BEHAVIORAL AND BRAIN SCIENCES (1994) 17, 293–355
*Printed in the United States of America*

# Motion perception during self-motion: The direct versus inferential controversy revisited

**Alexander H. Wertheim**
*TNO Institute for Human Factors, 3769 ZG Soesterberg, The Netherlands*
**Electronic mail:** *wertheim@izf.tno.nl*

**Abstract:** According to the traditional inferential theory of perception, percepts of object motion or stationarity stem from an evaluation of afferent retinal signals (which encode image motion) with the help of extraretinal signals (which encode eye movements). According to direct perception theory, on the other hand, the percepts derive from retinally conveyed information only. Neither view is compatible with a perceptual phenomenon that occurs during visually induced sensations of ego motion (vection). A modified version of inferential theory yields a model in which the concept of extraretinal signals is replaced by that of reference signals, which do not encode how the eyes move in their orbits but how they move in space. Hence reference signals are produced not only during eye movements but also during ego motion (i.e., in response to vestibular stimulation and to retinal image flow, which may induce vection). The present theory describes the interface between self-motion and object-motion percepts. An experimental paradigm that allows quantitative measurement of the magnitude and gain of reference signals and the size of the just noticeable difference (JND) between retinal and reference signals reveals that the distinction between direct and inferential theories largely depends on: (1) a mistaken belief that perceptual veridicality is evidence that extraretinal information is not involved, and (2) a failure to distinguish between (the perception of) absolute object motion in space and relative motion of objects with respect to each other. The model corrects these errors, and provides a new, unified framework for interpreting many phenomena in the field of motion perception.

**Keywords:** direct perception; efference copy; inference; motion perception; self-motion; velocity perception; visual-vestibular interactions

## 1. Inferential versus direct perception

How do we maintain the visual percept of a stable world while images of our environment move across the retinae during eye movements? Answers to this question can be classified in two main theoretical approaches. According to the traditional view, here called inferential theory, we perceive the motion or stationarity of an object, or of the visual world itself, on the basis of the outcome of a comparison between two neural signals (see e.g., Helmholtz 1910; Jeannerod et al. 1979; MacKay 1972; Mittelstaedt 1990; Sperry 1950; Von Holst & Mittelstaedt 1950). One signal, here to be called the *retinal signal,* consists of retinal afferents encoding the characteristics of the movement of the objects' image across the retina. The other signal, encoding concurrent eye movement characteristics, is usually termed the *extraretinal signal,* because it does not derive from visual afferents (Matin et al. 1969; Mack 1986; see also Matin 1982; 1986). The comparison mechanism treats the two signals as vectors (see, e.g., Mateeff et al. 1991; Wallach et al. 1985) and applies a simple rule: *when they differ, object motion is perceived; when they are equal, object stationarity is perceived.* Wertheim (1981) has shown that when a smooth pursuit eye movement is made across a visual stimulus pattern, the magnitude of the retinal signal corresponds to the velocity of the retinal image flow of the pattern. Similarly, the magnitude of the extraretinal signal corresponds to the velocity of the concurrent eye movement as "estimated" within the perceptual apparatus (see sect. 5.1). In the present target article, eye movements are mainly of the smooth pursuit type. Hence, the terms "magnitude" or "size" of retinal and extraretinal signals will refer to these velocity vectors. We see a stable world during eye movements because retinal and extraretinal signals are equal: the velocity of the image of the world across the retinae equals the velocity of the eyes.

The alternative theoretical view, here called the theory of direct perception, which originated from Gibson (1966; 1979), has no need of the concept of an extraretinal signal (Gibson 1968; 1973), as it assumes that the perception of motion derives exclusively from afferent retinal information. Its point of departure is that in everyday circumstances perception is veridical (it should be: the organism's chances of survival depend on it – for this reason the approach is also called the ecological theory of perception). Hence the perceptual mechanism functions as an unbiased sampler of external information from the real world (see Lombardo 1987). According to this theory, the visual world manifests itself as the particular pattern of light, called the *optic array,* that hits an observer's eye. The informational content of the scene is given in ("specified" by) particular invariant structural features of this light pattern. To perceive is to "pick-up" such invariants.

 0140-525X/94 $5.00+.00 

Thus, movement of an object may be specified by an invariant described as the concurrent appearance and disappearance of part of the array (specifying the background) along the two opposite borderlines of another part of the array (specifying the object). When the eyes move across the visual world as during (combined) eye, head, or ego movements, a coherent streaming motion of the optic array relative to the retinae usually occurs. The resulting retinal flow pattern has in recent years become the focus of research in the literature of direct perception theory. The basic assumption is that the brain is able to "pick up" from retinal flow those characteristics that are caused by invariants of the optic array such as the one mentioned above (they may be called "optic flow invariants"). A retinal flow pattern, however, may also contain characteristics that stem from movements of the eyes in space (caused by eye, head, or ego motion), but these invariants only specify how the eyes move (or are moved in space). When these invariants are "picked up," we only perceive (i.e., become aware of) the particular self- or ego motions that gave rise to these invariants.[1] This is called "visual kinaesthesis." For example, an invariant that specifies eye movements in the head could be motion of the dark middle area of the array – specifying the nose – relative to the outer boundaries of the optic flow field. Other invariants specify head or ego movements.[2]

Because the optic array stems from a stable world, retinal flow never holds optic flow invariants that could specify motion of the world. Consequently, the visual world cannot be perceived as moving. Recently, the question has been raised whether the visual system always needs to distinguish between optic flow invariants and self-motion invariants (Cutting et al. 1992b). Although strictly speaking this reflects a deviation from the original point of departure of direct perception theory, it does not affect the fundamental principle to be discussed in this paper: that the perception of object motion or stationarity stems only from retinal afferent information and not from comparing retinal and extraretinal information.

In neurophysiological research, awareness of ego motion is usually associated with the output activity of cells in particular areas of the brain, notably the vestibular nuclei and the vestibular cortex. These cells are driven by afferents from the equilibrium system and the somatosensory kinaesthetic system (here these will be called, collectively, *vestibular afferents*). Many of these cells are also driven by visual (image flow) afferents. One important pathway through which these visual afferents are conducted is known as the accessory optic pathway (see, e.g., Büttner & Buettner 1978; Büttner & Henn 1981; Cohen & Henn 1988; Dichgans & Brandt 1978; Dichgans et al. 1973; Henn et al. 1974; 1980). These visual afferents are complementary to vestibular afferents. Their function is to generate or sustain sensations of ego motion when the equilibrium system remains silent, that is, in the absence of an accelerating force acting on the equilibrium system (e.g., when traveling at constant velocity in a train). In the literature concerned with research in this area of so-called visual-vestibular interactions, a visually induced sensation of ego motion is termed *vection*, and the particular features of retinal flow that generate vection are not called "invariants that specify ego motion," but "optokinetic." The stimuli that generate them will here be termed "optokinetic stimuli," and the term "optokinetic pathway" will be used to denote in general the combined neural channels that convey the optokinetic afferents that generate vection and interact with vestibular afferents. To be optokinetic, a visual pattern must be large, have relatively low spatial frequency characteristics, move (not too fast) across the retinae, and remain visible for more than a very brief interval (see, e.g., Berthoz et al. 1975; Berthoz & Droulez 1982; Brandt et al. 1973; De Graaf et al. 1990; Dichgans & Brandt 1978).

It is the purpose of this target article to show that – within the domain of object-motion perception – an adapted version of inferential theory, in combination with knowledge from the research area of visual-vestibular interactions and ego motion, resolves the differences of opinion between inferential and direct theories of perception.

## 2. Problems for both theories

If vection is generated in the laboratory, some perceptual phenomena may occur that are incompatible with both direct and inferential theory. As an example, consider vection created with an "optokinetic drum," a large drum with vertical black and white stripes painted on its inside wall that can be rotated around an observer seated inside on a stationary chair. For the present purpose let us assume that the drum rotates with an angular velocity of 60 deg/sec around a stationary observer whose body, head, and eyes are fixed in space (using a small stationary fixation point attached to the stationary chair). Let us further assume that the lights inside the drum are extinguished (i.e., the observer sits in the dark and does not know that the drum rotates). If we now suddenly switch on the lights inside the drum, the observer will initially perceive the drum correctly as rotating and will experience no ego rotation. However, within a few seconds an illusory sensation of ego rotation in the direction opposite to that of the drum (called *circular vection*) gradually develops. During this period, ego velocity is experienced as increasing and the rotation of the drum appears to slow down. Finally, the drum is perceived as completely stationary in space and ego velocity does not seem to increase any further. Circular vection is then said to be saturated. The whole process – from the moment the lights inside the drum are switched on to the saturation of vection – may last between 4 and 6 seconds, depending on the velocity of the drum. At very low drum velocities, saturated vection may even be immediate, but in the present example, where drum velocity is considerably higher, it may take as long as 6 seconds or more before vection is completely saturated (for more details about the dynamics of circular vection, see e.g., Dichgans & Brandt 1978; Mergner & Becker 1990; Wong & Frost 1978).

The question that raises theoretical problems for both direct and inferential theory is: Why, during saturated circular vection, is the drum perceived as stationary in space? Direct perception theory has a simple answer: a coherent retinal flow of the entire environment is an invariant that normally specifies ego motion. When picked up, this yields a percept (an awareness) of ego

motion, not of drum motion. This reasoning poses two problems, however. First, how could the drum initially have been perceived as moving? That suggests the presence of an invariant that specifies environmental motion. Second, this anomalous invariant seems to dissipate in time (as drum rotation appears to slow down gradually) and disappears completely upon saturation, even though the optic array and the retinal flow characteristics remain physically identical.[3]

Inferential theory can explain why the drum is initially perceived as moving: its moving retinal image generates a substantial retinal signal, but the stationary eyes (focused on the fixation point) generate a zero extraretinal signal. The two signals therefore differ and the drum is seen to move. Hence for inferential theory the problem is that the drum appears to be stationary once vection is saturated.

## 3. An alternative model

The problems can be solved within the framework of inferential theory by reconsidering the concept of an extraretinal signal. This signal is usually defined as encoding ocular velocity and serves to determine to what extent retinal image motion is an eye movement artifact. The remaining image motion then reflects real object motion in external space. However, this reasoning only holds if the signal encodes eye velocity *relative to external space*, not relative to the head. The logic of this point has been recognized by many authors (see, e.g., Swanston et al. 1987; Swanston & Wade 1988; Wallach 1987) but its consequences for the nature of extraretinal signals have not been fully recognized.

Formally speaking, eye velocity in space ($V_{eye.s}$) corresponds to the vectorial addition of eye velocity in the head ($V_{eyes.h}$) and head velocity in space ($V_{head.s}$). Thus, it is here proposed that extraretinal signals actually consist of the vector sum of a $V_{eyes.h}$ and a $V_{head.s}$ velocity vector. The $V_{eyes.h}$ vector may derive from what is known as the "efference copy" – a neural corollary to the efferent oculomotor commands (Von Holst & Mittelstaedt 1950)[4] – and the $V_{head.s}$ vector most likely derives from vestibular afferents that result from head movements.

The implication of this reasoning is that during ego motion extraretinal signals must also be generated: although the eyes may not move in their orbits during ego motion, they do move in space and thus create artifactual retinal image motion.[5] How are these extraretinal signals generated? First, they most probably derive from the already mentioned vestibular afferents that encode $V_{head.s}$ during ego motion. There must be another source, however. The point is that in cases where the awareness of ego motion is sustained visually (vection), there are no such vestibular afferents: their function is taken over by the visual afferents that are induced by optokinetic image flow and pass through such channels as the accessory optic pathway. These pathways are referred to here by the general term "optokinetic pathway." Thus, it is proposed that such particular visual afferents may also generate (part of) an extraretinal signal. This obviously renders the term "extraretinal signal" incorrect. Therefore, from here on, the term *reference signal* will be used, which emphasizes only the evaluative function of the signal with respect to retinal image motion.

In summary, then, the present model holds that reference signals are compound signals, which may include (any combination of) an efference copy, a vestibular, and a visual component. Figure 1 illustrates how such reference signals may be generated.

The gating mechanism in the optokinetic pathway determines what aspects of visual afferents generate vection and thus generate or affect reference signals. The features that make a visual stimulus (its retinal flow) optokinetic have already been mentioned. They suggest that the gating mechanism acts as a low band-pass spatiotemporal filter.

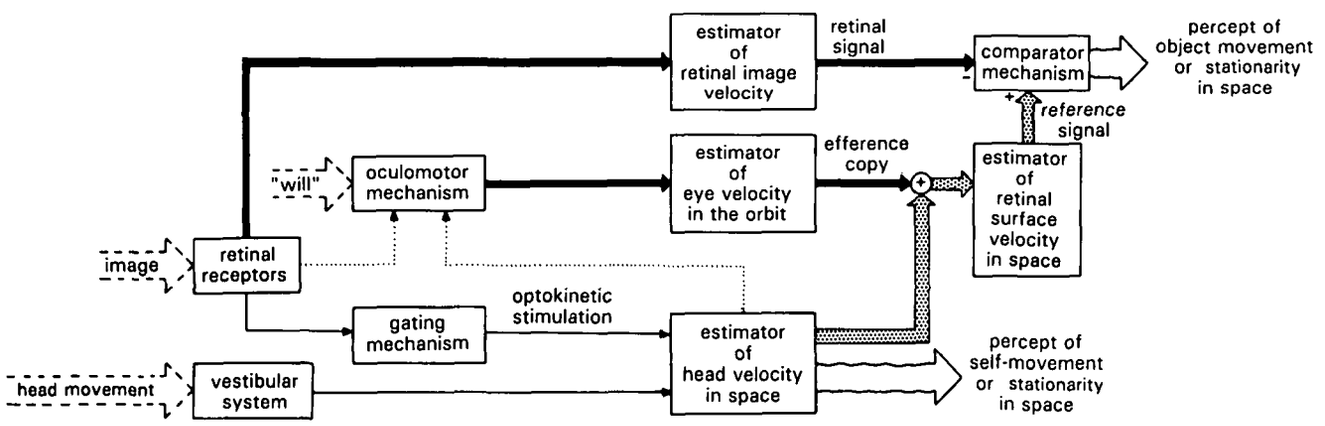


Figure 1. Functional model describing the generation of reference signals as they interrelate the percepts of ego motion and object motion in space. Thick lines reflect the traditional inferential literature on object-motion perception. Thin lines reflect the literature on visual-vestibular interactions and ego-motion perception. Gray arrows illustrate the contribution of the present model, connecting these two bodies of literature. Note that the reference signal (not the efference copy) feeds into the comparator mechanism and represents the estimate of eye velocity in space. The term "estimator" has no cognitive connotation but implies that retinal image, eye, and head velocities are not necessarily encoded correctly but with a particular gain. The dotted lines between the retina and the oculomotor mechanism and between the estimator of head velocity in space and the oculomotor mechanism represent the pathways that function to generate reflexive eye movements known as optokinetic and vestibular nystagmus (these are discussed in sect. 5.4).

A warning should be made here: the addition of a visual component to the reference signal is not meant to imply strict linear additivity. In fact, it is quite likely that the interaction between retinal and vestibular afferent information at the level of the estimator of head velocity in space is nonlinear (see, e.g., Barthélémy et al. 1988; Borah et al. 1988; Fletcher et al. 1990; Probst et al. 1985; Xerri et al. 1988).

The theoretical significance of the visual component in the reference signal – which can be conceptualized in cybernetic terms as a kind of feedforward signal – is that it implies a self-referential circularity or "strange loop" (Hofstadter 1980) in the perceptual system: retinal image motion may create (part of) a reference signal to determine its own perceptual interpretation. This circularity solves the problems associated with the development and saturation of circular vection: when the optokinetic drum starts rotating, the moving image of its stripes immediately generates a retinal signal (the eyes do not move in the head, as they remain focused on the stationary fixation point). But in the present example (in which the drum rotates at 60 deg/sec) vection develops only gradually, due to the low temporal band-pass characteristics of the gating mechanism in the optokinetic pathway. Hence, a (visually induced) reference signal is not immediately present, and the drum is initially perceived (correctly) as moving. When vection begins to build up, however, so does the reference signal. The difference between the (unchanged) retinal signal and this growing reference signal thus decreases gradually. If perceived object velocity is determined by this difference –as shown in section 5.2 – drum velocity will be seen as slowing down until saturation is reached, that is, until the reference signal has become approximately equal to the retinal signal. The drum is then perceived as stationary in space.

The relevance of this model for the discussion between direct and inferential theories of motion perception is that it provides a view that to a large extent agrees with both these theories, that is, it creates a compatibility between the basic presumptions of both inferential and direct theory. On the one hand, it agrees with the main inferential premise that information about how the eyes move (in space) is always necessary to perceive object motion or stationarity; on the other hand, it also agrees with three main assumptions of direct perception theory. First, the percept of object motion or stationarity may indeed stem exclusively from visual afferents (i.e., when reference signals only consist of a visual component). Second, retinal flow patterns may indeed specify ego motion and do not specify motion of the visual environment. Third, the gating mechanism in the optokinetic pathway (see Fig. 1) can be viewed as the mechanism responsible for "picking up" invariants from retinal image flow. Hence, in the light of the present model, the fundamental postulates of direct and inferential theory are no longer contradictory.

In the remainder of this article it will be shown that this also holds for the empirical database that has given rise to the controversies between direct and inferential theory, as well as to theoretical attempts to find a compromise between the two approaches (i.e., theories that propose that direct and inferential perception are not mutually exclusive but reflect two distinct modes of perception; such theories will be called "dual mode theories," see sect. 5). To make this clear, I will review empirical tests of predictions that follow from the present model but do not follow from dual mode theory or from either of the original two rival approaches themselves. First, however, an experimental paradigm must be outlined to serve as the frame of reference in terms of which the data obtain their significance.

## 4. Experimental paradigm

Imagine a subject looking at a screen in front of his eyes. On the screen a visual stimulus is projected. The stimulus can move in both horizontal directions with a fixed velocity, set by the experimenter. Assume also that the subject's head is fixed in space but that his eyes pursue a small fixation point sweeping horizontally (with another fixed velocity) across the moving stimulus. If we synchronize the beginning and termination of the motions of the stimulus and the fixation point, we can study the perception of stimulus motion or stationarity during a (pursuit) eye movement – made across the stimulus – of any given velocity. We will then use the following conventions: first, the terms "retinal image" or "retinal signal" will always be used to refer to the image of the stimulus, not the image of the fixation point. Second, retinal image velocity will be defined as the velocity of the eyes in space minus the velocity of the stimulus in space. This means that the directional sign given to the retinal image velocity vector (i.e., to the retinal signal, $V_{ret}$) will be such that in the case of a stationary stimulus it is the same as the sign given to the direction in which the eyes move in space ($V_{eyes.s}$).

Thus when, in the present example, the stimulus is stationary, the velocity of its retinal image equals $V_{eyes.s}$. If the stimulus is indeed perceived as stationary, retinal and reference signals must be equal too. Now imagine that we move the stimulus slightly in the *same* direction as the eyes. This reduces retinal image velocity and thus decreases the size of the retinal signal, which then becomes slightly smaller than the reference signal. If we further increase stimulus velocity, the difference between retinal and reference signals further increases until it become detectable within the perceptual apparatus. At that point the threshold is reached for perceiving stimulus motion during a pursuit eye movement. The retinal signal is then exactly one just noticeable difference (JND) smaller than the reference signal (see MacKay 1973; Wallach & Kravitz 1965; Wertheim 1981). This may be expressed as:

$$V_{retW} = V_{ref} - JND \quad (1)$$

where $V_{retW}$ is retinal signal size at the threshold for stimulus motion *with* the eyes (with-threshold), and $V_{ref}$ is the magnitude of the reference signal induced by the eye movement. Conversely, if the stimulus moves in the direction *opposite* to the eyes, retinal image velocity increases. The threshold for perceiving stimulus motion in that direction (against-threshold) is then reached when

$$V_{retA} = V_{ref} + JND \quad (2)$$

where $V_{retA}$ is retinal signal size at the against-threshold. It thus follows that

$$JND = \frac{|V_{retA} - V_{retW}|}{2} \quad (3)$$

Since retinal image velocity can be calculated as $V_{eyes.s} - V_{stim.s}$ (where $V_{stim.s}$ is stimulus velocity in space), this may also be written as:

$$JND = \frac{|V_{stimA.s} - V_{stimW.s}|}{2} \quad (4)$$

Hence, half the difference between the stimulus velocities at the two opposite thresholds for perceiving object motion can be used as an operational measure of the magnitude of one JND between retinal and reference signals.[6]

At the exact midpoint between the two opposite thresholds – which in this article will be called the point of subjective stationarity (PSS) – retinal image velocity ($V_{ret.PSS}$) corresponds to $V_{ref}$ because

$$V_{ret.PSS} = \frac{V_{retW} + V_{retA}}{2} = \frac{(V_{ref} - JND) + (V_{ref} + JND)}{2} = V_{ref} \quad (5)$$

Thus at the PSS retinal image velocity is not only proportional to the retinal signal but also the concurrent reference signal. We may therefore take retinal image velocity at the PSS as an operational measure of reference signal size.

The gain of a reference signal ($G_{ref}$) is the extent to which it registers the actual velocity of the eyes in space ($V_{eyes.s}$). It can be expressed as:

$$G_{ref} = \frac{V_{ref}}{V_{eyes.s}} \quad (6)$$

Since $V_{ref}$ was operationalized as $V_{ret.PSS}$, $G_{ref}$ may also be expressed as:

$$G_{ref} = \frac{V_{ret.PSS}}{V_{eyesPSS.s}} \quad (7)$$

$V_{eyesPSS.s}$ being the velocity of the eyes in space at the PSS. Since retinal image velocity equals $V_{eyes.s}$ minus stimulus velocity in space ($V_{stim.s}$), this may also be expressed as:

$$G_{ref} = 1 - \frac{V_{stimPSS.s}}{V_{eyesPSS.s}} \quad (8)$$

where $V_{stimPSS.s}$ is stimulus velocity in space at the PSS. Note that the PSS is the midpoint between two opposite thresholds. If they are equal, $V_{stimPSS.s}$ is zero. $G_{ref}$ then equals 1, which means that eye velocity in space is correctly registered in the reference signal.

What would unequal thresholds mean? Assume that the with-threshold is higher than the against-threshold. $V_{stimPSS.s}$ then differs from zero and is in the *same* direction (has the same sign) as $V_{eyesPSS.s}$. According to Equation 8, $G_{ref}$ is then smaller than 1, which means that the reference signal is too small, that is, that eye velocity in space is underregistered in the reference signal (to the extent of $1 - G_{ref}$). Conversely, if the against-threshold is higher than the with-threshold, the stimulus moves at the PSS in the direction *opposite* to $V_{eyesPSS.s}$. $G_{ref}$ is then larger than 1, and $G_{ref} - 1$ indicates the extent to which eye velocity in space is overrepresented in the reference signal. Hence asymmetric thresholds indicate an under- or overregistration of eye velocity in space in the reference signal depending on which threshold is higher, that is, on whether the PSS has shifted in the direction with or against the eyes.

## 5. Empirical tests of the model and their relevance for direct and inferential theory

**5.1. Thresholds for motion perception.** As mentioned in section 3, there have been some attempts to bridge the gap between the direct and inferential approaches in the form of a dual mode theory. This is basically the assumption that there exist two modes of visual perception: a direct mode, in which extraretinal signals play no role and which yields veridical percepts, and an inferential mode, which makes use of extraretinal signals and may yield illusions. For example, it is claimed that when a visual pattern is very large and covers most or all of the visual field, a particular mode of perception called *visual capture* becomes dominant. This mode needs no extraretinal signals and creates veridical percepts (see, e.g., Stark & Bridgeman 1983). Hence it can be viewed as a direct perceptual mode (e.g., Mack 1978). (It is also possible to view visual capture as a cognitive influence on perception, assuming that such patterns evoke a cognition of environmental stationarity because we *know* that our environment is normally stationary.)

