighting' extension of Howard and Kaneko's theory as formulated above. The deviations from zero are remarkably small. They are of the order of 7 deg slant, which is approximately equal to the standard deviations in subjects' judgments of the angle of perceived slant.

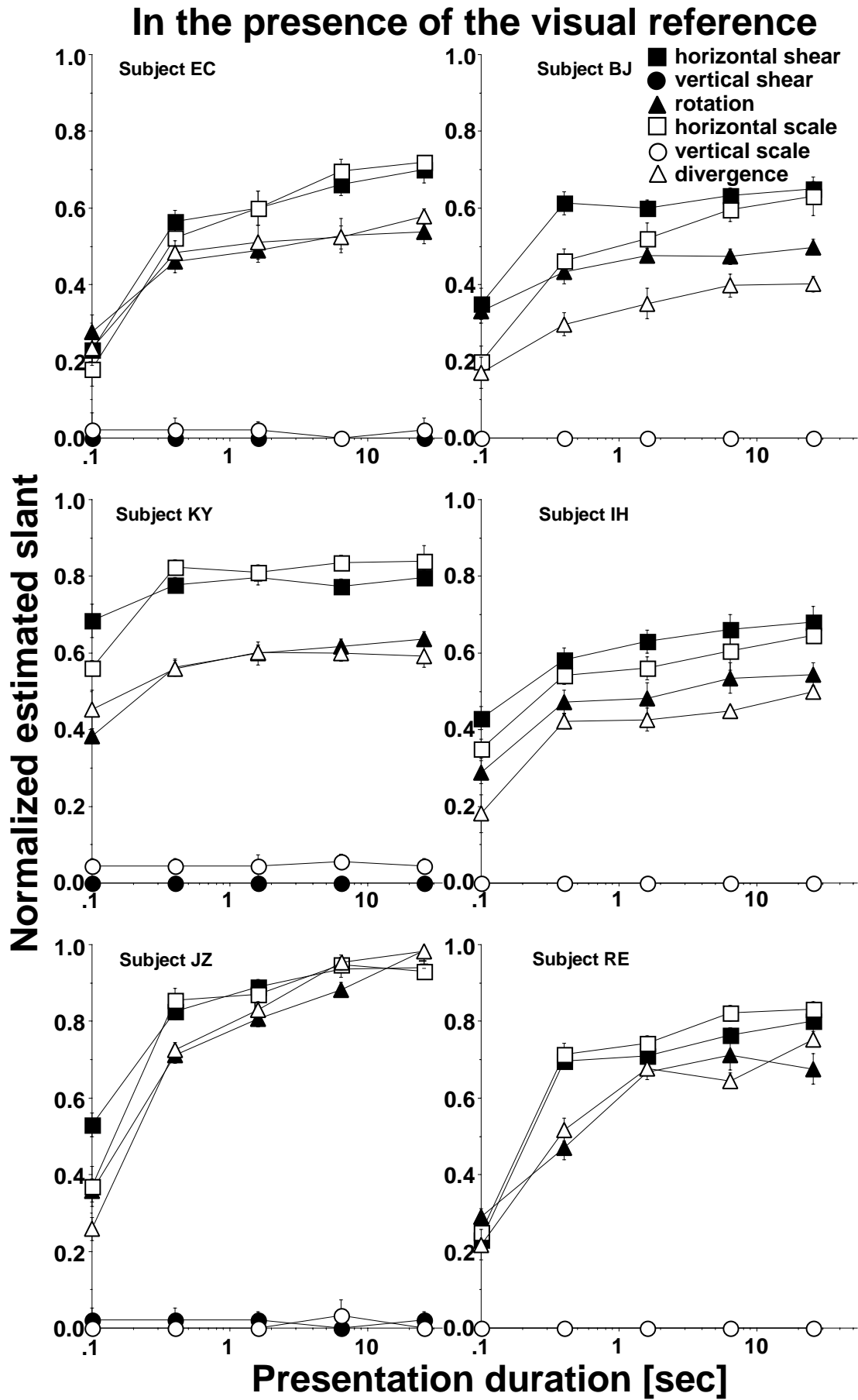


Figure 4: Same as figure 3 but in the *presence* of the visual reference. For the subjects BJ, IH and RE the symbols for vertical shear are hidden behind the symbols for vertical scale.