Dual mode theory (Mack 1978; 1986; see also Matin 1986) has developed from concepts originally formulated by Wallach (see, e.g., Wallach 1959) to explain the phenomenon of center surround induced motion (a stationary stimulus is seen to move when its surrounding background moves, irrespective of whether the eyes fixate the stimulus or track the surround; e.g., Shulman 1979). According to Wallach, there are two kinds of cues that may generate a percept of motion: "object-relative" and "subject-relative" cues (see also Shaffer & Wallach 1966). The "object-relative" cues stem from motion of objects relative to each other (i.e., from motion of object images relative to each other on the retina; see Matin 1986). These "object-relative" cues presumably overrule or suppress what Wallach called "subject-relative" cues, which stem from object motion relative to the observer. Center-surround induced motion is then explained as follows: the percept of surround motion, which is "subject-relative," is overruled by the percept of motion that stems from the "object-relative" cue of surround motion relative to the center stimulus. The impression of motion, however, is attributed to the smaller center stimulus, because – according to a Gestalt-like principle called the "stationarity tendency of large stimuli" (Duncker 1929) – a surround tends to act as a perceptual frame of reference (see, e.g., Mack & Herman 1978; Wallach 1959).

According to dual mode theory, "object-relative" and "subject-relative" cues somehow force the visual system to operate in a direct or in an inferential perceptual mode, respectively. The dominant direct mode is always operative in normal circumstances, because objects usually move relative to a full-field visually structured background – which implies the presence of "object-relative" motion cues – and the Gestalt principle mentioned above

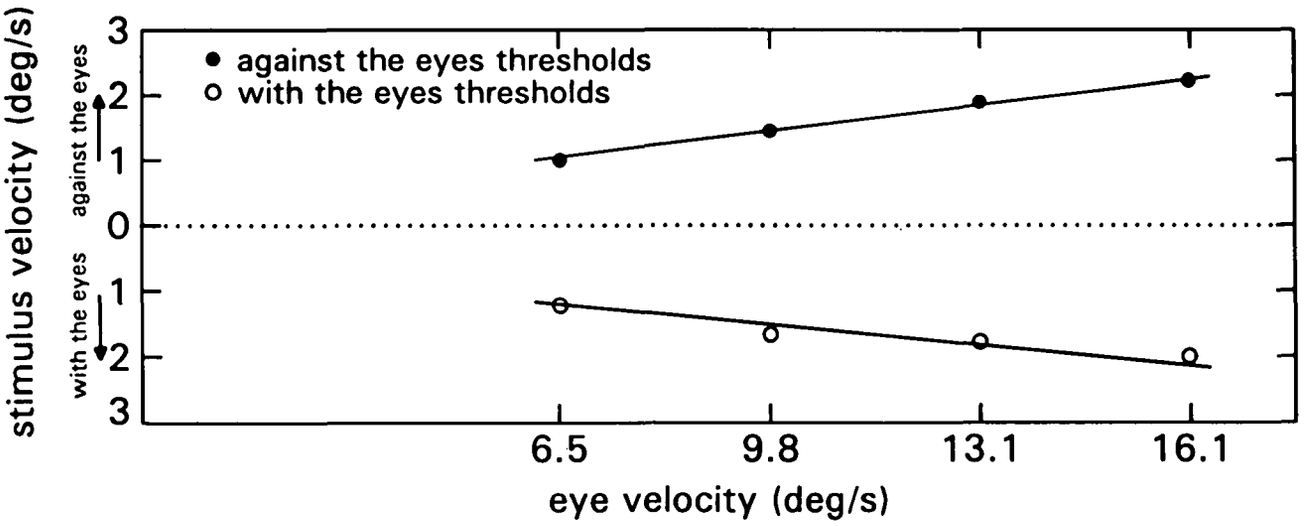


Figure 2. Stimulus velocity at the two opposite thresholds (with and against the eyes) for perceiving motion in space of a large stimulus pattern, during a pursuit eye movement to a fixation point sweeping across the pattern, as a function of eye velocity. Head is stationary in space.

always attributes the impression of motion to the smaller objects. The inferential mode, on the other hand, is seen as a kind of backup system, which uses extraretinal signals. It becomes operative if no "object-relative" cues are present (e.g., when objects move in a totally darkened environment). This mode produces illusions because of an underregistration of eye velocity in the efference copy.

Dual mode theory may be criticized on the basis of the argument that illusions of motion of the visual world often occur in situations where they should be prevented by capture (e.g., when dizzy, or when gently pressing a finger against the eyeball). But in the present section we will take a different approach, reviewing a number of experiments whose results show that the logic of dual mode theories is flawed, because the empirical criterion for distinguishing between the two modes is questionable.

The experiments concern predictions about thresholds for motion during eye movements. According to the new model, the difference between the thresholds with and against the eyes corresponds to twice the JND between the retinal and the reference signal (Equation 4). As JNDs increase linearly with signal size – Weber's law – *the distance between the two thresholds should increase linearly with eye velocity (in space).* Wertheim (1981) measured these thresholds for a large stimulus pattern (head fixed in space) and shows this to be true (Fig. 2). The dependency of the thresholds on eye movement velocity (rather than amplitude) implied that during pursuit eye movements the magnitude of retinal and reference signals corresponds to the encoded *velocity* of eye and image movements.[7]

In Figure 3 the data from the same experiment are plotted in terms of a relation between retinal image velocity and eye velocity (in space). The dashed line in this graph divides the vertical distance between the two threshold lines in half. It thus represents retinal image velocity at the midpoints between the two opposite thresholds, or $V_{ret.PSS}$, that is, it gives the magnitude of $V_{ref}$ at any eye velocity (in space), and according to Equation 7, its slope reflects $G_{ref}$.

In this particular experiment, $G_{ref}$ was approximately 1, that is, eye velocity in space was encoded more or less correctly in the reference signal. It should be noted that in this study the stimulus pattern was present on the screen throughout each pursuit eye movement that was made across it. Hence, during the eye movements there was always retinal flow. Therefore, the reference signal must – apart from its efference copy component – have contained a (relatively small) visual component. If the stimulus had been very small and had been visible only briefly during each pursuit eye movement, no such visual component would have been generated, because with

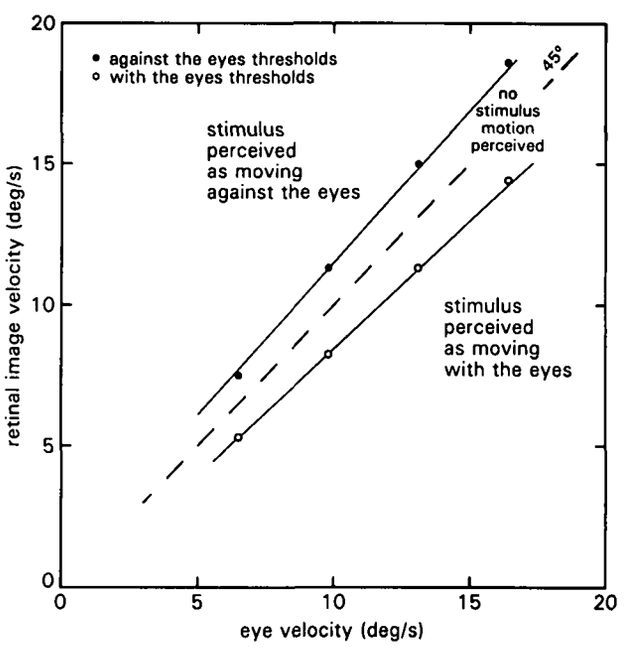


Figure 3. Retinal image velocity at the two opposite thresholds (with and against the eyes) for perceiving stimulus motion in space of a large stimulus pattern, during a pursuit eye movement to a fixation point sweeping across the pattern, as a function of eye velocity. Head is stationary in space.

such stimuli retinal afferents are too small and too short-lived to pass through the optokinetic pathway (given its low spatiotemporal band-pass gating characteristics). Consequently, it is predicted that with small and briefly visible stimuli the reference signal (its size and gain) should be less than with large stimuli that remain visible for a longer period.

Experiments with such small and briefly visible stimuli (performed in total darkness) have been reported by Mack and Herman (1978). These do indeed indicate the presence of undersized reference signals ($G_{ref} < 1$), because they yield high with- and low against-thresholds: at the PSS the stimuli always moved slightly in the same direction as the eyes. Since in these experiments reference signals could have consisted only of an efference copy, this is evidence that, during smooth pursuit eye movements, ocular velocity in the head is underregistered in the efference copy. In the Mack and Herman study the asymmetry between the with- and against-thresholds was quite strong. The against-threshold was often so low that it actually became "negative," that is, when stationary, the stimuli were still perceived as moving above threshold against the eyes (to reach the against-threshold they must be moved slightly with the eyes). This phenomenon is known as the Filehne illusion (De Graaf & Wertheim 1988; Filehne 1922; Mack & Herman 1973; Wertheim 1987). Its occurrence always implies a significantly undersized reference signal.[8]

The Wertheim (1981) study, however, does not necessarily prove the existence of reference signals that include a visual component. Since the stimulus was quite large (38 × 20 deg), the absence of a Filehne illusion could be explained as an instance where, according to dual mode theory, a direct mode of perception has occurred: visual capture may have happened or the "stationarity tendency" of large stimuli may have counteracted the Filehne illusion.

To test these hypotheses against the present one, the Wertheim (1981) study was replicated with a large but briefly visible stimulus pattern flashed on the screen for only 300 msec during the pursuit eye movement (Wertheim 1985; Wertheim & Bles 1984). Because briefly visible stimuli, whatever their size, cannot be optokinetic (do not pass the low temporal band-pass gating in the optokinetic pathway) they cannot generate a visual component in the reference signal (see Fig. 1). Hence the Filehne illusion should reappear. But according to a visual capture or stationarity-tendency hypothesis, no such illusion should occur with such a large stimulus. As shown in Figure 4, however, the illusion was observed.

Nevertheless, the support for the present model is still not definitive, because visual capture or a stationarity tendency might need more than 300 msec to build up. To test the model against this possibility, the experiment was repeated, but now with stimuli varying in optokinetic potential (Wertheim 1987). A very powerful optokinetic stimulus should induce such a large visual component that reference signals may become oversized ($G_{ref} > 1$). In terms of Equation 7, this means that to reach the PSS such a pattern should be moved *against* the eyes. If the effect is strong enough, an *inverted* Filehne illusion should be observed (the stimulus would, when stationary, seem to move *with* the eyes). No visual capture or stationarity-tendency hypothesis can be compatible with such a result. Various stimulus patterns were used. Each consisted of a large sinusoidal grating of a particular spatial frequency. Low spatial frequency patterns have a stronger optokinetic potential than high spatial frequency patterns (Berthoz & Droulez 1982; Bonnet 1982; De Graaf et al. 1990). Hence the former should create a larger visual component in the reference signal than the latter; and with very low spatial frequencies the reference signal might become oversized.

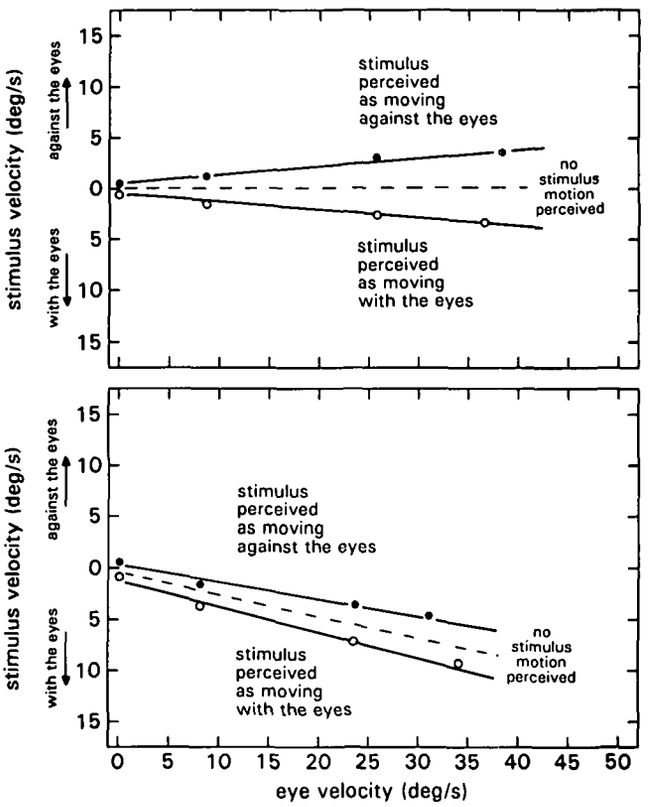


Figure 4. Stimulus velocity at the two opposite thresholds (with and against the eyes) for perceiving motion in space of a large stimulus pattern, during a pursuit eye movement to a fixation point sweeping across the pattern, as a function of eye velocity. Head is stationary in space. Upper panel: stimulus continuously visible. Lower panel: stimulus visible for only 300 msec. (Note the occurrence of the Filehne illusion.)

This indeed did happen: when the patterns were made visible long enough (1 sec) to generate a visual component in the reference signal, the lowest spatial frequency pattern created an inverted Filehne illusion and increasing spatial frequency reduced $G_{ref}$. At the highest spatial frequency $G_{ref}$ even became less than 1 again.[9] It is interesting to note that when the gratings were presented only briefly (300 msec) during the pursuit eye movement, the normal Filehne illusion was again always observed ($G_{ref}$ being approximately 0.8) and spatial frequency had no effect. This was in line with expectations because such briefly visible stimuli, whatever their spatial frequency characteristics, have no optokinetic potential.

The conclusion that reference signal gain can actually be modulated invalidates the empirical basis on which the compromise of dual mode theory rests. The point is that the empirical criterion, which makes it possible to identify whether a percept is direct or inferential, depends on

the issue of perceptual veridicality, an issue closely tied to the idea that extraretinal signals are always undersized.

The traditional claim of direct perception theory is that perceptual deviations from reality indicate a lack of information in the optic array, that is, particular invariants are absent, incomplete, or have changed structurally. Such instances do not reflect (deficient) characteristics of the perceptual picking-up mechanism but "impoverished" visual information in the environment, often believed to be an artifact of laboratory conditions. Normal, ecologically relevant percepts are thought to be veridical. (For some discussions of the central role of veridicality in direct perception theory, see Gyr 1972; Lombardo 1987; Ullman 1980.)

In inferential theory, the extent to which percepts deviate from reality reflects the extent to which the gain of extraretinal signals deviates from 1. Since the Mack and Herman (1973) studies on the Filehne illusion (see above) it has been assumed that extraretinal signals have a gain less than 1. Consequently, inferential theory has always found it difficult to explain instances of really veridical perception (see, e.g., Matin 1982).

These contradictory views have (implicitly) led to the decision rule of dual mode theory: if a percept is not veridical, this shows that it must have been mediated inferentially, that is, with the help of (insufficient) extraretinal information; and if the percept is veridical, it must have been mediated directly (for some examples of this reasoning, see Bridgeman & Graziano 1989; Mack 1978; Matin 1982; Stark & Bridgeman 1983). The evidence from the threshold experiments mentioned above shows the flaw in this argument: it is the implicit but mistaken belief that inferential perception should always be biased because the reference signal is always undersized. This is not true. Reference signal gain is not a constant. Hence inferential perception may or may not be veridical. Perceptual veridicality thus becomes a matter of degree and depends on whether or not (and how much) $G_{ref}$ deviates from 1. The present conclusion that reference signal gain is not fixed but can be modulated by retinal flow thus destroys the criterion for distinguishing between direct and inferential perceptual modes and thus invalidates its empirical base.

The present notion of a visual component in reference signals provides a new explanation (without the need for dual mode theory) of why under normal daylight circumstances no illusory motion of the world occurs during an eye movement: such illusions only happen if $G_{ref}$ differs significantly from 1. Although efference copy components in reference signals are indeed too small, the reference signals themselves usually are not: eye movement induced retinal image flow generates an additional compensatory visual component (the compensation need not be very precise: $V_{ref}$ must only be enhanced enough to make its difference with $V_{ret}$ less than one JND). Actually, the reason efference copies associated with pursuit eye movements are undersized may be that if they were not undersized, an eye movement induced visual component would oversize the reference signal, which could create illusory motion of the world.

The present model is also able to explain center-surround induced motion without using the concepts of "object-relative" and "subject-relative" motion: when the stationary center stimulus is fixated with the eyes, the moving surround induces image flow across the retinae and this generates a (relatively small) reference signal. The image of the center stimulus, however, does not move on the retinae and thus generates a zero retinal signal. The center stimulus is hence perceived as moving in space. When the surround is pursued with the eyes, the illusion corresponds to the Filehne illusion: the small stationary stimulus seems to move against the eyes during a pursuit eye movement (see sect. 5.3 for a quantitative treatment of induced motion).

**5.2. Velocity perception.** We are now in a position to investigate some basic assumptions of direct perception theory. To this purpose we will begin with a closer look at Figure 3. Imagine a horizontal line cross-secting this graph. Along this line $V_{ret}$ remains constant, which means that we always have the same retinal image flow characteristics (invariants): those present in the retinal image flow at the intersection between the vertical axis and the horizontal line. When we move from left to right along this horizontal line, however, the percept varies. First the stimulus is seen to move against the eyes, but then, with increasing eye velocity, the perceived velocity of the stimulus is reduced until, at a certain eye velocity, the (against-) threshold is reached. After this point the stimulus is seen as stationary across a certain range of eye velocities. At the end of that range the with-threshold is reached. Now the stimulus is again perceived as moving, but in the other direction (with the eyes) and now its perceived velocity increases with eye velocity. In other words, all percepts of motion, stationarity, direction, and velocity depend on the ratio between retinal image velocity and eye velocity (in space). This means that, contrary to the claims of direct perception theory, the invariants present in a particular instance of image flow themselves have no fixed perceptual significance. In defence of direct perception theory, it might be postulated that the invariant that must be "picked up" to perceive object motion could be a "higher order" one (similar to the one mentioned in Note 2), consisting of the ratio between a normal invariant present in the retinal image flow ($V_{ret}$) and eye velocity information. But that would be contradictory to the basic idea of direct perception theory that the percept of object motion derives exclusively from retinal information. The point is that such a "higher order" invariant actually represents the main inferential principle: in addition to retinal information, eye movement information is always necessary.

The claim that the above-threshold perceived velocity of a visual stimulus depends on the relation between retinal image velocity and eye velocity (in space) is incompatible with direct perception theory for a further reason. According to this theory, eye movements are considered exploratory information sampling activities, necessary to "pick up" invariants. They do not (i.e., should not) affect percepts of object motion. If anything, they might enhance the quality of such percepts, but they do not define them (see, e.g., Gibson 1979, p. 219).

In terms of the present model, the claim that perceived stimulus velocity depends both on how the image moves across the eyes and on how the eyes move (in space) can be formalized as follows: perceived stimulus velocity depends on how much the retinal and reference signals differ, minus the JND between them, or

$$V_{est.s} = V_{ref} - V_{ret} - JND \quad (9)$$

where $V_{est.s}$ is the subjectively perceived velocity of the stimulus in space and $V_{ref}$ and $V_{ret}$ the magnitudes of the concurrent reference and retinal signals respectively. The threshold is represented by the additional requirement that $V_{est.s}$ remains zero as long as $|V_{ref} - V_{ret}| \leq$ JND. Note that when the eyes move faster across a stimulus, $V_{ref}$ and $V_{ret}$ increase equally, so their difference remains the same; however, the JND grows (Weber's law), reducing $V_{est.s}$. Hence the present model predicts that during (faster) pursuit eye movements we should underestimate stimulus velocity in proportion to the increased JND, or, stated differently, *$V_{est.s}$ should depend on eye movement induced changes in the thresholds for motion.*

To test this prediction, a velocity magnitude estimation experiment was carried out in which stimulus velocity was judged while pursuit eye movements – of various velocities – were made across the stimulus pattern (Wertheim & Van Gelder 1990). The results showed that when the stimulus moved in the *same* direction as the eyes, $V_{est.s}$ was indeed underestimated as much as the with-threshold for motion was elevated.

When stimuli moved *against* the eyes the underestimation of $V_{est.s}$ was less pronounced and with high stimulus velocities it was even absent. One explanation is that the high retinal image velocity afferents that occur in against-the-eyes conditions may not so easily pass the low temporal band-pass gating mechanism in the optokinetic pathway (see Fig. 1). This would decrease the (visual component in the) reference signal, that is, reduce $V_{ref}$ in Equation 9. $V_{est.s}$ then increases, because the difference between $V_{ret}$ and $V_{ref}$ increases ($V_{ret}$ is always larger than $V_{ref}$ when stimuli are perceived as moving against the eyes – see Fig. 3). That counteracts the underestimation effect. Another explanation could be as follows: when a stimulus is perceived as moving in the same direction as the eyes, $V_{ret}$ is always smaller than $V_{ref}$ (see Fig. 3). Hence, in Equation 9, $(V_{ref} - V_{ret})$ is positive. As soon as it grows larger than one JND, $V_{est.s}$ increases from its initial zero level. But when stimuli are perceived as moving against the eyes, $V_{ret}$ is larger than $V_{ref}$ (see Fig. 3), which means that the factor $(V_{ref} - V_{ret})$ is negative. As long as the absolute value of the factor $(V_{ref} - V_{ret})$ remains less than one JND, $V_{est.s}$ remains zero, that is, below threshold, but as soon as it grows larger than one JND, the absolute value of $V_{est.s}$ in Equation 9 becomes larger than two JND. Thus a discontinuity may occur immediately above the against-threshold: $V_{est.s}$ does not gradually increase from zero but jumps to a higher level, canceling the velocity underestimation effect of the increased threshold.

An effect opposite to the threshold-related underestimation of stimulus velocity with stimuli moving in the same direction as the eyes should occur when the eye movement is stopped abruptly (e.g., when the fixation point sweeping across the stimulus pattern is suddenly arrested). This reduces the threshold and the stimulus should thus suddenly be perceived as accelerating, that is, as moving faster than when the eyes were still moving. This "acceleration illusion" was also reported by Wertheim and Van Gelder (1990), who showed it to be independent of other factors, such as the sudden change in $V_{ret}$ itself or in the relative velocity between the (images of) the stimulus pattern and the fixation point.

The underestimation phenomenon with stimuli that move in the same direction as the eyes explains the so-called Aubert-Fleischl phenomenon: the perceived velocity of a stimulus is less when it is pursued with the eyes than when it moves – with the same speed – across stationary eyes (Aubert 1886; 1887; Dichgans et al. 1969; 1975; Fleischl 1882; Gibson et al. 1957; Mack & Herman 1972). The phenomenon also occurs in a visually "rich" environment and has been recognized as anomalous in direct perception theory (Gibson et al. 1957). The present model explains the phenomenon as being identical to the velocity underestimation phenomenon during pursuit eye movements: when a stimulus is tracked visually, it moves in the same direction as the eyes and thus its velocity is underestimated. The fact that the stimulus is actually tracked with the eyes is irrelevant (for a quantitative analysis of this claim, see Wertheim & Van Gelder 1990).

This explanation obviates another slightly different version of dual mode theory, one originally designed to explain the Aubert-Fleischl phenomenon (Dichgans & Brandt 1972). According to this version, we perceive motion either in an "afferent mode" from image motion across (stationary) eyes or in an "efferent mode" by identifying object motion with ocular motion, that is, during ocular pursuit of the stimulus (actually, the "efferent mode" has also been considered as one of three modes of visual perception – see e.g., Wallach et al. 1982; Wallach 1987 – the other two being related to retinal image-motion cues and to object-relative motion cues). The "efferent mode" is presumably less precise, yielding slower velocity percepts. The modes have been identified with the direct and inferential modes mentioned earlier (Mack 1986; Mack & Herman 1972), the slower percepts of the "efferent mode" being explained as caused by the underregistration of eye velocity in the efference copy.

It is interesting to note that Dichgans et al. (1975) reported that the Aubert-Fleischl phenomenon was less pronounced with low than high spatial frequency stimuli. The reason was that the perceived velocity of gratings moving across stationary eyes was reduced with lower spatial frequencies and this did not happen when the gratings were pursued with the eyes (see also Diener et al. 1976). In terms of the present model this is explained as follows: when gratings move across stationary eyes they generate retinal flow, which induces a reference signal that consists only of a visual component. Low spatial frequency gratings are more optokinetic than high spatial frequency ones, however. Hence, the former should induce larger reference signals than the latter, that is, larger JNDs (Weber's law), and thus higher thresholds. Since, as explained above, higher thresholds create slower perceived velocities, low spatial frequency stimuli will appear to move more slowly across stationary eyes than high spatial frequency stimuli. When the gratings are tracked with the eyes spatial frequency has no effect, because there is no image flow across the retinae, that is, no visual modulation of reference signals.

This also explains the stationarity tendency of large stimuli: they are simply more optokinetic than small ones. They accordingly have higher motion thresholds and their perceived above-threshold velocities are correspondingly reduced. Thus, there is no need to assume that large

stimuli tend to act as perceptual frames of reference (Mack & Herman 1978) – an assumption that is in any case questionable: a frame of reference does not define its own motion or stationarity.

Patterns moving across the retinal periphery also seem to have more optokinetic potential than when they move centrally (Dichgans & Brandt 1978). Thus, when a stimulus moves continuously across the retinal periphery of stationary eyes it may gradually produce quite a large reference signal (composed of only a visual component). The difference between retinal and reference signal is hence gradually reduced, which should result in a decrease of perceived stimulus velocity. In some cases the difference may even become less than one JND, causing the stimulus to appear stationary. Such phenomena have indeed been reported (Cohen 1965; Hunzelmann & Spillmann 1984; MacKay 1982).

**5.3. Absolute versus relative motion perception.** So far, in referring to the present model, the terms "stimulus velocity," "threshold for motion," or "perceived motion" have meant motion of objects *relative to external space* (i.e., 3-D "Newtonian" space, as defined by the horizontal surface of the earth and its gravitational field). Henceforth, this will be termed "absolute motion." Now let us consider the perception of motion of objects *relative to each other*, which will be called "relative motion" (see Kinchla, 1971, for a similar use of the terms absolute and relative motion).

Assume that the eyes sweep across two stimuli, S1 and S2, moving relative to each other. According to Equation 9 (sect. 5.2), the subjectively estimated absolute velocity of an S1 ($V_{est1.s}$) equals the difference between the eye movement induced reference signal and the retinal signal ($V_{ret1}$), minus the JND:

$$V_{est1.s} = V_{ref} - V_{ret1} - JND \quad (10)$$

Similarly, with respect to S2 we may write:

$$V_{est2.s} = V_{ref} - V_{ret2} - JND \quad (11)$$

The subjectively perceived velocity of S1 relative to S2 ($V_{est1\Delta2}$) equals the difference between $V_{est1.s}$ and $V_{est2.s}$. Hence:

$$V_{est1\Delta2} = V_{est1.s} - V_{est2.s} = V_{ret2} - V_{ret1} - JND\sqrt{2} \quad (12)$$

This means that the perceived velocity of two stimuli relative to each other should be independent of how the eyes move (i.e., of reference signals), depending only on the difference between the two associated retinal image velocities minus a noise factor.[10]

Equation 12 is of course subject to the condition that $V_{est1\Delta2}$ remains zero (below threshold) whenever $|V_{ret2} - V_{ret1}| \leq JND\sqrt{2}$. In terms of Weber's law this means that at the threshold for relative motion between S1 and S2,

$$\frac{V_{ret1} - V_{ret2}}{V_{ret1}} = \text{constant} \quad (13)$$

This prediction was tested (Wertheim & Niessen 1986) by measuring the threshold for relative motion between two identical stimulus patterns while subjects tracked a fixation point sweeping (at various velocities) across both stimuli. The results (Fig. 5) confirm Equation 13.

This finding is theoretically important in the debate between direct and inferential theory. The point is that, since retinal image velocity is always equal to the difference between eye velocity in space ($V_{eyes.s}$) and absolute stimulus velocity ($V_{stim.s}$), Equation 12 can be written as

$$V_{est1\Delta2} = V_{stim1.s} - V_{stim2.s} - JND\sqrt{2} \quad (14)$$

Hence, not only does the percept of relative motion between objects depend exclusively on retinal afferents (Equation 12), but it is also always veridical, because it corresponds to the physical description of how the objects move in space – apart from a noise factor (Equation 14). These conclusions agree with the basic claims of direct perception theory, even though they follow from inferential reasoning; hence, with respect to relative motion, there is no disagreement whatsoever between the two approaches. It seems that the debate between the two theories actually reflects a failure to distinguish between percepts of relative motion (which are independent of reference signals) and percepts of absolute motion (which depend on reference signals). To state that both theories concern the perception of "motion" is to invite confusion. We should separate the concept of "motion" into absolute and relative motion, and correspondingly distinguish between percepts of absolute and relative motion (e.g., between seeing whether a car moves on the road and seeing whether it moves relative to another car).

In retrospect, this makes sense: inferential theory always concerns percepts of absolute motion, even if not mentioned explicitly (as for example in the literature on the Filehne illusion). Hence it refers to illusions caused by properties of reference signals. Direct perception theory is concerned with perception in natural "ecologically relevant" environments, that is, with the perception of relative motion of objects moving against a visual background. If the background is seen as stationary in space, the relative motion of an object against the background equals its absolute motion in space. Hence all percepts of motion become veridical. To illustrate this, let S1 be an object moving against a visual background S2. The subjectively perceived *absolute* velocity of the object S1 can be expressed as:

$$V_{est1.s} = V_{est2.s} - V_{est1\Delta2} \quad (15)$$

or

$$V_{est1.s} = V_{est2.s} - (V_{stim2.s} - V_{stim1.s} - JND\sqrt{2}) \quad (16)$$

or

$$V_{est1.s} = (V_{ref} - V_{ret2} - JND) - (V_{stim2.s} - V_{stim1.s} - JND\sqrt{2}) \quad (17)$$

Equation 16 shows that if a background is stationary ($V_{stim2.s} = 0$) and is also perceived as such ($V_{est2.s} = 0$), the absolute motion of the object, $V_{est1.s}$, is perceived veridically (apart from a noise factor). Equation 17 shows that this is true even in cases where the gain of reference signals differs from 1, if only the JND between $V_{ref}$ and $V_{ret2}$ is large enough to maintain a percept of background stationarity. Note that this is an example of a veridical percept of absolute motion in the presence of an inappropriately sized reference signal (visual capture).

If, on the other hand, the background moves in space, its estimated absolute velocity, $V_{est2.s}$, is usually not equal to $V_{stim2.s}$ (e.g., because of a size or spatial frequency induced stationarity tendency, or because it is perceived

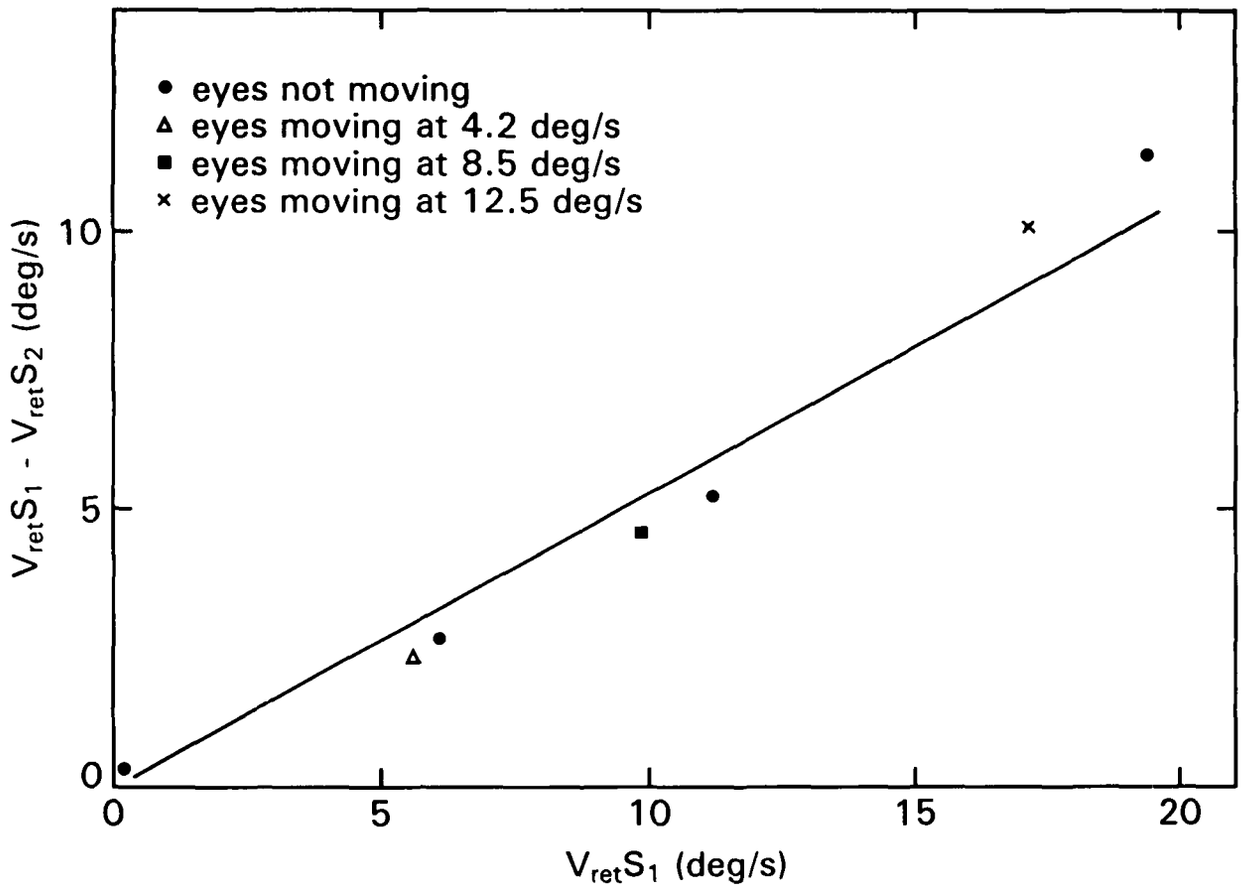


Figure 5. The difference between retinal image velocities of two identical stimulus patterns, S1 and S2, at the thresholds for detecting relative motion between them, as a function of retinal image velocity of one of the two patterns. Measurements taken with moving eyes (during ocular pursuit of a fixation point sweeping at various velocities across both patterns) and with stationary eyes (fixation point stationary on the screen). Head is fixed in space.

during an eye movement). Equation 16 shows that the percept of absolute object motion, $V_{est1.s}$, then becomes less veridical, that is, unequal to $V_{stim1.s}$. This provides a quantitative description of center-surround induced motion: a stationary object ($V_{stim1.s} = 0$), seen against a moving surround ($V_{est2.s} < V_{stim2.s}$), is perceived as moving in space ($V_{est1.s} \neq 0$).

Note that this view of induced motion differs from the one given by Wallach or dual mode theory (see sect. 5.1), according to which the crucial element of induced motion is the dominance of "object-relative" motion cues. Seen from the present perspective, however, induced motion is an illusion of absolute motion (see also Kinchla 1971). The illusion is not that the center dot seems to move relative to its surround (this is seen correctly), but that the center dot seems to move in space. This is illustrated by the fact that we can also express induced motion formally by substituting ($V_{eyes.s} - V_{stim1.s}$) for $V_{ret1.s}$ in Equation 10:

$$V_{est1.s} = V_{ref} - V_{eyes.s} + V_{stim1.s} - JND \qquad (18)$$

Thus, if we fixate the stationary stimulus ($V_{stim1.s} = V_{eyes.s} = 0$), it is seen to move *in space* with a velocity proportional to the visually induced $V_{ref}$ created by the surround image flow across the retinae (minus the JND). Note that in such circumstances the illusion should develop gradually, because the induction of a visual reference signal is a gradual process (actually, we should expect the duration of this process to become shorter with slower surround motion; see the discussion of the generation of vection in sections 2 and 3). If, on the other hand, the eyes track the surround, induced motion should be immediate, because $V_{ref}$ then consists of just an efference copy component (no image flow across the retinae), which is about 20% smaller than $V_{eyes.s}$. $V_{est1.s}$ is then proportional to ($V_{ref} - V_{eyes.s}$). If this is larger than one JND, induced motion (the Filehne illusion) occurs. There is indeed some empirical evidence (see Reinhardt-Rutland 1992) that supports this claim, that induced motion develops gradually when the eyes fixate the stationary center stimulus but is immediate when the eyes fixate the moving surround. A related prediction would be that no induced motion should occur if $V_{ref}$ approximates $V_{eyes.s}$ (see Equation 18), that is, if reference signal gain is close to 1. That may happen if the eye movement sweeps across the whole induced motion display in a normally illuminated environment, generating a visual component next to the efference copy.

Although relative motion between objects is not affected by reference signals, it may be affected by eye movements for another reason: eye movements made across various moving stimuli may increase retinal image velocities, that is, retinal signals. This would not affect the differences between these retinal signals, but it would increase the JNDs between them (Weber's law). Accord-

ing to Equation 12, this should elevate the threshold for relative motion of objects with respect to each other (see Murphy 1978; Nakayama 1981) and reduce perceived relative velocities. This may cause a "freezing illusion": imagine a screen on which various stimuli move relative to each other with different but not too high velocities. An eye movement across the screen will then increase the relative motion thresholds so much that the display seems to become motionless, as if suddenly frozen. Nakayama (1981) actually predicted that such a phenomenon should cause the disappearance of kinetic depth perception, which depends on the detection of small differences in relative velocity between many stimuli on a screen (e.g., Braunstein 1976; Wallach & O'Connell 1953).

**5.4. Interfacing ego- and object-motion perception and visual-vestibular interactions.** The main function of the vestibular apparatus is to signal head movements to the brain. According to direct perception theory, this is only confirmatory information because the visual system does the same through visual kinaesthesis. No particular interaction between these two kinds of information is postulated (but see Note 2) and the visual perception of object motion is thought to be independent of vestibular stimulation. As mentioned earlier, eye and head movements are viewed as exploratory information sampling activities, which, if anything, should only improve perception. In the direct perception literature there is only one exception. This is the case of overstimulation of the vestibular apparatus. Such overstimulation yields a percept, or awareness, of self-motion which differs from that of visual kinaesthesis: orderly percepts are disturbed and the observer experiences a sense of disorientation, part of which consists of perceiving the visual world as moving. However, since direct perception theory is mainly concerned with normal (ecologically relevant) perceptual conditions, it has no formal model for what happens in such cases, apart from the assumption that such conditions make retinal events "obtrusive" (Gibson 1968).

As shown above, the present model differs from this view. Although it agrees that percepts of relative motion may indeed be independent of vestibular stimulation, this is not the case with percepts of absolute motion. Here reference signals are always involved and they may include a vestibular component. The idea of a vestibularly induced kind of efference copy was first proposed by Sperry (1950), who called it a "corollary discharge." This term often features in the inferential literature (see, e.g., Jeannerod et al. 1979). To a certain extent the present model agrees with this idea. The difference, however, is that according to the present model vestibular stimulation does not generate an independent signal but a component in the reference signal. This is not just a matter of semantics, because the further assumption that reference signals may also include a visual component now introduces a new element: it implies an interaction between visual and vestibular information in the reference signal (i.e, in the brain's estimate of how the eyes move in space). As a result, the neurophysiological literature on visual-vestibular interactions, which consists mainly of research on ego-motion perception, now becomes relevant to the study of the visual perception of object motion.

Although this literature is much too large to review in the present target article, it should be mentioned that it often includes speculations about possible neural substrates of what we have called reference signals. For example, physical movements of the eyes in space – irrespective of whether they are caused by eye movements in the head, head movements, or both – have been recognized in the output activity of cells in the vestibular nuclei (Berthoz et al. 1981; Cohen 1981; Fuchs & Kim 1975; McCrea et al. 1981; Yoshida et al. 1981), in the flocculo-nodal lobe of the cerebellum (Cohen 1981; Lisberger & Fuchs 1987a; 1987b; Stone & Lisberger 1990a; 1990b) and in the vestibular cortex (Büttner & Buettner 1978; Büttner & Henn 1981). The activity of some of these cells is in fact modified by visual stimulation, that is, by retinal image motion or optic flow (see, e.g., Nagao 1988; Noda 1986; Waespe & Henn 1981; Waespe et al. 1981; Watanabe 1984). The time course of this modulation differs among cells, but seems to be slowest in the vestibular cortex. Hence the output activity of cells in that area might represent the neurological substrate of reference signals (see also Straube & Brandt 1987). The neural networks of which these cells are part have largely been charted out (e.g., Barthélémy et al. 1988; Berthoz & Melvill Jones 1985; Cohen & Henn 1988; Henn et al. 1980; Ito 1982; Precht 1982; Straube & Brandt 1987; Waespe & Henn 1979; Xerri et al. 1987; 1988). They are sufficiently complex to allow for a subsystem such as described in Figure 1 (or Fig. 7 below).

To illustrate how closely object-motion perception is linked with self- and ego motion, let us analyze the occurrence of saturated vection, not as described earlier for circular vection in an optokinetic drum (sect. 3) but as it occurs in an everyday kind of situation. Imagine a train engineer seated at the front of a train looking straight ahead and making no head movements. When the train begins to move, it accelerates. The vestibular apparatus, which only reacts to accelerations, responds. Integration of the response produces information about head velocity in space (recognizable at the level of single cell activity – e.g., Benson 1990); this, according to the present model, is used to generate a reference signal that provides the visual system with an estimate of how fast the eyes move in space (see Fig. 1). If the estimate is not too much in error, the reference signal will be approximately equal to the retinal signal evoked by the moving image of the visual world. Since small differences are masked by the JND, this keeps the world perceptually stable. When the train reaches a constant velocity, the vestibular apparatus becomes silent, but now vection takes over to maintain the sensation of ego motion, that is, the decreasing vestibular component in the reference signal is replaced by a gradually growing visual one. The reference signal thus maintains its size and the percept of a stable world remains. Without this visual-vestibular interaction the reference signal would decrease with the decrease of vestibular reactivity and the world would lose its stability, seeming to "rush" toward the observer. This illustrates an important ecological function of the visual-vestibular interaction taking place within reference signals: to interface the perception of a stationary world with the perception of ego motion.

Such interactions do raise a problem for the present model, however. It is well established that the time course of development of a vestibular response differs from that of an optokinetic one (see, e.g., Dichgans & Brandt 1972;

Henn et al. 1980). Hence the development of the various reference signal components is not always synchronous. For example, the vestibular apparatus reacts fast to relatively high frequency self-movements but it may take longer before a visual component is fully grown. Another problem is that the integration of vestibular information into velocity information is not perfect and depends on the frequency range within which the vestibular system responds (Benson 1990). Thus, it is unlikely that $G_{ref}$ is continuously close to 1 during activities like running or dancing. Nevertheless, we usually do not experience illusory motion of the visual world when engaged in such activities. The answer is probably that the JNDs of reference signals which include a vestibular component are very large. That would mask quite large unwanted differences between retinal and reference signals.

To investigate this issue, Wertheim and Bles (1984) measured the JND of reference signals during ego motion. They rotated subjects sinusoidally (0.05 Hz, various amplitudes) on a rotating chair inside a totally darkened optokinetic drum, which could be rotated independently around the subject. The inside of the vertically striped drum could be illuminated briefly (400 msec). This was done at the point where the subject rotated at peak velocity. Thus the drum wall served as a (full field) stimulus pattern that could be moved in space with or against the direction of the subjects' ego rotation in space. The two opposite thresholds for absolute motion of the drum wall were measured at various ego velocities, yielding JNDs of 35% of ego velocity. This is similar to the findings of Wallach (1985), who reported that the distance between the with- and against-thresholds for perceiving object motion in space is very large when measured with subjects walking alongside the stimulus. His results suggest JNDs that amount to 40% of ego velocity.

Such large JNDs should indeed facilitate a smooth interfacing of ego motion with percepts of environmental stationarity. However, the price is a dramatic increase of the perceptual thresholds for absolute object motion during ego motion and, because of that, a strong underestimation of absolute object velocity during ego motion. Such effects are indeed well documented (Berthoz & Droulez 1982; Büchele et al. 1980; Pavard & Berthoz 1977; Probst et al. 1980; 1984; 1986).

But the mechanism that serves the ecological function of interfacing percepts of ego motion and environmental stability has more drawbacks – percepts may be produced that are exactly opposite to what they should be: a really moving scene can erroneously be seen as stationary. This happens when we see a moving train close to the window of our own stationary train: the moving train acts as an optokinetic stimulus and creates a sensation of ego motion. It thus generates a reference signal that grows in size until its difference with the retinal signal (encoding the retinal image velocity of the moving train) becomes less than one JND. The moving train is then erroneously seen as stationary in space. This is basically the same phenomenon as the development of saturated circular vection in an optokinetic drum.

The opposite, illusory motion of an actually stationary scene, may also occur. A common example is what happens after a period of extreme vestibular stimulation: neural activity of cells in the central areas upon which the vestibular afferents converge dies out only gradually (as evidenced by a continuation of reflexive nystagmus eye movements, called "afternystagmus"; see, e.g., Henn et al. 1980). Hence, a residual vestibular component remains present in reference signals and oversizes them, causing illusions of environmental motion. Note that this explains the perception of environmental motion during dizziness. Hence, the present model differs from the traditional inferential view that such percepts are caused by an absence of efference copies during such reflexive nystagmus eye movements.

Similar reasoning may apply to the movement after effect, MAE (when a stimulus pattern is suddenly stopped after having moved for a while across stationary eyes, it is perceived as moving slightly in the opposite direction; the illusion may last many seconds, during which the threshold for object motion in the original direction is elevated). The most common, but still somewhat controversial, explanation of MAE is in terms of fatigued direction selective cells (see Denton 1977; Favreau 1976; Moulden 1975; Sekuler et al. 1982). The present explanation is different: when the lights in an optokinetic drum are suddenly extinguished, vection decays only slowly and reflexive nystagmus eye movements continue for a while. This suggests a continuation of central neural activity upon cessation of retinal flow (Henn et al. 1980). Hence visually induced (components in) reference signals may also decay gradually after retinal flow stops. As long as they last, however, a stationary stimulus, viewed with stationary eyes, will be seen as moving in space (see Equation 18). The JND associated with that residual reference signal explains the elevated threshold for motion in the direction of the original retinal flow (i.e., in the vectorial direction of the reference signal).

These examples show that, according to the present model, an appreciation of visual-vestibular interactions is needed to explain phenomena in the field of visual object-motion perception. The inverse is also true, however: the present experimental paradigm can serve as a tool in research on ego-motion perception and visual-vestibular interactions. The method for measuring reference signal magnitude (and gain) by measuring retinal image velocity at the PSS can be used to measure the gain of the response of the various parts of the equilibrium system (the semicircular canals, which react to angular accelerations, and the otoliths, which respond to linear accelerations of the head in space).

An example of such a study is the Wertheim and Bles (1984) experiment cited earlier, in which subjects were rotated inside an optokinetic drum. That experiment was not only designed to measure the JND between retinal signals and vestibularly induced reference signals; it also attempted to measure the response of the semicircular canals (neglecting possible kinaesthetic feedback) and its interaction with reflexive nystagmus eye movements during ego rotation in darkness. According to the present model, such ego rotation should induce reference signals that consist of the vectorial sum of $V_{head.s}$ (the response of the semicircular canals) and a $V_{eyes.h}$ component. The presence of a $V_{eyes.h}$ component stems from the reflexive nystagmus eye movements that occur during stimulation of the semicircular canals (in a normally illuminated environment nystagmus eye movements serve to stabilize the visual gaze in space during ego motion, but they also

occur in darkness). When a subject is rotated around the vertical axis on a rotating chair, nystagmus consists of slow phase smooth compensatory eye movements in the direction opposite to head rotation, alternating with fast phase recuperating saccades in the same direction as head rotation. Thus, during slow phase nystagmus eye movements, reference signal magnitude should be smaller than during the suppression of nystagmus (nystagmus suppression occurs when we ask the rotating subject to fixate the eyes on a small head-stationary fixation point), because $V_{eyes.h}$ then approximates zero.