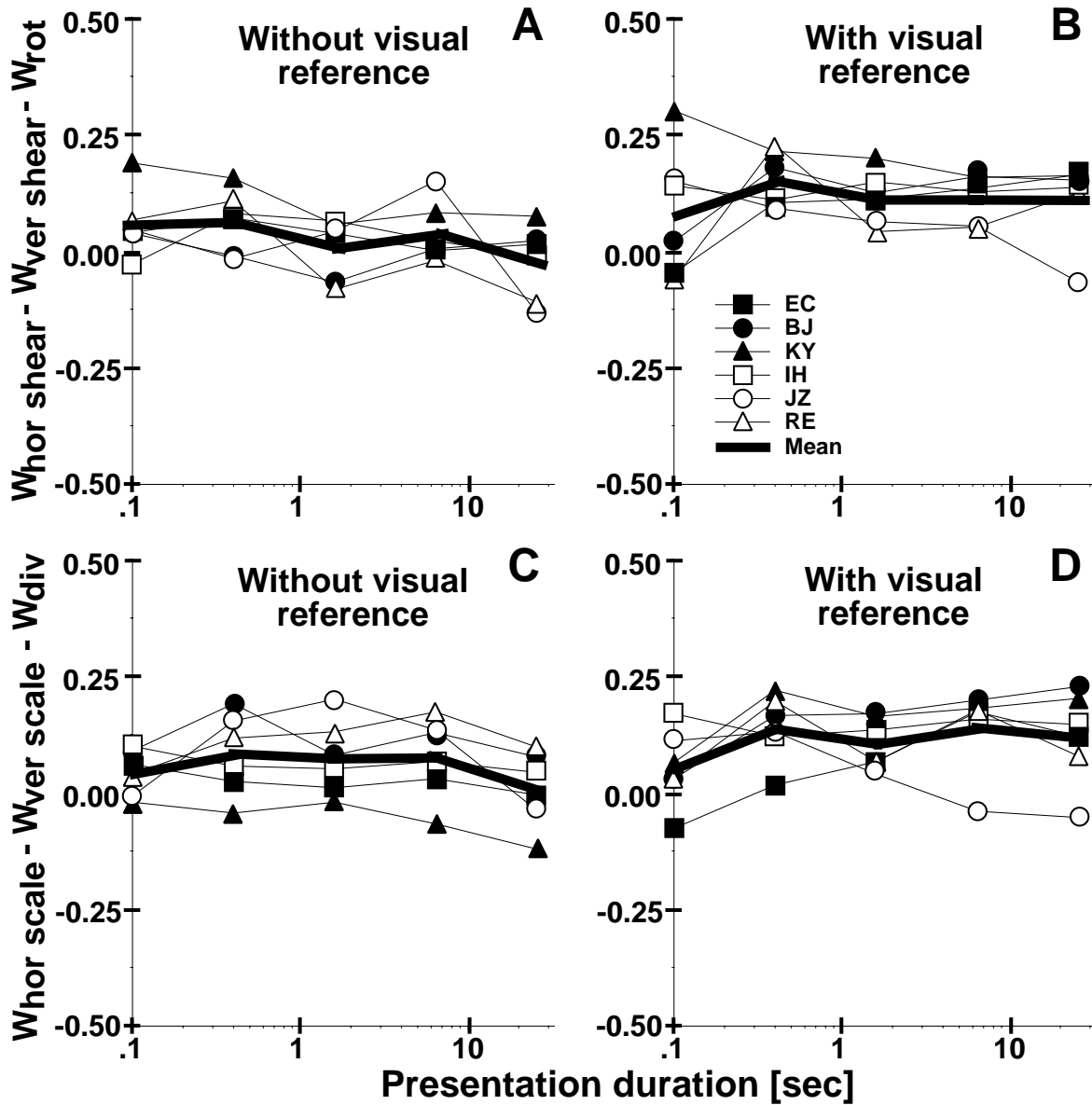


Figure 5: Subtraction of the weighting factors of the transformations presented for each subject and the mean of the six subjects. These weighting factors are determined from the data given in figures 3 and 4 for the complete range of presentation durations, both with and without visual reference. Hor shear and hor scale denote horizontal shear and horizontal scale, respectively. Ver shear and ver scale denote vertical shear and vertical scale, respectively. Div denotes divergence and rot denotes rotation. The standard deviation for each data point is about 0.15. A residue of 0.1 means that a rotation of, say, 3.3 deg is not perceived as a slant of zero deg but as a slant of 5.4 deg (because a horizontal shear of 3.3 deg corresponds to a predicted slant of 5.4 deg). In fact, the weighting factors of vertical shear and vertical scale are negative (see the end of the Methods section). However, to obtain the data in this figure all weighting factors of vertical scale and shear are taken to be positive such as given in figures 3 and 4.