Wertheim and Bles tested this hypothesis by performing their drum experiment with and without the suppression of nystagmus. They showed that (at the 0.05 Hz ego-rotation frequency used in this experiment)

$$V_{ref} = V_{ret.PSS} = -1.07\, V_{head.s} - 0.72\, V_{eyes.h} - 0.74 \quad (19)$$

Hence, $V_{ref}$ was indeed decreased by slow phase nystagmoid eye movements (during slow phase nystagmus eye movements the sign of $V_{eyes.h}$ is opposite to that of $V_{head.s}$) and increased when they were suppressed. Note that this means that slow phase nystagmus eye movements do in fact generate efference copies in reference signals in which only 72% of $V_{eyes.s}$ is registered, just as in the case of pursuit eye movements. This finding is at variance with the traditional view, mentioned earlier, that nystagmoid eye movements do not generate efference copy signals (Howard & Templeton 1966; Johnstone & Mark 1970; 1971; 1973; Kornhuber 1974; Leibowitz et al. 1982; Raymond et al. 1984; but see Bedell et al., 1989, and Mittelstaedt, 1990, for experimental findings and theoretical views that agree with the present observation).

Since in this experiment subjects were rotated along their vertical axis in total darkness, the $V_{head.s}$ term in Equation 19 actually reflects the gain of semicircular canal afferents (although some kinaesthetic feedback may also have been present). The small (7%) overregistration of head velocity in these afferents explains the oculogyral illusion (when an observer is rotated in complete darkness and nystagmus is suppressed with a head stationary fixation point, this fixation point, rotating with the observer, seems to move slightly faster than the observer, see, e.g., Elsner 1971; Graybiel & Hupp 1946; Howard 1982; Ross 1974; Whiteside et al. 1965): the velocity of the fixation point in space is overestimated because it corresponds to the difference between a zero retinal and a slightly oversized reference signal.

Recently, the characteristics of reference signals created by linear accelerations of the head in space – the characteristics of the otolith afferent response – have also been investigated in a series of experiments at our laboratory (Wertheim 1992a; 1992b; Wertheim & Mesland 1993; Zeppenfeldt 1991). Here $V_{ret.PSS}$ was measured with subjects moving forward or backward on a linear track sled between two screens on which the stimulus (a checkerboard pattern) was flashed (300 msec). The subjects looked straight ahead (a fixation point was placed several meters in front of the endpoint of the sled's track) and thus perceived the stimulus patterns peripherally. The sled moved sinusoidally (at 0.15 Hz and with a 109.5 cm/sec peak velocity) and the experimental room was completely dark to prevent the creation of a visual component in the reference signal (no retinal flow from the environment). Reference signals were measured with the monitors placed at various positions along the sled's track (i.e., at various phases of the sinusoidal sled motion) and the best-fitting sinus through these data was calculated.

As can be seen in Figure 6, the results showed that this particular ego-motion profile created undersized reference signals with a gain of 0.76 and a small phase lead of 3.8 deg (a similar phase lead – of approximately 6 deg – can be calculated on the basis of a mathematical model of the otolith system; see Grant & Best 1987; Marcus 1992). If such experiments are performed with linear ego-motion sinusoids of other frequencies and amplitudes, the full transfer function of the otoliths may become known (again, under the assumption of neglectable kinaesthetic feedback).

In a similar experiment (see also Wertheim & Mesland 1993) we measured $G_{ref}$ at the point of maximum sled velocity (109.5 cm/sec) in darkness, but now we compared it to a condition with the lights on in the experimental room. In darkness, the reference signal was again undersized ($G_{ref}$ being 0.8), but when the lights were on, allowing for the generation of a compensatory visual component in the reference signal, $G_{ref}$ became 1. This pattern of results is remarkably similar to the one discussed with relevance to the Filehne illusion (see sect. 5.1).

The same logic is used in a current research project, in which we investigate whether the otolith response changes after adaptation of the equilibrium system (adaptation is induced by rotating subjects in the gondola of a centrifuge such that they sustain a force of 3G for periods between 1 and 2 hours; see Bles et al. 1989; Ockels et al. 1989; 1990; Wertheim 1992a; 1993; Wertheim et al. 1989).

The present paradigm might even prove useful in the clinical diagnosis of vestibular deficiencies. For example, one study (Wertheim et al. 1985) tested the hypothesis that resting level activity of the central vestibular system is abnormally noisy in schizophrenia. Functionally, this implies very noisy reference signals, that is, abnormally large JNDs between retinal and reference signals, even if no head movements are made. Very high thresholds for motion were indeed observed with such patients. Findings were similar with patients who were not schizophrenic but had been diagnosed as having a noisy vestibular apparatus.

**5.5. Conclusions.** The controversies between direct and inferential theories of motion perception may have originated at least in part from different and sometimes contradictory philosophical views (Gibson 1973; Lombardo 1987). However, on the empirical level, most of the debate stems from the puzzling observation that the data gathered in everyday situations often differ from those gathered in strictly controlled laboratory conditions. The present model provides a theoretical alternative to the two approaches for two reasons. First, it explains the "puzzling" differences by showing that the two approaches actually reflect research on different topics: direct perception theory is concerned with the perception of relative motion and inferential theory with the perception of absolute motion. Second, it describes (quantitatively) how the two topics relate to each other.

As a result, a certain compatibility is created between premises from direct and inferential theories, premises

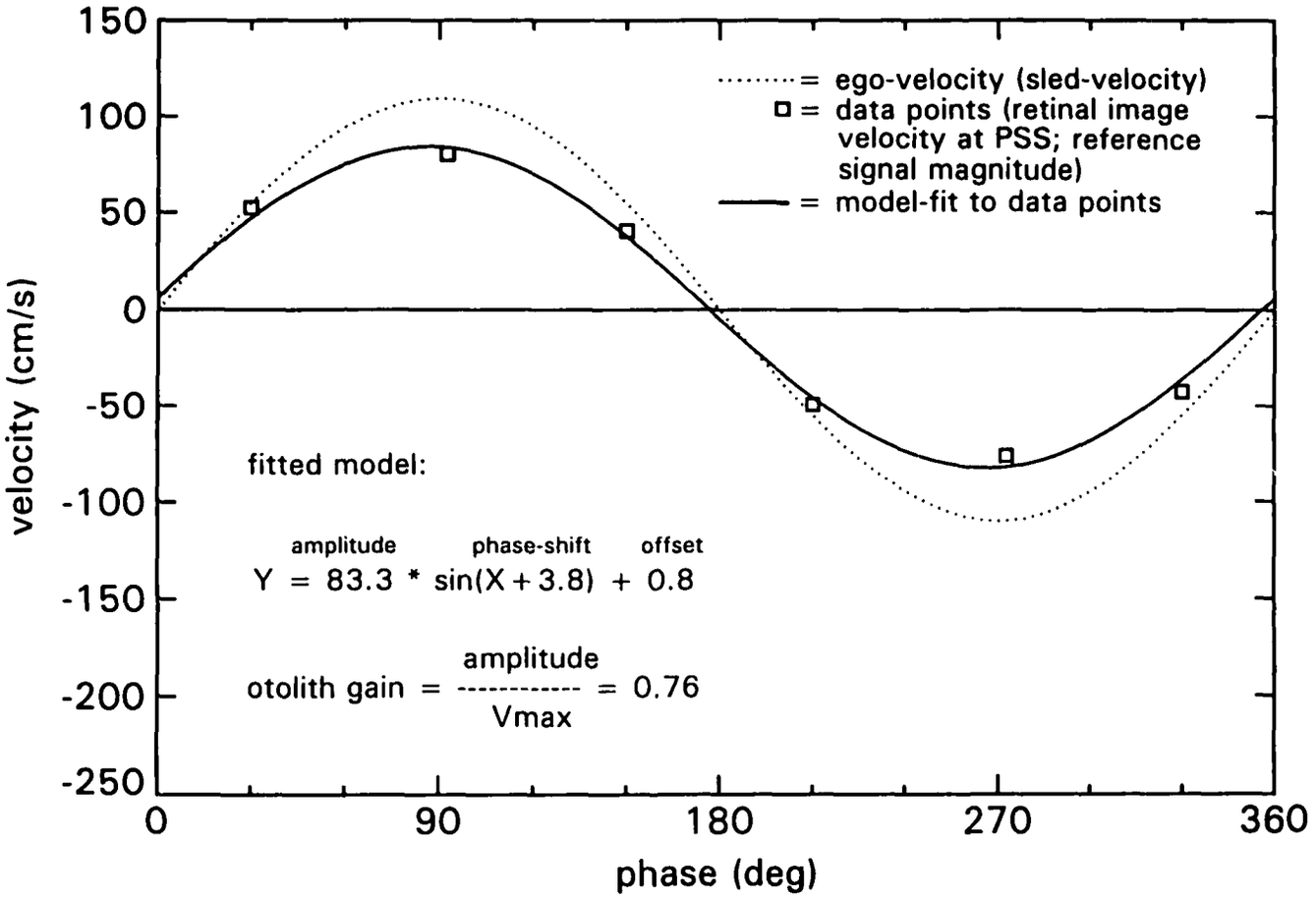


Figure 6. Modulation of the reference signal (drawn line) during sinusoidal linear ego motion.

that have traditionally been considered contradictory. Thus, in agreement with direct perception theory the perception of object motion may indeed stem exclusively from visual afferents, and retinal flow only contains information about ego motion and not about motion of the visual world. There is also agreement with the inferential assumption that information about how the eyes move in space is necessary for perceiving absolute object motion. However, on other issues the model diverges from direct and inferential theories. Thus, it disagrees with the direct perception assumption that self-motion is basically exploratory and only serves to upgrade perception. It also disagrees with the assumption that to perceive absolute motion the brain needs no estimate of how the eyes move in space. With respect to inferential theory, the present model replaces the concept of efference copy and corollary discharge with that of a (compound) reference signal. As it includes a visual component, the common assumption that it should be considered extraretinal and has a fixed gain is also abandoned. Finally, since the present model actually describes how percepts of self-motion and of object motion interface, it broadens the scope of the study of visual object-motion perception to include visual-vestibular interactions.

So far, this target article has been devoted to the description of the model and of an empirical paradigm which can be used to quantify its parameters and to test its predictions. The results of these empirical tests appear to support the model; their theoretical implications are shown to resolve most of the controversies between direct and inferential theory and seem to invalidate the theoretical rationale of dual mode theory. At the same time, new explanations have been given for many well-known phenomena in the field of motion perception. What remains now is to review theories and data that may point to deficiencies of the present model and directions for further research.

## 6. Problems and speculations

**6.1. The Post and Leibowitz model.** Post and Leibowitz (1985) have proposed a version of inferential theory that is at odds with the present model for two reasons. First, it assumes that a very large moving stimulus pattern always induces vection irrespective of whether its image moves across the retinae (according to the present model vection develops only through retinal image flow); in addition, such stimuli always cause reflexive optokinetic nystagmus. Second, according to Post and Leibowitz, efference copies – which, for the purpose of comparing their model with the present one, may be treated as reference signals – are proportional not to eye velocity, but to the effort invested in voluntary control of oculomotor activity. No effort is invested when eye movements are reflexive. Thus, optokinetic nystagmus generates no efference copies, but its suppression (by focusing the eyes on a head-stationary fixation point) takes effort, and this does evoke efference copies. With stronger optokinetic stimuli it presumably takes more effort to suppress optokinetic

nystagmus. Consequently, just as in the present model, efference copies are proportional to the force of vestibular and optokinetic stimulation. Many predictions derived from the present model therefore also follow from the Post and Leibowitz model.

The two models can be tested against each other, however, because they predict opposite effects when nystagmus is not suppressed. Consider a large stimulus pattern that moves sinusoidally in front of the observer, slowly enough and with an amplitude small enough to maintain continuous sinusoidal slow phase optokinetic nystagmoid eye movements without any fast phase saccadic eye movements. According to the Post and Leibowitz model, such conditions evoke sinusoidal vection, and since there is no efference copy (the eye movements are reflexive) and a zero retinal signal (no retinal image motion), the pattern should be seen as stationary in space. The present model predicts the opposite: first, in the absence of image flow across the retina no vection can develop; second, the absence of retinal image motion implies a zero retinal signal, but the slow phase eye movements generate nonzero (efference copy composed) reference signals. Hence, the pattern should be perceived as moving in space.

Such an experiment was recently reported (Mergner & Becker 1990). The stimulus consisted of a full-field shadow pattern moving sinusoidally across a semicircular screen. Subjects fixated a small fixation point, which was also projected on the screen and could move independently. It moved synchronously with the shadow pattern, having the same frequency but a different amplitude, that is, a different velocity. In their experiment, Mergner and Becker started out with the fixation point moving much more slowly than the shadow pattern, causing retinal image motion of the pattern. In this situation, sinusoidal vection always developed to saturation (at which point the shadow pattern appeared as stationary in space). They then gradually increased the velocity of the fixation point. Vection remained. However, at a certain moment fixation point velocity became equal to the velocity of the pattern, that is, it became part of the pattern. This is the critical condition, because in terms of the Post and Leibowitz model the slow phase reflexive nystagmus eye movements are now completely unobstructed by any voluntary effort to track the target. At this moment all subjects experienced a sudden elimination of vection, perceiving themselves as stationary and the pattern as moving in space, whatever the duration of the trial. This supports the present model and is contrary to the predictions of the Post and Leibowitz model.

**6.2. Retinal image flow and vection.** The Mergner and Becker (1990) experiment did indeed show that vection fails to develop in the absence of retinal slip (see also Fig. 1), but this poses a problem: in an optokinetic drum, circular vection occurs always (i.e., also when nystagmus is not suppressed). There may be two reasons for this.

First, if a full-field stimulus pattern is tracked with the eyes from extreme right to extreme left (as during the slow phase of optokinetic nystagmus) its image does not move across the retinae but illuminates different parts of them. To the visual system this is perhaps also a vection-inducing cue. Mergner and Becker used a shadow pattern with low contrast values that may have reduced the salience of this cue. We tried to test this idea with some pilot measurements: with a sinusoidally moving optokinetic drum (high contrast black and white stripes) vection always developed, including when the stripes were tracked with the eyes. However, one may also track the stripes of the drum with the head (eyes stationary in the head). The retinal image of the stripes then always illuminates the same retinal area. It appeared that in such cases vection did indeed fail to develop.

Second, it is possible that the repetition of brief instances of image flow during the fast phases of normal optokinetic nystagmus has the potential to induce vection. If so, the gating mechanism in the optokinetic pathway could be viewed as a velocity storage mechanism (Raphan et al. 1977) that can be loaded by brief repetitive instances of image flow across the retinae.

Another pertinent problem is that according to section 3, an optokinetic stimulus generates (a visual component in) reference signals because during vection the visual system assumes that the eyes move in space. But if this is so, should not vection always occur when an optokinetic reference signal (component) is generated? Clearly this is not always the case. For example, when pursuit eye movements across a visual background generate image flow across the retinae, vection usually does not occur, not even if the background consists of a strong optokinetic stimulus pattern (such as the one used by Wertheim, 1987, to invert the Filehne illusion). This suggests that the common pathway on which optokinetic afferents and vestibular afferents converge branches off in two directions, one generating ego motion, the other (its corollary) converging on the reference signal. Different gating mechanisms (i.e., different thresholds) may then be associated with the two branches (see Fig. 7).

**6.3. A "visual efference copy"?** Ehrenstein et al. (1986a; 1986b; 1987), using a briefly visible point stimulus, reported that the Filehne illusion increased dramatically with stimulus presentation times below 300 msec, implying a very strong reduction of reference signal size (in some cases even to zero). This poses a problem: without its visual component, reference signal magnitude should remain constant, as it still contains an efference copy component, encoding about 80% of eye velocity in the head. Ehrenstein et al. measured the PSS with a forced-choice method of constant stimuli using only two response alternatives (motion with or against the eyes), excluding "no motion" responses. Since with extremely brief stimulus presentations motion perception may become ambiguous or impossible (see, e.g., Algom & Cohen-Raz 1984; Bonnet 1982; Henderson 1971; Johnson & Leibowitz 1976), this may have caused a response bias. We were unable to replicate Ehrenstein's finding with a larger stimulus pattern (Wertheim & Bekkering 1991; 1992) using our standard staircase method of limits, in which the two opposite thresholds are measured separately, thus always allowing for "no motion" responses. Reducing presentation times to 150 msec never yielded large Filehne illusions ($G_{ref}$ remained approximately 0.8). However, the JND increased dramatically (suggesting that such brief retinal afferents are quite noisy). With presentation times below 150 msec, the JND became so large that retinal image velocity at the against-the-eyes threshold reached the upper limit for detecting image

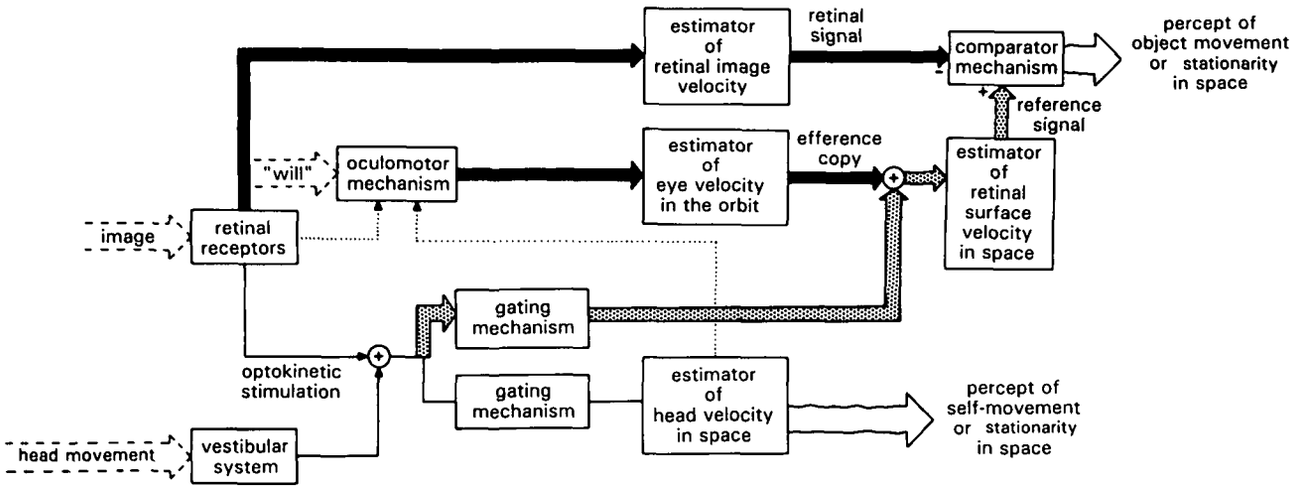


Figure 7. Adapted functional model describing the generation of reference signals and the interfacing of percepts of object and ego motion in space (see also caption, Fig. 1).

motion. In such cases subjects never perceive motion against the eyes, which means that the PSS cannot be measured.

In another experiment (De Graaf & Wertheim 1988), however, using a very high spatial frequency stimulus pattern (a little cloud of dots) visible for 300 msec in the retinal periphery, we did observe a very large Filehne illusion, suggesting a 0.5 reference signal gain. This is difficult to explain. One possibility is that the efference copy is not (as traditionally assumed) a neural corollary of the efferent command signals to the oculomotor musculature (Note 4) but stems from visual afferents (i.e., from visual kinaesthesis) that converge on the reference signal through a fast visual pathway with high spatiotemporal frequency band-pass gating characteristics. In the retinal periphery such a "visual efference copy" would be smaller than in the foveal area, because the retinal periphery is much less sensitive to high spatial frequencies. This could also explain the Ehrenstein effect: retinal afferents from a single-point stimulus that is visible extremely briefly might not generate such a "visual efference copy" at all, because they may not pass even this high spatiotemporal frequency gate.

Anatomically, such a fast visual channel could be included in the optokinetic pathway. There are some indications (Stone & Lisberger 1990a; 1990b) that the accessory optic pathway contains both a low and a high spatiotemporal band-pass gating mechanisms, as reflected in the different temporal characteristics of the simple and complex spikes of floccular Purkinje cells.

Although the idea of such a fast visual channel in the optokinetic (accessory optic) pathway is quite speculative, the idea is attractive, because it further blurs the distinction between direct and inferential theory: a "visual efference copy" would make the concept of visual kinaesthesis compatible with inferential theory. It might also help explain why vection can be instantaneous when an optokinetic drum is very slowly set into motion and optokinetic nystagmus is not suppressed.

On the other hand, the idea of a visual efference copy should not be embraced too easily because it also creates a serious problem: when a moving stimulus is properly tracked with the eyes in total darkness, there would be no efference copy, that is, no reference signal. Since there is also (almost) no retinal image motion, that is, no retinal signal, such a stimulus should be seen as stationary in space. This is not what happens: we do see stimulus motion under such conditions. This seems quite incompatible with the concept of a "visual efference copy."

**6.4. Signal magnitude.** Recently, we observed that the Filehne illusion is age dependent (Wertheim & Bekkering 1991; 1992): with very brief stimulus presentation times (150 msec), the usual illusion occurred with normal student subjects, but with older subjects it disappeared and with subjects over 50 years of age it was inverted (the correlation between $G_{ref}$ and age was approximately 0.7; n = 38). With longer stimulus presentations $G_{ref}$ gradually returned to approximately 1 for all ages. Thus, it appeared as if reference signals, (i.e., efference copies) grow (beyond proportion) with increasing age. A more plausible explanation, however, is that when people age it takes more time to register image velocity in the retinal signal. Extremely brief stimulus presentations would then yield undersized retinal signals. Very high retinal image velocities are then needed to augment retinal signals enough to make them larger than reference signals, that is, to reach the against-the-eyes thresholds. Hence stimulus velocity at the against-the-eyes threshold would become very high, and consequently $V_{ret.PSS}$ would increase, creating the impression of a very large reference signal.

This illustrates a particular complexity of the present model: in section 4, reference signal magnitude was operationalized as retinal image velocity at the PSS. However, this presupposes a proper encoding of retinal image velocity in the retinal signal. If retinal signals underregister image velocity (i.e., if their gain is less than 1), an inverted Filehne illusion should occur, which creates the impression that reference signals are oversized (see sect. 5.1). This means that the magnitude of retinal and reference signals cannot be assessed absolutely, but only relative to one another. Hence, to decide whether a particular condition really creates an increase or decrease in reference signal size, arguments over and above those mentioned in section 5 must be considered. For example,

the occurrence of a Filehne illusion with briefly visible stimuli can only evidence undersized reference signals, not oversized retinal signals, because briefer stimulus presentations are unlikely to increase retinal signals.

**6.5. The vectorial nature of retinal and reference signals.** The present experimental paradigm is based on stimulus motion collinear with self- or ego movements. It therefore uses simple subtraction and addition of retinal and reference signals and of the components within reference signals. The only exception in this respect is the visual component in reference signals, which is likely to show nonlinear interactions with the other components. The assumption that the other additivities in the present model are linear, however, could be viewed as a minimum requirement, at least as long as we do not discover evidence to the contrary. Thus, at present, the model considers calculations concerning these components as basically vectorial.

The somewhat more complex calculations, which ensue when retinal and reference signals are not collinear (i.e., when the stimulus and the eyes do not move collinearly), have recently been described by Mateeff et al. (1991). These could be extended to include 3-D motion in space: since vestibular afferents encode 3-D ego motion, they may induce 3-D (components in) reference signals.

Inferential theory was originally formulated to describe the perception of position of stimuli in space as a function of eye position in the head (see e.g., Helmholtz 1910; Matin et al. 1969; Mittelstaedt 1990). Since velocity relates mathematically to position, it might be possible to extend the present model to include the subjective perception of the position of stimuli in space, and perhaps also the perception of the direction and orientation of stimuli in space. In addition, a model similar to the present one could be developed to describe perception during saccadic eye movements (see Note 7).

**6.6. Other sensory domains.** Formally, the reasoning behind the present model applies to any perceptual system with which object motion can be perceived. For example, consider the tactile system: when our fingertips move across a tactile stimulus (e.g., a rough surface), its shearing velocity across the skin is encoded in an afferent tactile velocity signal ($V_{skin}$). To determine its perceptual significance, a reference signal ($V_{ref}$), encoding finger velocity in space ($V_{fing.s}$), should be created. The stimulus will then be felt to move in space if the difference between $V_{ref}$ and $V_{skin}$ exceeds one JND. Because of Weber's law, the tactile thresholds for stimulus motion with and against finger movements should grow wider apart when finger velocity increases, just as the two opposite thresholds for visually perceived motion grow wider apart with increasing eye velocity (Fig. 2). The shearing velocity of the stimulus across the skin at the midpoint between these thresholds ($V_{skin.PSS}$) would then indicate the magnitude of the tactile reference signal, and its ratio with $V_{fing.s}$ would express its gain.

Although no such experiments have been reported,[11] there is some evidence supportive of such a tactile model: tactile vibrational thresholds are elevated with increased velocity of the skin surface in space (Angel & Malenka 1982; Coquery 1978; 1981; Coquery & Amblard 1973; Dyhre-Poulsen 1978; Paillard et al. 1978; Rauch et al. 1985; see also MacKay 1973). It is interesting to note that recent data from our lab (Bles et al. 1994) show that skin stimulation may cause (illusory) sensations of ego motion. Hence, in the tactile domain there may also exist a self-referential (component in the) reference signal, analogous to the visual one in the present model.

Similar experiments can be performed in the auditory domain by measuring thresholds for hearing the motion of a sound source in space during self- and ego motion. The present model thus provides a theoretical framework for studying the perception of object motion (and stationarity) in any sensory domain, that is, in any perceptual system with a sensory surface that can move in space.

ACKNOWLEDGMENTS

This paper was written with support of the Royal Dutch Air Force (contract A86/KLu/0481). The helpful comments of Bruce Bridgeman of the University of California and Rik Warren of the Amstrong Aerospace Medical Research Laboratory, Ohio, on earlier drafts are gratefully acknowledged.

NOTES

1. In this target article the term "self-motion" denotes movement of parts of the body of the observer; the term "ego motion" denotes whole body movement.

2. Originally, Gibson (1966, pp. 283–84) recognized that vestibular (and somatosensory) afferents may also generate or contribute to percepts of head or ego motion. He proposed that such percepts derive from the covariation of visual and vestibular afferents, their correlation serving as a special kind of invariant. In a later paper, however, Gibson seems to let go of this idea, as he suggests that information about eye and head movements derives from visual kinaesthesis, vestibular information being confirmative only (Gibson 1968; see also Stoffregen & Riccio 1988). In fact, actual research in the tradition of direct perception theory has taken this line of thought and confined itself exclusively to the investigation of optic (or retinal) flow invariants (see, e.g., Andersen 1990; Cutting et al. 1992; Gibson 1979; Koenderink 1990; Koenderink & van Doorn 1987; Owen 1990; Warren 1990; Wolpert 1990).

3. One might assume that this anomalous invariant remains unchanged but that the percept changes. This would, however, imply that there is no one-to-one relation between an invariant and the percept that results when it is picked up, an implication that also seems incompatible with direct perception theory.

4. Some authors have proposed that the efference copy may instead stem from combining these corollaries with afferent oculomotor feedback (see, e.g., Matin 1982), or even from such feedback only (e.g, Skavenski 1972). More recent views, however, tend to agree with the original assumption of Von Holst and Mittelstaedt (see, e.g., Mittelstaedt 1990; Steinbach 1987).

5. Note that, although ego motion is usually treated as an effect, that is, as a perceptual result of a particular action or stimulation, here it is implied that it also causes some other effect.

6. Actually, when the retinal signal is just noticeably larger than the reference signal, the JND might be slightly larger than when the retinal signal is just noticeably smaller. Such a difference would be extremely small, however, and is unlikely to be reliably measurable. In a velocity discrimination experiment, Sekuler (1990) determined the JND between two moving stimulus patterns (since the eyes of the observer were kept stationary – they were fixed on a stationary fixation point which disappeared at the moment of stimulus presentation – this may have corresponded to the JND between a retinal signal and a zero reference signal). She did indeed find that the JNDs for velocity increments and decrements between the two stimuli were equal, but during measurements of only speed discrimination (without, or with degraded, directional motion information), Sekuler observed that the JND for increments and decrements

differed, depending on the speed on the criterion pattern. Note that since we have defined both the retinal and the reference signals in terms of velocity vectors, which include directional information, the latter finding need not be contrary to the present assumption that the JNDs in Equations 1 and 2 may be considered equal.

**7.** During saccadic eye movements made across a stimulus, the threshold for motion of that stimulus increases linearly with saccadic amplitude (Bridgeman et al. 1975). This suggests that the present model could also apply to motion perception during saccades, the dimension defining the size of retinal and reference signals being amplitude instead of velocity.

**8.** Mack and Herman (1978) proposed that during the brief moment when the small stimulus is visible its adjacency to the fixation point moving across the screen may add to the strength of the Filehne illusion. However, De Graaf and Wertheim (1988) showed that this is not the case. Hence this factor is not included in the present analysis.

**9.** There is at least one other report in which the Filehne illusion occurred with a large stimulus pattern that remained visible for a relatively long duration (Mack & Herman 1973). It is interesting that, just as in the present experiment, this pattern contained very high spatial frequency components.

**10.** $V_{ref}$ is the same in Equations 10 and 11, because only one eye movement is made. Thus its associated noise level, the JND, is also the same in these equations. However, a difference between two noise factors (assuming Gaussian noise) may be estimated as the square root of their sum of squares. Hence the noise factor in Equation 12, which represents the noise between two retinal signals, may be estimated as $JND\sqrt{2}$, that is, as slightly larger than the JND between retinal and reference signals.

**11.** J. Lackner and G. P. Amgott-Kwan performed an unpublished experiment in the Department of Psychology at Brandeis University, Boston, similar to the one proposed here. Although it served a different purpose and covered only a small range of stimulus and hand velocities, they did in fact observe a widening of the no-motion range with faster hand movements. Their data also suggest a slight overregistration of hand velocity in the tactile reference signal.

# Open Peer Commentary

*Commentary submitted by the qualified professional readership of this journal will be considered for publication in a later issue as Continuing Commentary on this article. Integrative overviews and syntheses are especially encouraged.*

## Analysis of information for 3-D motion perception: The role of eye movements

George J. Andersen

*andersen@ucracl.ucr.edu; Department of Psychology, University of California at Riverside, Riverside, CA 92521*

The model proposed by Wertheim suggests that the use of extraretinal information (eye movements) is central in the analysis of motion into object- and self-motion components. Although it is obvious that some type of eye movement information must be incorporated in the analysis of retinal motion, several issues are particularly problematic for the proposed model.

***1. Object motion, self-motion and eye movements.*** One of Wertheim's assumptions (also an assumption of other theoretical approaches of motion perception) is that the primary perceptual stimulus needed is retinal motion. Retinal motion includes transformations resulting from motion of objects, motion of the observer, and global transformations of the retinal image resulting from eye movements. Von Kries (1910/1962), in his notes on the writings of Helmholtz, and more recently McConkie and Farber (1979), argued that an analysis of retinal motion containing transformations due to eye movements and object motion would necessarily result in incorrect recovery of depth and 3-D shape from motion parallax. The present approach addresses this problem by comparing a retinal and reference signal. The accuracy of perceived object motion would depend on the accuracy of the reference signal. Any error in this signal would result in misperceived object motion. Although this may not be a severe problem for simple retinal transformations (e.g., horizontal motion of a frontal parallel surface) it would be particularly difficult to successfully recover object motion during complex transformations (e.g., translation in the depth of a slanted surface).

A second limitation is the use of eye movements in analyzing transformations resulting from self-motion. According to the proposed model, an OKN (optokinetic nystagmus) signal is generated during self-motion that is used with vestibular information during head rotation to derive a percept of stationarity in space. One unusual characteristic of the model is that self-motion, in terms of eye movements, is determined by activation of the vestibular system. In contrast, stationarity can be determined by eye movements activated by either an oculomotor mechanism (presumably an OKN signal) or the vestibular system. However, stationarity resulting from vection saturation necessarily implies that the observer perceives self-motion. Thus, self-motion can occur at either output in the model.

This issue aside, the accuracy of perceived stationarity would certainly depend in part, on the accuracy of the OKN signal in terms of magnitude and direction. The study by Brandt et al. (1973), however, suggests that the direction of OKN can be dissociated from the perceived direction of circular vection. Subjects were seated in a circular vection drum and required to fixate a grating pattern in the central visual field which moved in the same direction or the one opposite to the motion of the surrounding drum. They found that induced vection was always determined by the surrounding motion, whereas OKN was always determined by the direction of the central field stimulus. These results indicate that OKN can be dissociated from the perceived direction of induced circular vection (for a discussion of this issue see Andersen 1986).

A third limitation concerns the application of the model to complex transformations of the retinal image during locomotion. For the moment, assume that the proposed model does not have the limitations discussed above. The proposed model is theoretically sound if the retinal transformations involve homogeneous regions of velocity (e.g., transformations resulting from visual stimulation in a circular vection drum with or without the initiation of eye movements). However, it would have considerable difficulty with transformations resulting from observer motion through the environment, especially if the environment contained objects or surfaces undergoing motion independent of the observer's motion. It is the latter complex cases, in my opinion, that led Gibson and other researchers to propose analyses based primarily on visual information.

***2. Alternative models and formal analyses of optic flow.*** Wertheim argues that direct perception assumes that the percepts derive from retinally conveyed information alone. Gibson's analysis of transformations of the optic array, or optic flow, has traditionally been viewed as information not in the retinal image but the world or environment. This unique description of perceptual information has been a source of disagreement and discussion by advocates of both the direct and the inferential/information processing approach. In my opinion, this de-

bate is focused primarily on what a theorist emphasizes as information. Specifically, the question of whether information is in the environment or in the retinal image can be addressed by considering two different classes of models. Let us first examine them without regard to the semantically loaded labels typically associated with them. Class I models assume that the information is present in the transforming pattern of light and thus in the environment. Although it is true that information is first available to the visual system at the retina, and thus is subject to global transformations resulting from eye movements, this point is irrelevant, because the information, according to this class of models, is in the pattern of light. Thus, although this approach must acknowledge the existence of eye movements, it does not view the retinal image as the critical level of analysis and does not consider it to be the appropriate description of information.

According to class II models, on the other hand, the first stage at which information is available to the visual system is the retinal image. Hence this must be the first level at which information can be analyzed. This class of models (e.g., Koenderink & van Doorn 1976; Longuet-Higgins & Prazdny 1980) typically involves formal mathematical analyses of the transforming retinal image; but often, as a first stage of analysis, the transformation of the retinal image is decoupled into translation and rotation components. The rotational components, which would be the result of eye movements, are not analyzed. Instead, the remaining analysis is focused on the translation components.

It is important to note that although class I and class II models use information quite differently, they can both be considered as emphasizing the same information – the transformations of the optic array separate from the transformations due to eye movements. It is this point that distinguishes the model proposed in the target article from analyses concerned with optic flow. Wertheim's model suggests that the analysis must have as a central component the incorporation of extraretinal information (eye movements); the class I and II models consider eye movements to be of secondary importance.

In summary, the issues discussed above are particularly problematic for analyses based on retinal motion. One should not conclude, however, that eye movements are not important in perceptual processing. At some level of analysis the visual system probably has to use information regarding eye movements (whether it be an efference copy signal or a mathematical analysis involving the decomposition of the transformation into translational and rotational components) in analyzing transformations of the visual scene. Although information for eye movements from the vestibular and extraretinal signals certainly are important in our perception of both object and observer motion in the 3-D world, it seems likely, given the complexity of these transformations, that the analysis is primarily visual in nature.

## A theory of the perceptual stability of the visual world rather than of motion perception

Wolfgang Becker[a] and Thomas Mergner[b]

*[a]wolfgang.becker@medizin.uni-ulm.de and [b]mergner@sun1.ruf.uni-freiburg.de; [a]Sektion Neurophysiologie, Universität Ulm, D-89069 Ulm, Germany and [b]Abteilung Neurologie, Albert-Ludwigs-Universität, D-79104 Freiburg, Germany*

In his target article, Wertheim proposes a unifying theory of motion perception, which he feels reconciles previous theories that have variously stressed the role of extraretinal signals and of cues contained in the optic flow. According to Wertheim, the novel thing about this theory, illustrated by gray arrows in his Figure 1, is a link between (1) the perception of self-motion and (2) visuo-oculomotor mechanisms of motion perception. From a systems engineering point of view this is an almost trivial postulate: given that the head can move in space, a summation of a visual object's retinal slip with an extraretinal signal of eye-in-head velocity (efference copy) can only provide an indication of object versus head rotation. To arrive at a notion of object-in-space rotation we obviously need to know head movement in space. Such knowledge is forwarded by the vestibular system cooperating with a variety of other sensory channels (neck proprioceptive, optokinetic, arthrokinetic) and, last but not least, cognitive mechanisms (Mergner & Becker 1990). Although this cooperation results in perfect registration of self-motion, and hence of object motion, in most everyday situations, it can be misled in the laboratory when its fundamental premise – long-term stability of the visual surround as a representation of the "world" – is violated; the resulting illusions that afflict the perception of head motion in space also distort the perception of object motion, in a manner that can be predicted from its dependence on head-in-space representation (Mergner et al. 1992). Thus, we cannot but agree with Wertheim's basic postulate, which is a logical consequence of the head's mobility. However, as we shall discuss below, this is not to say that we agree with all of Wertheim's views.

Wertheim's experimental and conceptual approach mixes up two related yet not identical issues, (1) the perceptual stability of the world during eye movements and (2) the perception and quantification of object motion in space. It is debatable whether his experimental approach does not bias the study of (1), and it appears inappropriate for (2). In fact, except for saccades (and for some unphysiological conditions such as spontaneous nystagmus), people do not make spontaneous eye movements across a visual scene. Such movements occur only when they track a stimulus ("visual object") moving relative to the background ("scene"); in doing so, their attention is normally focused on the *object* but not the scene.

In asking his subjects (Ss) to report movements of the background (the stimulus in his terminology) while pursuing a moving fixation point, Wertheim must have dissociated their orientations of gaze and of visual attention. However, the perceptions of self-motion, object motion, and scene motion may depend strongly on where visual attention is directed. Under a variety of experimental conditions we have observed that Ss who concentrate on delivering concurrent indications of self-motion and object motion in space during stimulation will report that the scene was stationary when questioned on their perception of scene motion *after* termination of the stimulus; in contrast, when instructed to give a *concurrent* indication of the scene's behaviour they will perceive it as moving in space and will change their judgments of object and ego motion so as to restore compatibility with both the scene-in-space and object-versus-scene perceptions (Mergner & Becker 1990). Thus, if we assume, for lack of relevant details on experimental procedures, that Wertheim's Ss were instructed to watch background stability during their eye movements, the thresholds he reports may underestimate those occurring in "natural" situations.

By a similar argument, in order to investigate and quantify the perception of object motion we must, in the first place, focus the perceptual inquiry on actual objects rather than on the scene. Otherwise, given (1) that registration of self-motion is an indispensable prerequisite for judging object motion in space and (2) that scene motion conflicts with the working hypothesis of the movement-in-space channel (stationarity of the world), we risk obtaining a distorted picture of object behaviour. Clearly, we do not doubt Wertheim's experimental results, which are valuable and consistent as such, but we insist that they be correctly labeled: they concern the perceived motion of a visual stimulus the eye is *not* tracking because it is engaged in the pursuit of another stimulus; unless positive evidence is given, they are not applicable to the movement of the stimulus that is being pursued.

Wertheim realizes that the perception of self-motion depends on a visual-vestibular interaction. Unfortunately, his description

of the visual component of this interaction (the "gating mechanism") remains rather vague; except for some qualitative indications, we are given no estimates for its time constant, its gain, its dependence on stimulus characteristics, and the mode of interaction with the vestibular signal.

There are obvious reasons to ask for such details. We take it from Wertheim's description that with presentation times of more than 300 msec the gating mechanism begins to contribute significantly to the reference signal suggesting a time constant of less than 1 sec. Furthermore, its contribution is said to establish a total gain of the reference channel of about unity; given an efference copy of $0.7–0.8 \cdot V_{eye}$, one would infer a gain of 0.2–0.3 for the gating mechanism. On the other hand, the building up of activity in the optokinetic channel during sustained rotation of a scene, to a level where it can no longer be distinguished from the retinal signal, is viewed as the cause for the emergent perception of scene stationarity and the concomitant circular vection (CV); for this interpretation to hold, the optokinetic contribution ought to have a gain of about unity and a time constant considerably in excess of 1 sec.

Wertheim claims that the optokinetic contribution to his reference signal explains center-surround induced motion (Duncker illusion; Duncker 1929) and dissolves the distinction between "subject-relative" and "object-relative" motion. If we take this claim and Figure 1 at face value, the illusory perception of object motion in space that is evoked by an optokinetic movement behind the stationary object should disappear if the optokinetic contribution to the reference signal is balanced by an efference copy of opposite direction.

However, (unpublished) results from our laboratory strongly suggest that this is not the case. We presented Ss with a sinusoidal rotation of an optokinetic pattern, and with an (initially) stationary light spot in front of the moving pattern which they were to fixate. Invariably this stimulation led Ss to experience an object-in-space rotation. Ss were then to move the object, by means of a joy stick, in the opposite direction while a constant scene-to-object motion was maintained (by an appropriate coupling of the object and scene actuators), until the object would appear stationary. None of our Ss (N = 6) was able to perceptually stabilize the object by this procedure although, using a similar method, all could nullify vestibular and neck-proprioceptive illusions of object motion (cf. Mergner et al. 1992). Ss continuously increased the physical spot excursions up to a manifold of the retinal stimulus without major effects on perceived object motion in space. It is interesting that during the later course of such trials a strong CV and motion sickness would develop.

These observations suggest that visual capture and object-relative cues constitute a powerful and separate system of object-motion perception that is (1) not accessible to oculomotor efferent signals and may (2) contradict oculomotor and vestibular cues or (3) change their interpretation. To illustrate point (3) we note that naive Ss who are presented with an object and a scene that rotate together in a sinusoidal manner will veridically perceive both object and scene as rotating in space and will experience no CV. In contrast, when the object is moving at a different velocity or when it is fixed in space, a CV develops, but only secondary to the (illusory) percept of object-in-space motion evoked by the relative motion between object and scene (Mergner & Becker 1992, p. 231).

A further argument against Wertheim's explanation of the Duncker illusion comes from the observation that center-surround motion also induces marked percepts of object-in-space motion at frequencies (e.g., 1 Hz) that would seem to exceed the cut-off frequency of his gating mechanism (again, one would like to know its characteristics).

To explain a variety of phenomena Wertheim invokes his hypothesis of a dead zone ("JND") affecting the reference signal and growing in proportion to its magnitude. It is unclear, however, whether he also envisages a JND for his retinal signal – if Weber's law is invoked for the reference signal why not also for the retinal one?

In the context of relative motion perception it becomes unclear where in the topology of Figure 1 we are to assume the JND. By estimating the JND for the perception of velocity differences to equal about $\sqrt{2}$-times the JND for perception of movement in space, Wertheim seems to suggest that the signals $V_{ref}$-$V_{ret1}$ and $V_{ref}$-$V_{ret2}$ are first subtracted from each other (the difference then would have $\sqrt{2}$-times the noise of $V_{ret}$) before applying the dead zone, which would therefore be located after his comparator mechanism. The preceding paragraphs, however, suggest that the dead zone operates on the $V_{ref}$-input to the comparator, sparing the retinal input. Also, on the basis of Equation 12 one would conclude the constant in Equation 13 to be of the order of $JND \cdot \sqrt{2}/V_{ret}$ or 0.14 (Fig. 2 suggests a value of $0.1 \cdot V_{ret}$ for JND); however, the value taken from Figure 5 is of the order of 0.5.

A final comment concerns the vestibular contribution to the hypothesized reference signal. It is surprising that the vestibular registration of head velocity in space should have a gain of more than unity (cf. Equation 19) at a frequency as low as 0.05 Hz. We and others have observed clearly lower gain values (cf. Mergner et al. 1991, p. 393; 1992, p. 659). We suspect that the presentation of the visual scene at regular intervals and always during the same phase of the vestibular stimulus may have helped to establish almost the same notion of head-in-space velocity as its permanent visibility would have evoked.

Finally, still in the context of visual-vestibular interaction, it is surprising that some well-considered theories of visual-vestibular interaction such as the conflict model of Zacharias and Young (1981) or descriptions of basic principles of perceptual stability of the visual world (Bischof 1974) are not even mentioned by Wertheim.

ACKNOWLEDGMENT

This work was supported by Deutsche Forschungsgemeinschaft, Sonderforschungs bereich 325, No. DFG SFB 325.

## Frame and metrics for the reference signal

Victor I. Belopolsky

*vbelop@ipras.msk.su; Institute of Psychology, Russian Academy of Science, Moscow 129366, Russia*

In the attempt to reinforce the behavioral validity of the traditional inferential theory of perception, Wertheim has reconsidered the nature of the extraretinal signal. The reference signal, instead of the extraretinal one, is assumed to encode information about eye movement relative to external space together with visual spatiotemporal and vestibular motion information. Although this novel position appears reasonable and is capable of explaining some contradictory experimental findings, it faces serious problems, both old and new.

The basic assumption of the traditional inferential theory is the egocentric, mostly oculocentric, representation of the external world; that is, to be compatible, the metrics of both the retinal and the hypothesized extraretinal signals must be defined in visual angles. Wertheim accepts that "we see a stable world during eye movements because retinal and extraretinal signals are equal: the velocity of the image of the world across the retinae equals the velocity of the eyes" (sect. 1, para. 1). However, even under conditions in which the eye rotates around a fixed center, the two signals do not operate in the same metric.

The above claim is derived from the fact that the eye's rotation center does not coincide with the eye's nodal point. So when the eye rotates a certain angular distance, the corresponding shift in the retinal image of a stationary object will actually be a fraction

of that angle. In terms of visual directions, the ratio could be up to 1:2. The same ratio will hold for foveal velocity during smooth pursuit of a moving target and for the target's retinal velocity relative to the stationary eye. Jung (1972), who first advanced this line of argument, used it to explain the Aubert-Fliesch phenomenon. This can also be considered the source of the Filehne illusion. As a consequence, we must assume that recalibration of the extraretinal signal is a precondition for visual stability during eye movements.

The discrepancy between the center of rotation and the nodal point of the eye high-lights the role of another factor that is usually underestimated by the inferential theory, the target's distance. Two targets at different distances which were previously on one visual axis lose this identity of their visual directions after an eye movement (Howard 1982, p. 278, Fig. 7.1). This means that retinal velocity during eye movement is a function of the target's distance. This factor could play an important role in contradictory demonstrations of the Filehne illusion: it is much easier to experience the illusion when the moving target and the background are located in different planes than when they are in the same plane.

The situation becomes more complicated when someone tries to define the metrics of eye movement relative to external space. Wertheim introduces the vector signal of head movement in coordinates of 3-D "Newtonian" space (sect. 5.3, para. 1). It is not difficult to show that the metrics of head rotation differ from the oculomotor and retinal vector metrics even when a subject is sitting upright and turning his head and eyes around a vertical axis. If we also take into account translational head movements (because the head has 6 degrees of freedom), this raises the question of the metrics of the reference signal. We must agree that Wertheim's Equation 9 is accurate only if the dimensionalities of its terms are the same. However, $V_{est.s}$ and $V_{ref}$ are encoded in the metrics of the exocentric coordinate system whereas $V_{ret}$ is encoded in oculocentric terms. Moreover, in general, there is no universal transfer rule for transforming one coordinate system into another: the rule depends on the relative positions and movements of the observer and the external objects. This point was in fact crucial for Gibson's rejection of geometrical optics in favor of his ecological position (Gibson 1979).

The acceptance of a visual component of the reference signal reflects the proved significance of visual feedforward in visual stability processing in addition to visual feedback (e.g., Belopolsky 1978; MacKay 1973). Unfortunately, this claim cannot be formalized in terms of vector algebra (see sect. 6.5, para. 1) and leads to the redundant duplication (or triplication?) of visual pathways in the proposed model in the target article.