## 5 Discussion

Although the weighting factors determined for the two experienced subjects (RE and JZ) are consistent with the above formulated predictions, the two subjects RE and JZ showed estimated slant from both rotation and divergence to be larger than estimated slant from vertical shear and vertical scale. As far as we know, this is the first time that such large slants evoked by whole-field rotation and divergence have been reported in the literature. As mentioned before, these results are not in accordance with the premises (see Introduction) of Howard and Kaneko's theory. Howard and Kaneko reported that estimated slant from rotation is zero and that estimated slant from divergence is very small. The results of subjects RE and JZ may have been due to a practice effect which enabled these subjects to perceive large slants evoked by divergence and rotation. Recent investigations in our lab (van Ee, Backus & Erkelens, 1996) have shown that with practice the majority of unpractised subjects show a considerable improvement in their estimates of slant evoked by horizontal scale and horizontal shear without a visual reference but not with a visual reference<sup>11</sup>. The improvement occurs even without feedback.

In Figure 6 we concentrate on the averaged weighting factors of the four inexperienced subjects. Figure 6a shows that, in the absence of a visual reference, rotation between the half-images of the stereogram evokes only little perceived slant. The magnitude of estimated slant due to rotation is similar to the difference between estimated slant from horizontal shear and estimated slant from vertical shear for the entire range of presentation durations. Perceived slant evoked by rotation deteriorates over time (fig. 6a). This has previously been observed by Rogers and co-workers (personal communication, April 1996). The fact that Howard and Kaneko (1994) and Kaneko and Howard (1997) did not find perceived slant from rotation could be due to the rather long presentation durations they used. Figure 6b shows that in the absence of a visual reference, divergence between the half-images of the stereogram evokes significant perceived slant. In addition, this figure shows that there is a clear difference between the estimation of slant due to horizontal scale and vertical scale (see also Gillam, Chambers and Lawergren, 1988). The magnitude of estimated slant due to divergence is almost similar to the difference between estimated slant from horizontal scale and vertical scale for the entire range of presentation durations, which is again in accordance with the 'unequal weighting' extension of Howard and Kaneko's theory.

Figure 6c shows that in the presence of a visual reference, vertical shear between the half-images of the stereogram does not evoke perceived slant. This is in accordance with Kaneko and Howard's (1997) view that vertical shear disparities are derived globally. Figure 6c also shows that the magnitude of estimated slant due to horizontal shear is larger than estimated slant from rotation for the entire range of presentation durations. This finding is not expected from Kaneko and Howard's (1997) work since equal magnitudes of rotation and horizontal shear contain an identical magnitude of horizontal shear. This finding also seems to differ from our own previous reported results. With regard to slant perception in the presence of a frame of reference we reported earlier that equal magnitudes of rotation and horizontal shear led to identical slant settings (van Ee & Erkelens, 1995). The reason for the non-replication of the results could be that in the present study subjects had to *estimate* angles of perceived slant, whereas in the study by van Ee and Erkelens (1995) they had to *match* slant (with a reference plane adjacent to the test plane). Figure 6d shows that in the presence of a visual reference, vertical scale between the half-images of the stereogram evokes little perceived slant. This result is in accordance with Kaneko and Howard's (1996) view that vertical scale disparities are derived regionally and also with the results of Stenton, Frisby and Mayhew (1984). Figure 6d also shows that the magnitude of estimated slant due to horizontal scale is larger than slant estimated from divergence for the entire range of presentation durations. For reasons similar to those given above, this finding is not expected directly from the work of Kaneko and Howard (1996) or from the work of van Ee and Erkelens (1995).

---

<sup>11</sup> It is difficult to say precisely what the subjects learn. One reason for the learning effect could be that in our set-up experienced subjects are able to use the framework of the anaglyph glasses as a visual reference. To test for the latter possibility subjects RE and JZ did a similar slant estimation experiment in the haploscope of Prof. Martin Banks' lab at the School of Optometry in Berkeley (which does not make use of anaglyph glasses). However, although estimated slants from both rotation and divergence were smaller, the pattern of results was similar to that found in Utrecht.