According to the proposed model, the reference signal has no sensory correlate; it is used only to cancel, completely or partially, the retinal signal. As a result, its role is purely visual rearrangement. The sensations of self-movement or self-stability come through a parallel branch of the information processing system, although the reference signal itself contains all necessary data. The model does not provide any special mechanism for the coordination of visual and ego movement in space. For example, the model is rather efficient in explaining the time course of background motion perception (assuming variability in $V_{ref}$ gain) but fails to explain movement of eye/head/body egocenters during the circular vection illusion. It is worth noting that direct perception theory manages this problem by considering the observer's body parts as the context of the optical array (Gibson 1979).

The most challenging problem for both theories remains the nature of eye or, more appropriately, gaze positional sense, by which I mean the human's ability to hold, reorient, and locate the position of attentional focus in space. Direct perception theory does not indicate the body's landmark connected with the gaze direction (this cannot be derived simply as the center of the optical array). On the other hand, this sense cannot be identified with an extraretinal or a proposed reference signal. Two examples will illustrate this idea. First, when the stabilized retinal image subtends a large visual angle (> 40 deg of arc), eye movements do not change its apparent spatial location, as occurs with a smaller image. Especially interesting is the fact that, in both cases, subjects experience their gaze as moving in space (Belopolsky 1985; Zinchenko & Vergiles 1972). Second, when voluntary eye movements are made as the subject examines meaningless texture patterns through an artificially reduced (up to 3–5 deg of arc) central visual field, the visual world is perceived as movable relative to the stationary gaze (Belopolsky 1978).

In summary, Wertheim's target article provides a subtle analysis of the direct versus inferential perception controversy. However, the attempt to resolve the controversy on the basis of inferential theory has made it too complicated and flexible to be an effective tool for predicting perceptual experience in certain conditions.

## Biological perception of self-motion

Ronald G. Boothe

*boothe@rmy.emory.edu; Departments of Psychology and Ophthalmology and the Yerkes Regional Primate Research Center, Emory University, Atlanta, GA 30322*

The schematic in the upper right hand corner of Figure 1 illustrates a view of a retinal image in which an element moves from point X to point Y. There are three ways this retinal image movement could have come about. The first, illustrated in the upper left of Figure 1, is that the head and eye remained steady and that an object in the environment moved from A to B. The second, illustrated in the lower left, is that the object and the eye remained steady, but that the head moved from A to B. The third, illustrated in the lower right, is that the object and the head remained steady, but the eye rotated in the socket from position A to B. Thus, the perceptual system of a biological organism that has such an eye is confronted with an ambiguity. How does it decide what this movement from X to Y an the retinal image means in terms of what is happening in the environment?

There are two general theories about how this perceptual problem might be solved. Direct perception theories state that there are lots of other elements moving in the retinal image at the same time as this particular one we are focusing our attention on, and that the information potentially available from all these moving elements (the optic flow) is sufficient to disentangle the various environmental sources that gave rise to these movements in the retinal image. Inferential theories state that the brain uses extraretinal sources of information to estimate head movement (derived from vestibular system input) and eye movement (derived from copies of motor efference). When these extraretinal sources are subtracted from the retinal motion, the remainder can be interpreted as movement of objects in the environment.

Wertheim points out that both these theories have problems accounting for some perceptual illusions that involve vection. He offers an alternative model that attempts to solve these problems by introducing the concept of a "reference signal." His alternative model relies on input from the optokinetic system to supplement the extraretinal information used by the inferential theories. He cites empirical data obtained in laboratory settings that support the predictions of his model. There are problems in generalizing Wertheim's model to more general ecological conditions, however, because the optokinetic system does not provide the requisite information except under specialized conditions.

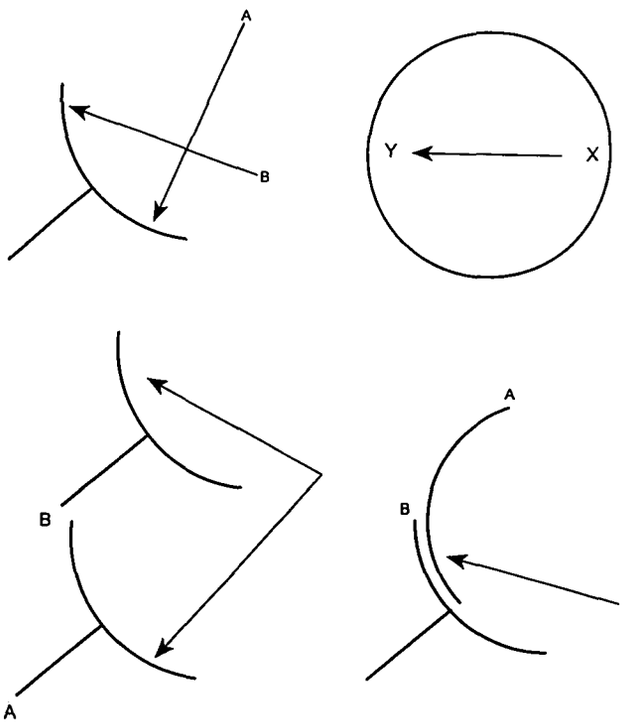


Figure 1 (Boothe). Schematic illustration of three environmental sources of retinal motion.

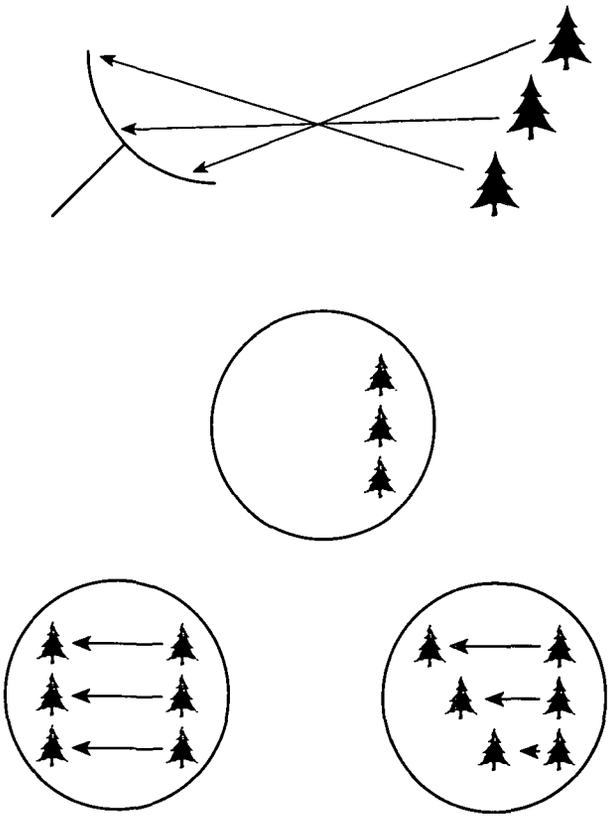

Figure 2 (Boothe). Schematic illustration of optokinetic stimulation produced in the ordinary environment when the eye is rotated, but not when the head is translated.

I illustrate the reason for this in Figure 2. The schematic at the top of the figure shows an eye that is viewing three large trees. Low spatial frequencies from these trees form images in the retina, illustrated in the schematic in the middle of the figure. The left bottom schematic illustrates what happens to the retinal image when the trees are stationary and the eye or head is rotated. This retinal motion provides the input to the optokinetic system that Wertheim uses in his model to explain the vection illusion. However, the right bottom schematic illustrates what happens to the retinal image when the trees and the eyes are stationary and the head translates, as would happen in the example Wertheim uses of a train engineer. There are three different velocities generated here, one for each tree. Which of these velocities is Wertheim going to plug into the optokinetic system to help generate $V_{ref}$? It is obvious that the information needed by Wertheim's model is not going to be provided by the optokinetic system under general ecological conditions. This only works for specialized conditions in the laboratory, such as when the movements feeding into the optokinetic system are all generated by objects that happen to be at the same distance from the observer.

Mother Nature has no interest in conforming to our theories of perception. She is only concerned with picking up information that has been demonstrated to be important for survival over evolutionary time. She is therefore unlikely to care whether the information available to the brain is of retinal or extraretinal origin (a distinction that is important to direct perception theorists). On the other hand, she is not likely to be very interested in trying to figure out the meaning of information that would only be available in artificial laboratory environments and not in the environment in which evolution took place (conditions that are sometimes used to test predictions of inferential theories). When presented with a complicated set of retinal and extraretinal stimulations in a laboratory environment, a biological system is likely to interpret the information in terms of what external conditions would have to have been present in the environment in which evolution took place in order for that pattern of stimulation to have occurred. Given this rationale, there is no reason we should be surprised by the fact that when seated in a drum in a laboratory in which the retinal stimulation is as in the lower left of Figure 2, and the efference copy in the brain indicates that the eyes are still, and the vestibular system input indicates that there is no acceleration, and this pattern of stimulation continues for some time, vection occurs. The only condition that could have occurred in the environment in which our species evolved that would give rise to this sustained pattern of stimulation would be if the world were stationary and we were spinning around at a constant rate.

## Extending reference signal theory to rapid movements

Bruce Bridgeman[a] and Jean Blouin[b]

*[a]bruceb@cats.ucsc.edu and [b]laboscm@frmrs11.bitnet; [a]Department of Psychology, University of California at Santa Cruz, Santa Cruz, CA 95064 and [b]Laboratoire de Controle Sensorimoteur CNRS, Département de Neuroscience, Université de Provence, Marseille 13397, France*

Wertheim's view of motion perception brings together a wealth of observations and theory that have accumulated over several decades, synthesizing them into a comprehensive model. The present model is aimed at smooth self-motion and continuous eye movements. We would like to begin the process of extending the model to ballistic bodily movements and saccadic eye movements as well, and on the sensory side to address perception of abrupt jumps of visual targets.

Reference signal theory predicts that the relative contribution

of retinal and extraretinal signals for motion perception is a function of the presence or absence of relative motion cues during movement. In a normal structured visual environment, retinal signals provide the dominant information processed by the brain for perceiving the motion of a visual object. In a dark room (i.e., when no relative motion cues are provided), however, extraretinal signals necessarily dominate. In the latter condition, Wertheim assumes that the reference signal for processing space constancy emerges from a combination of signals from different sources capable of providing an estimate of how the eyes move in space (e.g., efference copy of the oculomotor outflow, vestibular signals).

Problems of information pickup around saccadic eye movements are fundamentally different from those around smooth movements because vision and velocity estimation continue during smooth movements while both of these are interrupted or suppressed during saccades. In an accompanying BBS target article, Bridgeman, Van der Heijden, and Velichkovsky (this issue) propose that quantitative subtraction of retinal and extraretinal signals before and after a saccadic eye movement is not necessary – rather, all sources of spatial information are simply used during a new fixation to provide perceptual and motor orientation. Nevertheless, we suggest that a richness of information sources exists for calibrating perception and visually guided behavior surrounding saccadic and other ballistic movements.

Recently, we have examined empirically whether signals that do not provide information about how the eyes move in space can nevertheless inform the perception of spatial position. More specifically, in a dark room we tested whether saccadic suppression of displacement is modified in combining eye and arm movements toward a small target shifted during a saccade (Blouin et al., submitted). The arm movements were rapid and visually open-loop. After they jabbed a pointer at the target's position, subjects judged whether a small visual target had been displaced during the saccade. We found that the perceptual threshold of target displacements increased when subjects responded with combined eye and arm movements compared to a situation with only saccadic eye movements. Signals arising from the sensorimotor arm system hence reached the spatial processing system and increased perceptual stability of visual world.

In this experiment, subjects corrected most of their arm trajectory toward the displaced targets. Error messages about the programmed movements (i.e., eye and arm) therefore reached the eye and arm motor control systems without being perceived. Nevertheless, they served to amend the motor responses rapidly in relation to the new target position through short latency secondary eye movements and online corrections of the arm trajectory. Following the movements, the null error signal issued from the sensorimotor arm system presumably overlapped the postsaccadic retino-oculomotor error signal at the conscious level and a stabilized environment was perceived. Terminal hand position error was then interpreted as being within the range of normal endpoint dispersion associated with the production of rapid arm movements rather than a change of the environment. All these processes involve evaluation of retinal and extraretinal signals after the completion of the saccade and the arm movement, rather than pre- and postmovement comparisons.

These results are in agreement with Wertheim's model and provide further evidence that space perception is not restricted to the visual modality, especially in the absence of relative motion cues. Rather, the results argue for a supramodal spatial processing in which multiple nonvisual mechanisms are closely involved, including those that do not provide information about the position of the eyes in space. Specifically, this experiment reveals that the processing and updating of spatial information seem not only to take into account movements that change position of the eyes in space, such as eye, head, or locomotor movements (as emphasized in Wertheim's target article), but also spatially goal-directed arm movements that do not affect eye position in space.

An additional result speaks to the influence of extraretinal signals on perception. Judgments that the target had not jumped were most frequent when the target indeed failed to jump, that is, perception of target displacement during the saccade was veridical except for the increased displacement threshold. This was true only when the eye and arm were moved simultaneously, even if the arm movement was mechanically interrupted just as it began. Thus an efference copy of the corrected arm movement must have informed the visual system about the extent of the target displacement. When only the eye moved, however, judgments that the target had not jumped were most frequent when the target actually jumped slightly in the direction opposite the saccade. This result shows that the arm movement did not serve merely to increase displacement thresholds nonspecifically, but altered the perception of position. The underconstancy is consistent with the finding of Bridgeman and Stark (1991) that the combined gains of efference copy and oculomotor proprioception are less than one in an unstructured field.

The difference between our arm movement condition and our no arm movement condition is analogous to Wertheim's measure of the midpoint of the subthreshold range and again shows that arm movement efference copy and proprioception contribute to perceptual judgment. The details of the saccadic and the smooth movement system are different, but the general principle of combining multiple information sources seems to be the same.

## Ego- and object-motion perception: Where does it take place?

U. Büttner and A. Straube

*u7x31aa@sunmail.lrz-muenchen.dbp.de; Department of Neurology, Ludwig-Maximilians-University, Klinikum Grosshadern, 81377 Munich 70, Germany*

Vestibular stimulation leads to the sensation of ego (self)-motion, whereas visual stimulation can lead to both object- and ego-motion perception. Thus a multimodal visual and vestibular interaction should be the neuronal basis for ego-motion perception. There is a vast literature on neuronal activity related to visual-vestibular interaction that includes many articles referring to object and ego motion. In this commentary it will be argued that despite the vast and "sufficiently complex" literature on this topic, there is still a lack of satisfactory evidence for a signal that might be related to the sensation of ego motion. In particular, the question of a primarily cortical or subcortical interaction is still under discussion.

The fact that neurons that respond to both visual and vestibular stimulation have been recorded in the vestibular nuclei in the brainstem (Waespe & Henn 1987) has been taken as an indication that a reference signal for visually-induced ego motion is present in the brain. However, there are several points of evidence against such an assumption. In particular, the following factors must be considered:

1. *Eye movements:* Vestibular and visual stimuli that induce ego motion also lead to eye movements, that is, nystagmus. In many studies, a distinction cannot be made whether activity might be related to the eye or to a reference signal for ego-motion velocity. Visually induced ego motion (vection) of considerable magnitude also occurs during the suppression of nystagmus (Bötzel & Grüsser 1982). Under certain conditions, slow phase nystagmus velocity and vection, which are normally in opposite directions, can be dissociated (Brandt et al. 1974). No studies specifically addressing these points have been performed and hence results indicative of an ego-motion reference signal have not been established (Buettner & Büttner 1979). It

should be stressed that in subcortical structures highly specific signals can also be encoded; for example, Purkinje cells (Pcs) in the oculomotor vermis of the cerebellum have been found to encode target velocity during head, eye, and target motion (Suzuki & Keller 1988).

2. *Location:* Neuronal correlates of a reference signal for ego motion should be expected specifically in cortical structures rather than in the brainstem. In recent years, many vestibular cortical areas have been delineated (Guldin et al. 1992): the parieto-insular vestibular cortex (PIVC), area 3aV, area 2v, and parietotemporal association area T3. In most of these areas neuronal evidence for visual-vestibular interaction has also been found (Büttner & Buettner 1978, Büttner & Henn 1981; Grüsser et al. 1990a; 1990b). For most of these areas, however, a specific functional role has not been established, particularly in relation to a reference signal for ego motion. They are probably not primarily involved in the generation of vestibular system-related eye movements. Some might be involved in the vestibular and visual control of posture, and others might encode motion; however, there are no results indicating a specific relation to the encoding of vection.

Another important but unanswered question concerns where the convergence of vestibular and visual signals relevant for vection takes place. It has been postulated that visual-vestibular signals from the vestibular nuclei are transferred to cortical areas via thalamic nuclei. However, only very sparse projections from the vestibular nuclei to the thalamus have been found (Lang et al. 1979). At present, it appears likely that visual signals relevant to vection originate in cortical areas and converge at the cortical level with vestibular signals (Straube & Brandt 1987). As one would expect from these concepts, visually induced ego-motion sensation is lost in patients with parieto-temporal lesions when visual stimulation is restricted to the ipsilateral visual cortex (Straube & Brandt 1987).

As mentioned above, it is not quite clear by which routes the vestibular signals enter the cortex. Descending pathways from cortical vestibular areas to the vestibular nuclei have been recently demonstrated (Guldin et al. 1992). These pathways might be involved in the cortical control of posture, but their relation to vection is unknown.

3. *Temporal and physical aspects:* To demonstrate a neuronal signal encoding vection, it is also important to consider temporal aspects of vection. Vestibular neurons carry a signal related to the velocity-storage mechanism during optokinetic stimulation (Cohen et al. 1977; Waespe & Henn 1987). Such signals have not been encountered in cortical vestibular areas (Büttner & Henn 1981; Grüsser et al. 1990a; 1990b). Velocity-storage signals can be best related to eye movement parameters rather than to vection. They usually have a buildup of several seconds, whereas under certain stimulus conditions ego-motion sensation can have latencies of less than 1 second (Straube et al. 1990). It has also been shown that visually induced ego-motion perception depends on stereo-optical perceived depth (Wist et al. 1975). Such stereo-optical visual responses are found only in the visual cortex.

Recently, attempts have been made to train monkeys to signal a distinction between visually induced ego and object motion (Mergner, personal communication). With these experiments it appears feasible to record neuronal activity which can be related either to ego or object motion. Thus, at present, the challenge is to provide neurophysiological evidence for the theories and concepts as shown in Wertheim's Figure 7. Until then, the present neurophysiological data cannot be taken to support of any of the theories related to vection.

## Sensor fusion in motion perception

David Coombs

*david.coombs@nist.gov; Robot Systems Division, National Institute of Standards and Technology, Gaithersburg, MD 20899*

A goal of robotics is to discover principles that enable systems to behave robustly in complex domains, and animals offer rich examples of such systems to study. Wertheim's theory elegantly unifies empirical data and existing theories. It is satisfying to see the theory grounded in the neurophysiological structures that have so far been implicated, since implementation feasibility is important. Nevertheless, speculation on additional cues to motion perception might be reasonable. Vision is the primary cue to perceiving motion of objects, but there are several possible clues to ego motion (motor efference from eye, neck, and body, vestibular and otolith afference, oculomotor kinesthesis, visual afference, and expectations) and these may influence the perception of motion in general if they help inform ego-motion estimation. Ideally, a robot would perceive motion and behave appropriately. The system would attempt to assimilate sensory data from several modalities to estimate the motion of the creature itself, and of nearby objects. Human motion perception can be considered from this viewpoint.

Suppose that one drives a car over a bumpy road. The largest visual flows indicate forward translation, yet the vestibular and otolith systems do not. They indicate accelerations of the head with respect to a coordinate system (CS) that is a low-pass filtered (smoothed) version of the (bouncing) car CS, since the subject has reached a steady velocity and is no longer accelerating. There are three possibilities. The vestibular and otolith signals fully describe the subject's motion and a large segment of the world is moving rapidly past the subject; or there is a steady state component of the subject's motion to which the vestibular and otolith organs no longer respond; or both the subject and the scene are moving with respect to the world. How does the system choose a combination of ego motion and object motion that accounts for the observations?

Two obvious policies are to take the most likely interpretation based on experience, and to make the most conservative judgment with respect to the creature's safety. The creature must presume that it can influence its motion, so it might as well attribute perceived coherent large-field motion to ego motion and behave accordingly. If it cannot control the relative motion, its response will be irrelevant, but if it can, it might avoid colliding with trees, for instance. It might be argued, then, that the safe interpretation is to attribute the recent history of sensory cues to ego motion.

A crucial element of Wertheim's model is the incorporation of the sensory data histories and the resulting interactions. The temporal characteristics of each cue might be considered. What determines the spatiotemporal properties of the optokinetic signal that contributes to perception of ego motion? Certainly the range of image flows that can be used is limited by the range that can be perceived, but is use of the data further limited? It has been noted (Howard 1982) that the sense of vection approximately follows the time course of the decay of the vestibular system's response to constant rotation of the body. These spatiotemporal response characteristics of complementary signals such as vestibular data might be natural for limiting the use of visual signals to charge up the eye ego-motion reference signal. Suppose that image flows corresponding to ego motions below the sensitivity of the vestibular system contribute to the reference signal immediately and faster flows contribute to the reference signal only after the vestibular system could be expected to fall silent. Then vection would be induced immediately at accelerations below the sensitivity of the vestibular system. Vection would result later for larger accelerations only after the history of conflicting vestibular data is sufficiently old that the subject may have been slowly accelerated, undetected

by the vestibular system. Thus, motion would be attributed to both the subject and the scene until enough time has passed so that the subject's motion alone could account for the visual motion.

Consider the extinction of vection in an ordinary situation. Suppose one is sitting in a train and an adjacent train begins to move, inducing vection. When one looks up and sees that another train (and not one's own) is moving, the vection is extinguished. What factors could have extinguished vection? It could not have been vestibular sensations resulting from looking up, or they would also suppress the veridical perception of ego motion when one's train does begin to move. Again, there are three cases. The other train moves, one's own train moves, and both trains coincidentally begin to move at the same time. When one looks out the window, one can presumably determine whether either or both of the trains are moving against the ground and trees. (It might be interesting to know whether there is a difference in the reaction times to extinguish vection and to confirm veridical ego motion.) This example suggests that some sensory fusion occurs at a high level even if it is not necessarily required for behavioral responses. It further confirms that in ordinary behavior, the creature may actively seek out additional information to assess the situation. There is a related question I would like to consider. When a person experiences vection in a parked car, a common reaction seems to be stomping on the brake pedal before looking up to determine whether the car is in fact rolling. Does a train passenger ever stab at an imaginary brake pedal at the onset of vection? If not, then this suggests that low-level "reflexive" reactions to ego-motion perception are influenced by high-level contextual expectations. It may be impossible to decouple models of motion perception from models of behavior and expectation; it seems likely that higher-level models will be necessary to predict many observed behaviors.

## Direct perception theory needs to include computational reasoning, not extraretinal information

Niels da Vitoria Lobo

*niels@bohr.cs.ucf.edu; Department of Computer Science, University of Central Florida, Orlando, FL 32816*

I will argue two points. The first is that neither direct perception theory nor Wertheim's theory needs extraretinal (oculomotor) information to achieve the explanatory power desired. Second, both theories need a stronger role for inferential reasoning, along the lines of the target theory's comparator mechanism. [See also Ullman: "Against Direct Perception" *BBS* 3(3) 1980.]

Examine Wertheim's functional model (outlined in his Fig. 1). The main disagreement I raise concerns how it is that eye velocity becomes known (to generate the reference signal); I will argue that retinal motion information suffices for the task. If one considers the model in Figure 1 minus the efference copy pathway from the oculomotor mechanism, then, with one exception, all of the claimed new explanatory power still holds. The exception appears to be the argument presented at the beginning of section 5.2 regarding a horizontal line through Figure 3. It would seem that this argument about the results in Figure 3 necessitates oculomotor input. This is deceptive. Examine carefully the conditions under which the experiments pertaining to Figure 3 were performed. They (and others with similar findings) took place in front of screens with visible borders. Hence, when nonzero eye velocity is used, border effects allow the computation of eye velocity from *purely* retinal image flow information. The claim that retinal velocity (of the stimulus) is kept constant (when considering the horizontal line through Fig. 3) is accordingly true only for a portion of the retinal array. Invariants present in retinal flow do not stay the same when moving across the horizontal line of Figure 3.