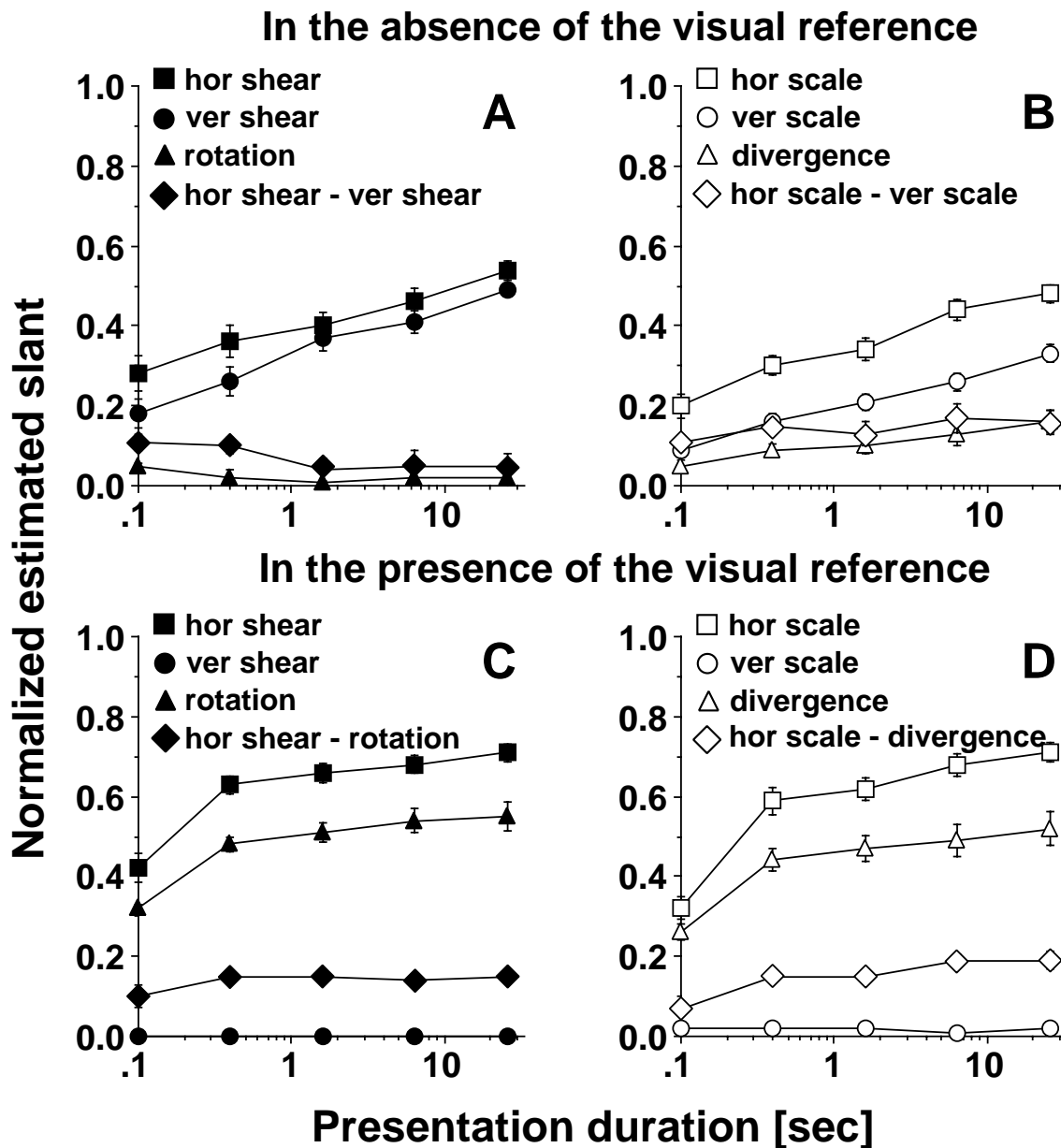


Figure 6: Averaged normalized estimated slant of the four inexperienced subjects EC, BJ, KY and IH versus the presentation duration. The error bars represent the standard error. Hor shear, ver shear, hor scale and ver scale denote horizontal shear, vertical shear, horizontal scale and vertical scale, respectively.