There is computational evidence that retinal flow does suffice for obtaining ego motion of the eyes (which can generate the reference signal needed for the comparator mechanism). Novel methods for computing ego motion and alternatives to the target theory's comparator mechanism for detecting (independent) object motion from purely retinal flow have been presented recently (da Vitoria Lobo 1992; da Vitoria Lobo & Tsotsos 1991). [See also Tsotsos: "Analyzing Vision at the Complexity Level" *BBS* 13(3) 1990.] The key idea behind the alternatives to the comparator mechanism can be expressed as follow: the arguments are made at the level of description of three-dimensional motion. For a purely translating eye, any retinal flow (at a point in the image) with a component perpendicular to a radial line from the FOE (focus of expansion) must be due to independent object motion. For an eye that translates and rotates simultaneously, a three-point computation that cancels rotation suffices to detect independent motion. The computation of ego motion by this approach of cancelling rotation is similarly simple and efficient. The only departure from direct perception theory in that work is that the notion of computation (and hence representation) is needed, something that die-hard direct perception proponents appear to deny.

The second point of this commentary is related to the above. The target article seems to be right in suggesting that the stationarity tendency of large stimuli arises from their large optokinetic potential. The problem with relying on this as the complete story is that other phenomena will not be explained. Consider, for example, the explanation offered for the moving train experience: the window containing the moving train acts as the optokinetic stimulus for experiencing ego motion. First, one can observe that one does not typically experience ego motion every time one sits in front of a movie screen (even though its field of view is at least as large as that of a moving train). Second, one can experience ego motion when peering out the small window of a stationary docked airplane if the adjacent plane begins to move. Both these observations seem to suggest the use of expectation as a factor in deciding the final interpretation of retinal input data, the final interpretation being one that satisfies the most factors and wins out among competing interpretations (hypotheses). It appears that computational inferential reasoning is necessary to achieve this.

## Ambiguities in mathematically modelling the dynamics of motion perception

Robert A. M. Gregson

*rag655@cscgpo.anu.edu.au; Department of Psychology, Australian National University, Canberra ACT 0200, Australia*

This is a valuable review of experiments whose ingenuity is impressive. It is always a pleasure to read carefully described results which irrevocably undermine simplistic theories of perception; without such results the exercise of constructing mathematical models can become vacuous.

Wertheim's coverage of motion perception, which interlocks in ways we do not fully understand with the perception of space (Dzhafarov 1992), advances analyses at two levels; vector summation of velocities and flow diagrams. It is purely with his use of mathematics that I seek clarification. I have no objections to his use of deterministic as opposed to stochastic algebra, but my first difficulty arises in his Equations (17) and (18) in section 5.3. Let us rearrange (17) to read

$$V_{est1.s} = V_{ref} - V_{ret2} - V_{stim2.s} + V_{stim1.s} + .414JND \quad [17A]$$

There is a problem here in that a JND is traditionally interpreted (as Wertheim does in the last paragraph of sect. 5.2) as a

minimum sort of quantal jump which is modulo 1 (that is, it can only take integer values), so a quantity less than unity is by definition zero. Perhaps Wertheim is not using JND in that sense here, but as some second-order discontinuity in the system. Now, (17)-(18) gives

$$V_{ret2} + V_{stim2.s} = V_{eyes.s} + \sqrt{2}JND \qquad [17B]$$

But if we do rearrange to get [17B], how do we interpret this? Is there a constant difference between $V_{eyes.s}$ and the l.h.s. of [17B], or is this a vector sum equation in which the resultant of the l.h.s. always predicts $V_{eyes.s}$ with a fixed (vectorial) disparity, which Wertheim calls a noise factor in his discussion of (12) in section 5.3? Equation [17B] is as I have written it; if I can in his algebra add JNDs linearly and not vectorially, then I can take .414 JND to be zero and get

$$V_{ret2} + V_{stim2.s} = V_{eyes.s} + JND \qquad [17C]$$

If I cannot take (17) and (18) as simultaneous equations, then this is an interesting algebra, but there seem to be implicit rules in the manipulation of the equations which need spelling out, perhaps as some sort of axiom.

I accept that the kinetics of visual motion involves eyes movement, object movement, and perceived motion, but I find ambiguities in how to represent the system's properties algebraically. I find more interesting Wertheim's observations that the process of moving into dynamic equilibrium in motion perception is not instantaneous (of course, in a biological system it would be amazing if it were); the vector equations are thus representations of the outcome of a process, and not a process description in themselves. Delays are sometimes sufficient in themselves to introduce nonlinearities in system dynamics, which may be what Wertheim is trying to capture by adding in JNDs. I should have thought that the appropriate algebra to describe processes which can be destabilised and can restabilise themselves under continuous input, was delay-differential equations. Let us consider the status of Wertheim's Figure 7 (in sect. 6.2), where he gives a lines-and-boxes description of the system. This isn't a mathematical model; it tells us what a system might do if it existed. I counted 14 functional links between boxes, some of which contain scalars (velocities) some estimates of scalars, and others mechanisms or subsystems which are probably already known to be nonlinear. To test this theory, I would ideally want to simulate Figure 7 on a computer, and to do that I need the 14 delay-differential equations of the links between boxes, plus three inputs, to get two outputs, taking the mechanism boxes as given a priori. Some of the boxes (the velocities) can be coalesced with their input-output links, thus reducing the order of the system. There are no feedback loops in Figure 7 (unless they are within the "mechanism" boxes), so if the system can become stable, it has to be interpreted as embedded in a larger system with feedback through "will," "image," and "percepts." Whether one can defensibly represent such things as "will" by scalars is by no means certain. If Figure 7 is completely open-loop (admittedly with step functions in some places) it could be reduced to one equivalent equation with discontinuities allowed in input-output mappings, but it is not clear that this would be the vector calculus which Wertheim uses.

Now, how can such a system be externally identifiable? As Walter (1982, p. 96) comments, "the computational burden imposed by the test of a model for structural global identifiability grows very rapidly with the model order," and "the chosen parameters do not correspond to the degrees of freedom of the problem."

There is a defence for what Wertheim is trying to do, in that he is studying an adaptive system, and modelling such systems can sometimes be facilitated by splitting the dynamics into slow dominant components and residual high frequency parasitic frequencies (Ioannou & Kokotovic 1983) which would presumably correspond to the brief periods in which motion vision runs to stability. As the exercise stands at present, it is not a mathematical model, but it lists some properties a model lying between physiology and psychophysics would have to reproduce.

## Computational aspects of motion perception during self-motion

Itzhak Hadani and Bela Julesz

*hadani@gandalf.rutgers.edu; Department of Psychology and Laboratory of Vision Research, Rutgers University, New Brunswick, NJ 08903*

Wertheim has made two significant contributions to the field of space perception. First, he has diagnosed some theoretical difficulties in both inferential and direct perception approaches that try to explain the phenomena of visual stability, object motion, and self-motion. Second, he has advanced a theoretical model that assumes the existence of a reference signal which gets inputs from visual, ocular, and vestibular outputs. The main problem with this model is that it is physically unsubstantiated and can serve only as a qualitative framework. This is mainly because the perception of visual stability, ego motion, and object's motion cannot be accounted for kinematically without considering ego distance and structure of visual objects; these elements are missing in the model.

To see the problem of what is required from a moving observer in order to perceive real motion of objects in space, let us turn to the simpler case of perceiving static objects as static. To obtain visual stability of objects, one needs to stabilize the retinal image against eye movements in space. We distinguish between two cases: (a) Stabilization is carried out solely on the basis of retinal information, as assumed by the direct perception theory. In computational vision this capacity is called *passive navigation* (Bruss & Horn 1983). (b) Stabilization is carried out with the aid of extraretinal signals, which we call *active navigation*. Wertheim correctly concludes that both types of navigation play a role in space perception. It is clear, however, that extraretinal signals are insufficient to account for the static appearance of objects because they are inaccurate (due to noise in the ocular control mechanism) or partially missing (due to insensitivity of the vestibular organ to linear motion).

We therefore restrict our arguments to passive navigation. Several approaches have been used to address the issue; the discrete approach, the differential approach, and the least square approach. Works in all categories show that the optical flow depends upon the six motion parameters of the eye, on the (ego)distance, and on the 3-D structure of the objects (as compared to the single angular velocity vector considered by Wertheim). Works in the discrete and differential approaches are also characterized by analyzing the minimum conditions under which an ideal observer can solve the navigation problem. These minimal conditions are given in terms of the number of points and views required to extract the six movement parameters of the eye and the distance (or structure) of objects' points. The most rigorous solution was advanced by Tsai and Huang (1985). They show that seven points and two views are required to recover uniquely the distance and the motion parameters, but this solution is only good up to a scalar in the translation vector. The consensus in computational vision is hence that the problem of passive navigation in a static environment has no unique solution but a solution up to an arbitrary affine transformation (Koenderink & van Doorn 1991). This is called the "*indeterminate scale problem*" and has direct implications for visual perception in general and Wertheim's model in particular, because it means that a visual system cannot uniquely recover the reference signal from the optical flow even when the scene is static.

It is clear that the problem of passive navigation becomes more complicated mathematically and less determinate when

some of the visual objects are moving. To solve for this case the visual system has to extract from retinal projections three additional motion parameters in the space of each object's point (Hadani et al. 1980).

Interestingly enough, our main argument is not against the physical inaccuracies of Wertheim's model but against the consensus in computational vision about the indeterminacy of the passive navigation problem. In contrast to this view, it has recently been shown that the absolute distance of objects and the eye movement parameters can be uniquely and metrically extracted from the optical flow as suggested by Gibson's direct perception concept (Hadani et al. 1993; Hadani & Kononov 1993). Furthermore, to account for visual stability the later work shows how a moving observer can reconstruct dynamic (time changing) retinal projections of static objects into a representation given in time invariant space coordinates. This was done by solving the optic-flow equations in both the discrete and differential approaches. It was shown that the differential case requires only a single visual point to carry out all the computations for passive navigation, because even the smallest visual point (say a star) has an extension on the retina due to the light point spread function. Such theoretical developments may give Wertheim's model a robust support.

Another difficulty in Wertheim's scheme is the ambiguity with respect to the kind of representation his model attributes to the organism; is it egocentric, exocentric, or retinocentric? The definition of the reference signal leads us to believe that what he means is exocentric representation. This is because the reference signal is defined as encoding the motion of the retinal surface in space. However, in Wertheim's paper percept stationarity is often associated with stationarity of the image on the retina, particularly when the eyes accurately pursue a smoothly moving target. This may lead to the wrong interpretation that a stable percept is retinocentric. The problem becomes more ambiguous when the perception of moving objects is analyzed. In this case the reality is that both frames of reference, the object and the observer, are moving relative to a third frame of reference, which is the inertial space. Wertheim "solves" the problem of relative motion perception, however, by eliminating the reference signal in his equations. It is then unclear what remains of the frame of reference. These conceptual difficulties could be eliminated if Wertheim clarified the differences between events observed in different frames of reference and the attributes (coordinates) of their representations.

## The inferential model of motion perception during self-motion cannot apply at constant velocity

Richard Held

*held@id.mit.edu; Infant Vision Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139*

***The traditional model.*** Wertheim considers the scope of inferential theory to extend beyond the visual changes produced by eye movement within the orbit, for which it was developed, to movements of head and body. According to traditional theory, signals from the movement of retinal images converge together with control signals from movements of the eye in its orbit. At some center of convergence these signals are evaluated with respect to each other. Retinal signals that are discounted by control signals from movements of the eye are seen as stationary while those that are not discounted are seen as moving. One might say that the convergence center of these paired signals incorporates expectancies about the relation between body movements and visual changes. (Parenthetically, these expectancies can even be altered by a process of adaptation that has been demonstrated in rearrangement experiments; Welch 1986). Wertheim claims that, by analogy, retinal signals accompanying movements of the head that carry the eyes with it are evaluated with respect to signals about the head movement (Wertheim's reference signal). Accordingly, an observer should report visible objects as stationary when reference signals balance signals from retinal image movement whereas sufficiently discrepant signals should elicit perception of visual motion.

***Failure of the model.*** Wertheim takes up the case of a stationary and fixating observer inside a rotating drum which he believes raises problems for the traditional inferential theory. According to him, the moving image of the drum on the retina should be seen as moving since reference signals from the stationary and fixating observer are inadequate to discount the retinal signals. Instead, observers often see the drum as stationary while feeling themselves in rotation – the familiar phenomenon of circular vection. Wertheim claims that seeing the moving drum as stationary contradicts traditional inferential theory which, consequently, requires modification. His modification incorporates the strange claim that in the absence of a reference signal from body movement, the retinal motion signal must provide the missing reference signal. Thus he postulates that the retinal signal itself provides the reference signal that discounts the motion signal it also provides. The evidence that such a reference signal actually exists is dubious, as is the existence of a "low band pass" pathway allegedly filtering this signal with a time constant measured in seconds. Apart from these considerations, however, one wonders why such contorted reasoning is necessary at all.

***Wertheim's oversight.*** I suspect that the problem derives from an uncritical extension of the traditional inferential model of eye-in-orbit to head and body movements. Let us reconsider the extended inferential theory. If an observer's movements are restricted to those of the eyes, the traditional model applies. Efferent signals that control displacement, velocity, and acceleration of the eyeball provide a reference signal against which any movement of a visual target can be compared and either discounted or not. If the observer is restricted to movements of the head on torso, neuronal signals that control and monitor the kinematics of these head movements will also provide a reference signal against which any and all movements of a visual target can be compared and either discounted or not. Now consider an observer who moves with head rigidly connected to the torso (head-body). To the extent that movement of the head-body entails acceleratory components, a reference signal will be generated that can be compared with movements of a visual target. However (and this is the crucial issue), in this case of head-body movement no reference signal, independent of vision, exists for movements of constant velocity. The sensors of the vestibulum and perhaps the viscera respond only to changes in velocity. Efferent commands to the limbs which can move the head-body, and might play a role as reference signals, activate musculature that does not maintain movement at constant velocity. The net result is that under the conditions of Wertheim's example of prolonged drum rotation at constant velocity (vection), no evaluating reference signal exists. The reference signal is not merely of zero value, hence failing to discount the retinal signal, but of no value. Consequently, the conditions for applying the inferential theory are absent. Wertheim fails to appreciate the inapplicability of the model in this case; and, in an effort to preserve its relevance, he invents the visual reference signal.

It is the absence of a reference signal which makes attribution of the source of the retinal motion signal quite ambiguous, unlike in the case of movement of the eye in its orbit or movement of the head on the torso. The motion is equally likely to have resulted from either head-body movement or drum movement at constant velocity or any combination thereof that sums to the required value. The occurrence of the trade-off between perception of self-movement and object movement – the waxing and waning and alternations between object motion and self-

motion – under these conditions confirms the validity of this implication. Many investigators have discussed this bistability of vection, which seems counter to the predictions of Wertheim's new model (not to mention direct perception ideology). As a consequence of its ambiguity, the attribution of motion is subject to attentional shifts, transient movements, perturbations in visual stimuli, and habituation. For example, the perception of drum motion at the initiation of rotation may be attributed to the discrepancy between the rapid change in the retinal signal (acceleration?) and the absence of a discounting signal from the stationary head-body. Several seconds may elapse before saturated vection is achieved. Such visual-vestibular interactions commonly have time constants of the order of several seconds, unlike most purely visual phenomena. Perhaps one should add here that vection can also be produced by sinusoidal movements of the drum which entail accelerations of the retinal signal. These motions should be evaluated in conjunction with reference signals. Presumably, they may at times be so evaluated, but vection is also frequently reported. The answer here is that the visual system appears to be relatively insensitive to acceleration of the retinal image. It appears likely that low levels of acceleration are not distinguishable from constant velocity signals.

In conclusion, once the inapplicability of the inferential model to movement of the retinal image at constant velocity is recognized, no convincing need remains for the addition of the visual reference signal proposed by Wertheim. Some form of the extended inferential model appears to apply in all cases in which a reference signal (nonvisual) is defined and in no cases in which it is not.

## Some problems with the gain of the reference signal

Hitoshi Honda

*psyhonda@cc.niigata-u.ac.jp; Department of Psychology, Faculty of Humanities, Niigata University, Niigata 950-21, Japan*

Many perceptual phenomena related to saccade and/or pursuit eye movements have been explained by the classical subtraction (inferential) theory. It is evident, however, that this theory is too simple to deal with the complex aspects of these perceptual phenomena; there have accordingly been many attempts to correct the theory by adding complementary psychological mechanisms such as visual masking, visual adaptation, information from various visual cues or higher cognitive knowledge about the visual world, and so on.

Wertheim's new model is based on a reconsideration of the concept of an extraretinal signal, replacing the extraretinal signal by a reference signal that includes (any combination of) an efference copy, a vestibular component, and a visual component. He claims that this model can successfully explain various perceptual illusions that occur without eye movements (such as motion aftereffect and circular vection) as well as perceptual phenomena that are observed during pursuit eye movements (the Filehne illusion and Aubert-Fleischl phenomenon).

1. Wertheim's model was originally designed to describe the perception of object motion or stationarity during self-motion. However, as the author himself mentions, it can be extended to include the perception of the position of external objects in space. In this regard, I would first like to comment on the gain of the efference copy signals from the pursuit eye movement system. In one of his experiments, Wertheim (1987) examined the possibility that the size of reference signals would be affected by optokinetic stimulation, estimating that an efference copy signal encoded about 80% of eye velocity in the head. This gain magnitude of 0.8 is nearly consistent with the 0.8–0.9 I estimated as the gain of efference signals available for judging the distance tracked by pursuit eye movements (Honda 1990). Mack and Herman (1972) and Miller (1980) estimated similar magnitudes. Also, in both Wertheim's (1987) and Honda's (1990) studies, it was shown that the gain is independent of the velocity of the eye. It should be noted here that Wertheim estimated the gain of efference copy signals for eye velocity whereas Honda estimated gain for the distance tracked by the eye.[1] This coincidence suggests the visual system knows that distance is mathematically related to velocity.

It is interesting to compare the gain of the efference copy signal from the pursuit eye movement system with that from the vestibular system. Wertheim estimated reference signal gains ($G_{ref}$) when subjects were moved linearly without eye movements in the total darkness. The $G_{ref}$ for the otolith afferent response was 0.76–0.8. Again, this is about the same as the $G_{ref}$ estimated for efference signals from the pursuit eye movement system. The coincidence of the $G_{refs}$ between the otolith and the pursuit eye movement system is plausible, because in each system the undersized reference signal is compensated by visually induced optokinetic stimulation.

On the other hand, the gain of the semicircular canal afferents was shown to be 1.07 (Wertheim & Bles 1984). It is not apparent why there is a discrepancy between the gain of the reference signal from the otolith system and that from the semicircular canal system. One possible explanation may be that unlike forward or backward linear self-movement, ego rotation (head rotation) is usually accompanied by slow (reflex) eye movements.

2. As mentioned earlier, Wertheim's model can successfully explain many perceptual phenomena that occur during slow eye movements. There are some problems, however, which need further explanation.

My first question concerns how the size of the visual component in a reference signal is adjusted to a level appropriate to maintaining the stability of the visual world. As already mentioned, the gain of efference copy signals from the slow eye movement system is about 0.8. According to Wertheim's model, the underregistration of the efference copy signal is compensated by a visually induced optokinetic component. Therefore, during slow eye movements, the size of compensation should be about 0.2, never exceeding this value. Wertheim supposes that this regulation of visual components in $G_{ref}$ is accomplished with the help of a gating mechanism in the optokinetic pathway which acts as a low band-pass spatiotemporal filter. However, Wertheim (1987) showed that, in at least one experimental situation, this mechanism does not work appropriately: with a low spatial frequency visual stimulus the $G_{ref}$ become oversized ($G_{ref} > 1$). A solution to this problem may come from the fact that in normal daylight circumstances there is usually not such a low spatial frequency pattern. However, the overregistration dramatically increases with eye velocity (35.6 deg/sec vs. 23.3 deg/sec in Wertheim's experiment). This overregistration, which happens during relatively fast eye movements (> 30 deg/sec), seems to be overcome by the additional fact that in normal everyday life we seldom track a fast moving object with smooth eye movements; instead, the tracking is usually interrupted by many corrective saccades.

The second problem, as Wertheim is aware, is that according to his model image flow across the retinae generates optokinetic stimulation which, on the one hand, produces a visual component in reference signals, and on the other hand, is used for perceiving self-movement or stasis in space. There is hence a possibility that pursuit eye movements across a highly optokinetic stimulus pattern induce impressions of self-movement (vection). Wertheim solves this problem by postulating two gating mechanisms with different thresholds, one for generating ego motion and the other for converging on the reference signal. This explanation is insufficient. What one needs to know is the neurobiological or behavioral basis on which the thresholds are determined.

NOTE

1. A pursuit eye movement is frequently interrupted by small corrective saccades (Collewijn & Tamminga 1984; Puckett & Steinman 1969). The interruption by saccades occurs when the target velocity exceeds 30 deg/sec. In Wertheim's (1987) study, the target moved at 25 deg/sec or 40 deg/sec. Hence it seems that in his experiment the tracking of a moving target contained many small saccadic components. To estimate the gain of efference signals precisely, one must know about the role the saccadic component plays in determining the apparent velocity and position of external objects. A clue to the solution of this problem comes from the generally accepted idea that saccades are predominantly triggered by a displacement of targets, whereas pursuit eye movements are triggered by a change in target velocity. Based on this idea, we can assume that the reference signal gain ($G_{ref}$) of saccades is 1.0 when it is used for position judgment (Honda 1990) but 0 when it is used for velocity judgment.

## Optical foundations of perceived ego motion

Nam-Gyoon Kim and M. T. Turvey

*gyoon@uconnvm.bitnet; Center for the Ecological Study of Perception and Action, University of Connecticut, Storrs, CT 06269-1020*

Motivated by the results of experiments on human observers in rotating drums, Wertheim hypothesizes that for an organism to perceive motion (that is, self-motion and object motion), a special reference signal is needed which encodes the time derivative of the position of the retinal surface relative to the environment. There are, however, reasons to believe, that the optical conditions characterizing the empirical phenomenon motivating Wertheim's model have been insufficiently analyzed. Any analysis of the optical structure at a point of observation *O* must identify what is exterospecific and what is propriospecific. In the rotating drum, as the striped drum rotates about *O*, there is accretion and deletion of the optical solid angles adjacent to the solid angle corresponding to the fixation region, a small surface or light spot interposed between *O* and the drum, on which the observer is required to focus. These optical solid angles constitute the optic array at *O*. Because there is no disruption or deformation of internal optical structure as would follow from a perspectival transformation (displacement of *O*), it specifies one surface (the drum wall) passing behind another (the fixation surface). At the same time, there are no widenings or narrowings in the surface layout between the edges of the occluding fixation surface and the peripherally extreme edges of the drum. The optical solid angles whose bases are the alternating black and white stripes of the drum undergo continuous changes in their adjacency structure but undergo no change in their distribution relative to the fixation solid angle. This particular global pattern of change and nonchange is specific to a rotation of the head-body in pursuing the movement of an object in a uniformly textured environment.

In sum, the optical structure at *O* contains (a) exterospecific information about two surfaces, one in front of the other relative to *O*, with the further surface displacing relative to the nearer, and (b) propriospecific information about the rotation of the observer. The optic array in the rotating drum specifies an ecological contradiction (Gibson 1979, p. 215) and the often vague, nonuniform, vacillating responses of experimental subjects (see Mergner & Becker 1990) suggest that they respect this contradiction perceptually. Subjects behave as one would expect if the normal (lawful) complementarity of optical exterospecific and optical propriospecific information had been violated.

What of the time scales Wertheim considers to be critical? Properties of the optic array specific to properties of an organism-environment system are revealed over spatiotemporal transformations (Gibson 1966; 1979). Their time course depends on the properties in question and, in the laboratory, on how faithful the experimental conditions are to those of the organism-environment systems they are meant to capture. Nevertheless, there are reasons to believe that the claim that there are different lag times may be overstated. Other experimental results suggest a closer temporal fit between optical conditions and perceived self-motion and object motion (e.g., Lishman & Lee 1973; Ohmi et al. 1987).