In figures 3 and 4 the differences in results between subjects are clearly larger in the absence than in the presence of a visual reference. In addition, underestimation of slant is more pronounced in the absence than in the presence of a visual reference. These observations are in accordance with the idea that stereopsis is relatively insensitive to whole-field slant perception (Gillam, Flagg & Finlay, 1984). A possible reason for the relative insensitivity of slant perception without a visual reference is that the stereoscopic visual system is utilized primarily for relative slant judgments (van Ee and Erkelens, 1996b). It could be that perspective cues are weighted more heavily relative to disparity cues without a reference than with a visual reference. The presence of conflicting perspective cues can also explain why, generally, subjects underestimate perceived slant. Nonetheless, in our stimulus we attempted to have a minimum of perspective cues, which probably explains why we did not find a marked anisotropy between perceived slant evoked by horizontal shear and scale (Ryan & Gillam, 1994).

Cagenello and Rogers (1990) reported on the influence of cyclovergence on estimates of slant evoked by shear disparities. They found that the magnitude of estimated slant evoked by vertical shear increased over time. They attributed this increase to the equal and opposite torsional movements which

the eyes might have undergone so that the vertical shear between the images was converted into horizontal shear. Howard and Kaneko (1994) argued that the perceived slant about the horizontal axis evoked by vertical shear is not due to the torsional state of the eyes since 'a strong asymmetry evident in the cyclovergence of four subjects was not reflected in their psychophysical data'. Kaneko and Howard (1997) noted that adaptation to a torsional state prior to the presentation of shear stimuli hardly affected perception of slant. In this study we show that there is another reason why the explanation of Cagenello and Rogers (1990) is unlikely to be correct. In our study we found that in a number of subjects the estimated magnitude of perceived slant evoked by horizontal shear increased over time in almost the same way as perceived slant evoked by vertical shear.

In the literature on disparity theories, when authors refer to disparity they often mean *retinal* disparities. However, in the literature on disparity experiments when authors refer to disparity they often mean *screen* disparities. Although there is a one-to-one mapping between screen coordinates and retinotopic coordinates, descriptions in the two domains are only equal for the first order disparity terms limited to an infinitesimally small region of the retina (Koenderink & van Doorn, 1976). Howard and Kaneko (1994) and Kaneko and Howard (1997) suggested that 'specialized neural circuits will be discovered for each of the differencing operations'. Although in essence we do not disagree with this remark, we stress that the strategy for searching for neural circuits is not trivial. It is important to note that the theory of Howard and Kaneko refers to *size* and *shear* disparities between the half-images of a stereogram projected on a frontal screen, but not to *size* and *shear* disparities in the retinal domain. In this paper [and in our earlier work (van Ee and Erkelens, 1995)] we also referred to transformations on a frontal screen. In order to stress the difference between screen transformations and retinal transformations, one could calculate the retinal disparities caused by a pure vertical shear between the frontally projected half-images used by Howard and Kaneko (or used in this paper). In addition to retinal vertical shear disparities one finds (van Ee and Erkelens, 1996b) retinal horizontal scale disparity (since the projection on the screen is not along the theoretical horopter) and a certain retinal vertical scale disparity (since the points in the left and the right part of the median plane are at different distances from each eye). In view of the fact that Howard and Kaneko's theory takes into account regional and even global characteristics of the disparity field, their theory, if expressed in retinotopic coordinates, would not be equal to, and would not be as elegant as their theory which is expressed in screen coordinates. There is no firm mathematical basis for their theory. Thus, it is not yet clear how an underlying mechanism (neural network) of the visual system could process local horizontal scale and shear disparities, regional vertical scale disparities and global vertical shear disparities (but see Erkelens & van Ee, 1997). If we want to develop a neural network which incorporates the scheme of Howard and Kaneko it will have to include a stage which converts retinotopic coordinates into screen coordinates. We recall that vertical scale and vertical shear as presented on a frontal screen can never represent a real object under the given oculomotor conditions. This is easy to see because the light rays of corresponding features of the half-images never intersect in space: vertically shifted corresponding features on a screen do not obey the epipolar constraint. This makes it the more intriguing why horizontal scale and vertical scale evoke predictable magnitudes of estimated slant. It could be that vertical disparity is a trivial influence of the scaling of horizontal disparities, autonomously and to a certain extent independent of conflicting oculomotor information (Gillam & Lawergren, 1983).

Finally, we conclude that the experimental results for all our subjects and for all tested presentation durations are in agreement with the above-formulated proposed extension to Howard and Kaneko's theory. The extension makes Howard and Kaneko's theory more explicit in that their proposals are translated into linear mathematical expressions. These expressions contain weighting factors which allow both for slant evoked by rotation or divergence and for the subject-dependent underestimation of slant. A possible future line of research in stereoscopic slant perception is to determine whether perceived slant about oblique axes without a visual reference, in other words, slant evoked by a combination of whole-field scale and shear disparities, is successfully predicted by the proposed extension.

**Acknowledgments**---The work was supported by a grant from the Foundation for Life Sciences (SLW) of the Netherlands Organization for Scientific Research (NWO). We thank Dr. Ian Howard for many helpful comments. We thank Drs. Marty Banks and Ben Backus for using their haploscope and insightful comments. We also thank Drs. Cliff Schor and Tom Freeman for helpful comments, and Pieter Schiphorst for developing part of the computer software.

## 6 References

- Backus, B. T. & Banks, M. S. (1997). Slant-cue conflicts explain why the induced effect does not scale with distance. *Investigative Ophthalmology and Visual Science*, 38, S904.
- Cagenello, R. & Rogers, B. J. (1990). Orientation disparity, cyclotorsion, and the perception of surface slant. *Investigative Ophthalmology and Visual Science*, 31, S97.
- 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.
- Erkelens, C. J. & van Ee, R. (1997). Are stereopsis and ocular vergence based on head related disparity. *Investigative Ophthalmology and Visual Science*, 38, S955.
- Gillam, B., Chambers, D. & Lawergren, B. (1988). The role of vertical disparity in the scaling of stereoscopic depth perception: An empirical and theoretical study. *Perception & Psychophysics*, 44, 473-483.
- 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.
- Gillam, B., & Lawergren, B. (1983). The induced effect, vertical disparity, and stereoscopic theory. *Perception & Psychophysics*, 34, 121-130.
- Gillam, B. & Rogers, B. J. (1991). Orientation disparity, deformation and stereoscopic slant perception. *Perception*, 20, 441-448.
- Howard, I. P. & Kaneko, H. (1994). Relative shear disparities and the perception of surface inclination. *Vision Research*, 34, 2505-2517.
- Howard, I. P. & Zacher, J. E. (1991). Human cyclovergence as a function of stimulus frequency and amplitude. *Experimental Brain Research*, 85, 445-450.
- Kaneko, H. & Howard, I. P. (1996). Relative size disparities and the perception of surface slant. *Vision Research*, 13, 1919-1930.
- Kaneko, H. & Howard, I. P. (1997). Spatial properties of shear disparity processing. *Vision Research*, 315-323.
- Koenderink, J. J. & Van Doorn, A. J. (1976). Geometry of binocular vision and a model for stereopsis. *Biological Cybernetics*, 21, 29-35.
- 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.
- 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. & Koenderink, J. J. (1986). Monocular aniseikonia: a motion parallax analogue of the disparity-induced effect. *Nature*, 322, 62-63.
- Ryan, C. & Gillam, B. (1994). Cue conflict and stereoscopic surface slant about horizontal and vertical axes. *Perception*, 23, 645-658.
- Stenton, S. P., Frisby, J. P. & Mayhew, J. E. W. (1984). Vertical disparity pooling and the induced effect. *Nature*, 309, 622-623.
- Stevens, K. A. (1983). Slant-Tilt: the visual encoding of surface orientation. *Biological Cybernetics*, 46, 183-195.
- Swash, S., Rogers, B. J., Bradshaw, M. F. & Cagenello, R. B. (1995). The role of cyclovergence in the perceived slant of stereoscopic images related by vertical shear, rotation and deformation. *Investigative Ophthalmology and Visual Science*, 36, S368.
- van Ee, R., Backus, B. T. & Erkelens, C. J. (1996). Perceptual learning in stereoscopic slant estimation. Internal Report of the Helmholtz Institute, UU-PA-pmi-088.
- van Ee, R. & Erkelens, C. J. (1993). Slant perception in relation to the geometry of binocular vision: Effects of local and global disparity fields. 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*, 24, 299-314.
- van Ee, R. & Erkelens, C. J. (1996a). Temporal aspects of binocular slant perception. *Vision Research*, 36, 43-51.
- van Ee, R. & Erkelens, C. J. (1996b). The stability of stereoscopic depth perception with moving head and eyes. *Vision Research*, 36, 3827-3842, 1996.
- van Ee, R. & Erkelens, C. J. (1996c). Binocular slant perception from scale, shear, rotation and divergence. *Investigative Ophthalmology and Visual Science*, 37, S289.