There is apparently more to the optical structure in the rotating drum experiment than Wertheim recognizes. The preceding description suggests, among other things, that the resultant perceptions and their time scales depend nontrivially on characteristics of the fixation region, the number and uniformity of the stripes, and the rotational frequency and radius of the drum – all considered in relation to each other. The theory advanced by Wertheim seems to be a case of overinterpreting an insufficiently developed empirical data base. At the very least, a careful exposition and experimental analysis of the optical conditions for the "circular vection" phenomenon is needed.

At best, Wertheim's mechanism addresses the fact that movement (of self and/or a part of the surroundings) is taking place. It addresses neither the kind of movement (e.g., rigid vs. nonrigid) nor its direction (e.g., heading; Warren & Hannon 1988). Within this restricted domain, the proposed mechanism seems limited to a single object displacing relative to the observer. A "retinal signal" is said to encode the velocity of the retinal image; if this signal is matched by the reference signal then it is ego movement that is taking place, not object movement. Suppose that many objects are moving relative to the observer, in different directions with different speeds – as is commonly the case. What could it then mean to speak of perceived object velocity as due to the magnitude difference between retinal and reference signals minus their JND (just noticeable difference)? Our suspicion is that Wertheim's model requires many ad hoc adjustments to extend to real-world examples.

To illustrate some of the above points, we consider one candidate strategy for modeling the optic flow field and its perceptual implications. (This candidate does not consider occlusion and disocclusion.) Suppose *O* is moving along a circular path parallel to the ground plane in a stationary environment. Optically, this corresponds to the ground plane rotating beneath a stationary *O* and each surface element moving along a circular path. Then, at each point in the ground at every instant, the velocity of the element passing through the point can be represented as a vector *V*. Moreover, if *O* is moving at a constant speed, the velocity vector at any point is constant. A time-independent flow field is called *steady*. In this case, particle trajectories can be determined by the fact that the velocity of a particle is everywhere tangent to the trajectory. Trajectories are obtained by plotting curves, called *streamlines*, so that their direction at each point agrees with the direction of the velocity *V* at that point. Each streamline can be represented as $\Psi(x, y) = c$. Assigning different values to $c$ gives a family of streamlines whose totality is a *one-parameter family* of curves with $c$ the *parameter* and with $\Psi$ the *stream function*. Given, $\Psi$, the velocity of flow at a given point is

$$u = \frac{\partial \Psi}{\partial y} \text{ and } v = -\frac{\partial \Psi}{\partial x}$$

where $u$ and $v$ are projections of $V$ on the coordinate axes. In hydrodynamics, no internal friction between particles means no particle rotation. The result is *irrotational* flow. When the flow is irrotational, there is another family whose curves intersect the streamlines at right angles. This family of curves, $\Phi(x, y) = c$, has its velocity as

$$u = \frac{\partial \Phi}{\partial x} \text{ and } v = \frac{\partial \Phi}{\partial y}.$$

$\Phi$ is the *velocity potential* of the flow. A comparison of the velocity components of $\Phi$ and $\Psi$ reveals that they satisfy the Cauchy-Riemann equations

$$\frac{\partial \Phi}{\partial x} = \frac{\partial \Psi}{\partial y} \text{ and } \frac{\partial \Phi}{\partial y} = -\frac{\partial \Psi}{\partial x},$$

which means that $\Phi$ and $\Psi$ are the real and imaginary parts of a single analytic function

$$F(z) = F(x + iy) = \Phi(x,y) + i\Psi(x,y) \quad (1)$$

$F(z)$ is the *complex potential* of the flow. The curves of these two families are mutually orthogonal and their trajectories are the orthogonal trajectories of the given curves. The curves $\Phi(x, y) = c$ are the *equipotentials*, and the orthogonal curves $\Psi(x,y) = c$ are the *streamlines*. The velocity of the flow can be obtained by differentiating (1) and using the Cauchy-Riemann equations

$$F'(z) = \frac{\partial \Phi}{\partial x} + i\frac{\partial \Psi}{\partial x} = \frac{\partial \Phi}{\partial x} - i\frac{\partial \Phi}{\partial y}$$
$$= u - iv.$$

It follows that

$$V = u + iv = \overline{F'(z)}.$$

The analytic function techniques facilitate investigating various flow patterns because the real and imaginary parts of an analytic function of the complex variable $z$ are the velocity potential and stream function, respectively, of a steady flow. An analytic function, split into real and imaginary parts, produces an infinite variety of patterns of streamlines.

Returning to circular optical flow; as the observer moves along a circular path, the corresponding flow field is a family of concentric circles, where the observer's path is the circle beneath $O$, with radius $r$ given from the center of rotation. These circles correspond to the field's streamlines. Consider an analytic function

$$F(z) = \frac{K}{2\pi i} \ln z,$$

where $K$ is a real constant. From $z = re^{i\theta}$, we obtain

$$\Phi = \frac{K}{2\pi}\theta,$$

$$\Psi = -\frac{K}{2\pi} \ln r, \quad (2)$$

and,

$$V = u + iv = -\frac{K}{2\pi i}\frac{1}{\bar{z}}.$$

Equation (2) shows that the streamlines are the circles $r = c$ (see Fig. 1), while the velocity is constant on every streamline. When $K > 0$, the flow goes counter-clockwise, specifying forward locomotion of $O$; when $K < 0$, the flow goes clockwise, specifying backward locomotion. Given this, the question is whether these properties in optical flow can be detected by an observer (Warren et al. 1991).

How does an animal distinguish its own movements from those of objects? Consider a *streamtube* – a closed curve in the flow with all the streamlines passing through the curve. Thus, a two-dimensional streamline pattern becomes a plane view of three-dimensional tubes as if there were rigid walls inside which fluid flows. When the observer moves, the array undergoes a global transformation, which for circular movement results in a torus-like streamtube. Conversely, when an object moves, a local region of the array transforms. This corresponds to a local streamtube within the global streamtube. Local streamtubes are distinguished in the flow field because streamlines do not cross each other, being everywhere parallel to the direction of flow. This existence of nonparallel flows in optical flow means there is more than one streamtube. The local streamtube means an object moving independently of the observer.

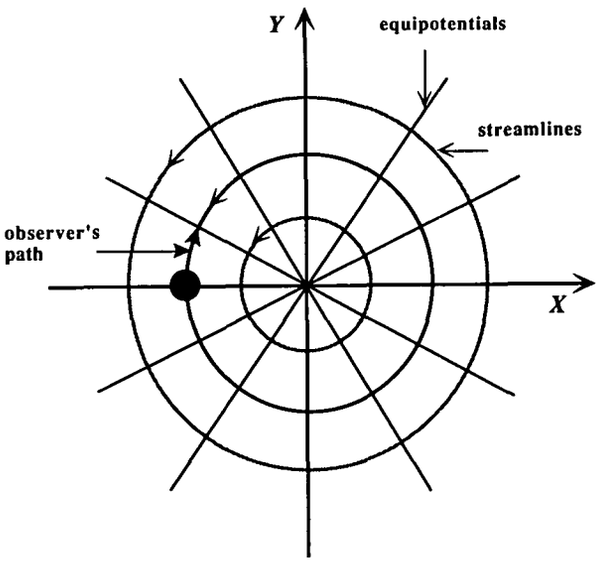


Figure 1 (Kim and Turvey). Circular flow field defined by the complex potential $F(z) = K/(2\pi i)\ln z$. The equipotential lines are the straight lines and the streamlines are the concentric circles. The flow circles around the origin, from which the observer's path lies with radius $r$.

Streamlines created by the movement of a neighboring train completely occupy one's view through the window. The impression is of one's own train moving. The limited view eliminates the distinction between global and local streamlines, resulting in the otherwise local streamlines going proxy for the global. Take another example. For a bird to stay in the same place in the wind or a fish to maintain its position in a flowing river, they must do the opposite – they must move to remain stationary. Optically, this means they have to cancel the velocity vectors of the flow field. These examples suggest that ad hoc devices such as optokinetic stimulation are not needed to explain optically induced ego motion. The so-called illusions are the lawful outcomes of ecological optics.

What, then, is the role of the eyes? Why do they rotate? Optical flow is described by various patterns of streamlines. Specifically, optical flow in circular locomotion results in a torus-like streamtube. The retina meets these streamtubes orthogonally, taking their cross sections – the eye coincides with the cross section of the flow field. The cross section of any vector field is called a Poincaré map. A map constructed from a flow via a cross section is generally unique, allowing the study of how visual perception is specific to flow. Animals with retina-like optical systems should benefit from the ability to rotate them, because through rotation they can sample a larger region than otherwise would be allowed. That is, animals rotate their eyes not to generate efference copies, but to detect as much information in the optical flow as possible.

## Wertheim's "reference" signal: Successful in explaining perception of absolute motion, but how about relative motion?

S. Mateeff[a] and J. Hohnsbein[b]

*[a]percept@bgcict.bitnet and [b]hohnsbein@ifado.arb-phys.uni-dortmund.de; [a]Institute of Physiology, 1113 Sofia, Bulgaria and [b]Institut für Arbeitsphysiologie, 4600 Dortmund, Germany*

We like Wertheim's target article. He has succeeded in assembling many pieces of the human spatial orientation puzzle into a

clear unitary picture. As longtime admirers of the "inferential" approach developed by such authors as Rock (1977; 1983) and Epstein (1973), we are pleased to see Wertheim developing a model that regards motion perception as the result of cooperative contributions from different sources of information: retinal, oculomotor, and vestibular. Wertheim has replaced Rock's "unconscious inference" by operations of addition and subtraction, which makes his model more explicit.

Wertheim has been especially successful in two points. First, he has shown how to manipulate the visual input to change the strength of the Filehne illusion, and even to invert it. These experiments strongly support the existence of a "reference" signal that contains a substantial visual component. Second, he has clearly pointed out that the visual input per se has no fixed perceptual significance, since stimuli with the same retinal velocity may lead to quite different percepts depending upon eye velocity: for example, to percepts of stimulus stationarity or of stimulus motion in a certain direction. This demonstrates convincingly the existence of an extraretinal, eye-movement related component in the "reference" signal.

But Wertheim's model has a strong competitor: the theory of Post and Leibowitz (1985). Decisive data in support of Wertheim's model seem to have been provided recently by Mergner & Becker (1990), who showed that no vection occurred when a low-contrast pattern was sinusoidally moved with low velocity and amplitude in front of the subject and was concomitantly pursued with the eyes. According to Wertheim, this finding can be predicted by his model, but not by the theory of Post and Leibowitz: slow eye movements produce an efference copy that is strong enough to elicit a perception that the pattern which is otherwise stationary on the retina moves in space and that the observer's body is stationary. This explanation, however, poses additional problems arising from the common occurrence of circular vection (see sect. 6.2 of the target article).

Wertheim's model can successfully predict and explain experimental data on the perception of the "absolute" motion of objects and patterns, that is, visual motion relative to external space. In our opinion, some problems may arise when perception of "relative" motion (motion of two or more objects relative to each other) is considered. According to Wertheim, the velocity of relative motion is perceived as the difference between the perceived absolute velocities of the two objects, or as the difference between the two retinal image velocities (see his Equation 12). His model implies, however, that the percepts of the absolute velocities of the two objects are also (automatically) given (Equations 10 and 11). The question then arises: What happens when two objects move relative to each other and relative to the external space; will the observer see the two absolute motions or the relative motion?

To clarify the question, let us consider an example solely from the standpoint of Wertheim's model, without reference to any other existing theory. A bicycle moves along a street and a stationary observer pursues the axle of one of the rolling wheels with his eyes. There is a mark on the tire. The axle moves linearly relative to external space; thus, the absolute trajectory of the mark is cycloidal. Pursuing the axle with one's eyes usually results in a percept of a linear motion of the axle and a circular rather than a cycloidal motion of the mark relative to the axle (Stern & Emelity 1978). Wertheim's model explains how the absolute motion of the axle (its retinal image being stationary during the pursuit) is perceived. The model also explains how the circular motion of the mark relative to the axle is veridically perceived, but it tells us nothing about why the absolute, cycloidal motion of the mark is not perceived. This again raises the question of how the visual system makes a "choice" as to which of the three motions will be perceived. Is the choice determined by some parameters of visual stimulation, or is it a result of focusing attention on certain visual details and features?

In our opinion, both factors may determine the choice. As Mateeff et al. (1987) have shown, when a small (0.35 deg) test object moves too fast (15 deg/sec) during ocular pursuit of another target, the degree of constancy may decrease to zero; the observer perceives the retinal image trajectory of the test object, that is, its motion relative to the pursued target. Increasing the size of the test object and decreasing its velocity gradually lead to a recovery of constancy and the motion of the test object is experienced in space rather than relative to the pursued target. Thus, perception of the trajectory of the test object is determined by the ratio between its size (in deg) and its velocity (in deg/sec); when this ratio becomes more than 300 msec, the absolute motion of the test object is perceived. This result is in a good quantitative agreement with the data of Ehrenstein et al. (1987) that are discussed by Wertheim; both results show the importance of the parameters of the stimulation on motion perception during self-motion.

The attentional factor may also play a role in the choice between perceiving relative or absolute motion. Mateeff (1980) demonstrated that when a small test object moves at a high velocity noncolinearly relative to a point that is tracked and the observer has to report the direction of motion of the object, its retinal direction is reported. Using the same physical stimulus parameters but changing the task so that the observer now has to localize the beginning and the end points of the test object's trajectory, the observer's reports correspond much more to the external, absolute motion rather than the retinal, relative motion of the test object. The choice of which motion to perceive appears to depend on the observer's task.

Finally, Gogel and Sharkey (1989) used horizontal physical motion of objects to induce perceived horizontal motion in a vertically moving test spot that was pursued visually. They showed that attending to or ignoring the inducing objects resulted in a significant change in the direction of the perceived trajectory of the pursued test spot. This finding also indicates the need for Wertheim to extend his welcome model by including at least an input for direct top-down influences.

## Inferring the visual reference

Fred H. Previc

*previc%kirk.decnet@hqhsd.brooks.af.mil; Crew Technology Division, Crew Systems Directorate, Armstrong Laboratory, Brooks Air Force Base, TX 78235-5104*

In his target article, Wertheim does a commendable job of arguing that ambient visual information serves as an important reference for differentiating ego motion from object motion in space. I would like to add two caveats that qualify his general assertion that the existence of an ambient visual reference is consistent with the theory of "direct perception." The first of these seems to relate to Wertheim's description of the same visual reference for both oculomotor phenomena (and related ones such as induced motion) and "whole-body" percepts such as vection. The second caveat relates to the inferential processes involved in selecting the visual reference.

Wertheim seems to invoke the same visual reference signal for such diverse phenomena as vection, oculomotor control, induced motion, and even motion aftereffects. Yet there are many experiments showing that the visual stimulus driving eye movements and induced motion is fundamentally different from the one controlling self-motion perception. Most of these differences relate to spatial parameters of the visual stimulus, although optokinetic eye movements and induced motion also differ from vection in their temporal dynamics (Heckmann & Post 1988). Brandt et al. (1973), for example, showed that optokinetic nystagmus and vection are driven primarily by the central and peripheral portions, respectively, of a surround whose central and peripheral sectors move in opposite directions. We have shown a similar relationship between induced

motion – a phenomenon that is related to nystagmus-suppression and other aspects of pursuit tracking – and vection: no roll vection was elicited in our study using the same central region (~50 deg) that gave rise to potent induced motion (Previc & Donnelly 1993). Heckmann and Howard (1991) have further shown that a coplanar stimulus eliciting superior induced motion does not stimulate the vection mechanism nearly as well as a more distant stimulus (see also Brandt et al. 1975; Delorme & Martin 1986; Ohmi et al. 1987). Finally, motion aftereffects are also strongly elicited by very small visual stimuli, but require higher spatial frequencies (>0.5 c/deg) than those which produce good vection (Cameron et al. 1992).

The possibility that different references exist for oculomotor and related phenomena versus whole-body percepts such as vection is probably less detrimental to Wertheim's theory than are the reasons *why* such different references exist. According to Dichgans and Brandt (1978), the most distant and peripheral portions of the visual field are used to infer self-motion because motion in such regions is "the inevitable consequence of movement of the body in space" (Dichgans & Brandt 1978, p. 778). By contrast, motion in the central 60 deg of peripersonal (near-visual) space is composed mainly of movements of the arm and of objects we manipulate; we must disregard these as references for head or body motion in space. Oculomotor outputs and induced motion, on the other hand, are typically more affected by *centrally located surrounds* in or slightly beyond the plane of convergence (Heckmann & Howard 1991; Howard & Marton 1992), mainly because foveally tracked objects rarely appear in front of a background surface. The induced motion imparted to a tracked object by a coplanar or more distant central surround is apparently useful in both pursuing and manipulating objects in peripersonal space; this explains why induced motion is closely related to visually mediated manual control mechanisms (see Previc & Donnelly 1993). The above distinction between oculomotor and whole-body responses may also have neurophysiological reality, in that separate parietal areas seem to mediate oculomotor control and reaching in peripersonal space (the inferior parietal area) versus whole-body movement through more distant space (the superior posterior parietal area; Brain 1941; Grusser 1983).

Much evidence indicates that the selection of a visual reference for establishing whether or not self-motion has occurred itself requires a set of perceptual decisions. This is illustrated by the fact that a moving surround perceived as the background elicits stronger vection than one that is perceived as the foreground even if they are at the same *actual depth* (Ohmi et al. 1987). Likewise, vection is reduced when subjects *attend to* (but do not actually fixate) a more distant surround lying behind a proximal fixation target (Mergner & Becker 1990) – which makes sense given that the perception of ego motion is carried out by *ambient* (nonattentive) visual processes. Another indication of the inferential nature of vection is its much shorter latency when the subject has been previously rotating around a vertical axis (Mergner & Becker 1990) even if the lag of the endolymph in the semicircular canals has dissipated and the nonvisual sense of rotation has subsided. Finally, we have recently demonstrated that large visual fields which ordinarily elicit good vection may be unable to overcome illusory orientational percepts created by shifting the gravitoinertial vector in a motion-based simulator if they are presented in a head-mounted visual display that subjects believe to be *part of their frame-of-motion* (i.e., a moving vehicle; Previc et al. 1992). Conversely, a much smaller "real-world" scene that is located outside the frame-of-motion can be quite effective in overcoming these illusions. For example, the illusory climbing sensation during takeoff in an aircraft is not eliminated by looking at a large subject-fixed image such as the bulkhead of the aircraft, which is interpreted as lying within the frame-of-motion, whereas a view through a small window at an external reference (e.g., the horizon) can completely break this illusion. Similar problems in interpreting spatial orientation have been noted when no outside visual reference is present in a 0-g environment (Lackner 1992; Oman 1988).

In summary, Wertheim fails to recognize that we use different portions of the visual world to reference oculomotor behavior and related phenomena on the one hand, versus spatial orientation and vection on the other. Nor does he acknowledge the critical *inferences* made in adopting such references, especially in the case of whole-body ego motion. By contrast, I have suggested an "ecological" visual approach highlighting the important decision-making that helps establish the perceptual "frame" for everyday visuomotor activities, even though such inferences become dramatically apparent only during abnormal acceleratory and other sensory-conflict situations.

## Why another alternative optokinetic model?

Thomas Probst

*probstt@ze8.rz.uni-duesseldorf.de; Department of Experimental and Clinical Neuropsychology, Heinrich-Heine-University of Düsseldorf, D-40225 Düsseldorf, Germany*

Wertheim's target article is a comprehensive review of motion perception, particularly self- and object-motion perception. It corresponds largely to Chapter 9 (Wertheim 1990) of the book edited by Warren and Wertheim (1990), including both the functional and the adapted functional model describing the generation of reference signals and the interfacing of percepts of object motion and ego motion in space (Figs. 1 and 7 of the target article). The additional connections between the "oculomotor mechanism" and the "retinal receptors" on the one hand and the "estimator of head velocity in space" on the other, drawn in as dotted lines in both figures are no substantial improvement of Wertheim's model because they represent pathways for the generation of optokinetic and vestibular nystagmus which are well known and described in detail in the literature on the subject.

I have difficulty with Wertheim's view of "optokinesis." The optokinetic response (OKR) is an ocular reflex that produces eye movements following a slow movement of the visual surroundings. In stationary visual surroundings, the OKR also acts to stabilize eye position because otherwise there would be continuous drifting and rapid resetting of eye movements. Eventually, vestibulo-ocular reflex (VOR) relay neurons mediate the OKR. Thus, the VOR and OKR share the same vestibular relay neurons and act conjointly to stabilize retinal images during head movement (e.g., Ito 1987). Although effects on vestibular nuclei neurons and optokinetic nystagmus (OKN) are strongest when the whole visual periphery moves, it is not necessary to have full-field motion to elicit OKN; much smaller fields are sufficient (Henn et al. 1980). Nevertheless, extending the stimulus field to the (horizontal) retinal periphery has a strong effect in increasing the angular velocity of the slow phases of eye movements during OKN (Dichgans et al. 1973). If the optokinetic stimulus is large enough ("area effect"), however, or stimulates mainly the periphery of the retina ("retinal location effect"), then this stimulus is suitable for eliciting visually induced self-motion perception called vection (e.g., Dichgans & Brandt 1978). So, Wertheim's definition to the effect that "to be optokinetic, a visual pattern must be large, have relatively low spatial frequency characteristics, move (not too fast) across the retinae, and remain visible for more than a very brief interval" is misleading.

Wertheim describes briefly and adequately two theoretical views of motion perception, the theory of "direct perception" and the "inferential theory." One of his basic questions is, why is the objectively rotating visual surround perceived subjectively

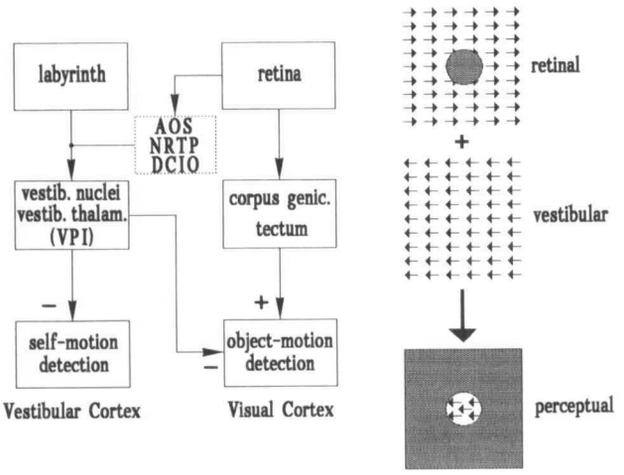


Figure 1 (Probst). Schematic diagram of visual-vestibular interaction explains the fact that during visually induced self-motion perception, small contrasts that are stationary on the retina such as a fixation point, connected to the subject's frame of reference, seem to move with the observer. In this instance the visual perception of stationarity of the objectively moving visual surround would be reevaluated on the basis of a signal originating from the vestibular system (after Dichgans & Brandt 1978). The dotted-line box is in itself a complex neurophysiological circuit mediating the main routes from the accessory optic system (AOS) including the nucleus of the optic tract and the three terminal nuclei (dorsal, lateral, medial) via the dorsal cap of the inferior olive (DCIO) and cerebellar structures to the vestibular nuclei. An additional pathway runs from the AOS to the nucleus reticularis tegmenti pontis (NRTP) and from here both directly toward the vestibular nuclei and through cerebellar structures (flocculus) (Berthoz 1981).

as stationary in space during saturated circular vection (CV)? I disagree with Wertheim's opinion that the inferential theory has a problem explaining this phenomenon. His elegantly formulated and differentiated "alternative model," which he introduces here again, is accordingly not necessary, at least not to explain the subjective stationarity of the physically moving visual environment inducing saturated self-motion perception in the opposite direction. Dichgans and Brandt (1978) introduced a schematic diagram of visual-vestibular interaction in object-motion detection and dynamic spatial orientation. They assumed that "a hypothetic explanation for the sensed stability of the actually moving visual environment during CV might be that the motion information arriving at the visual cortex is neutralized by a collateral signal deriving from the vestibular system" (p. 760), and they provided neurophysiological evidence for this. Figure 1 shows their model of intersensory neutralization, which should not be confused with the model of von Holst and Mittelstaedt (1950), who discussed cancellation of sensory inflow by an efference copy of *motor* commands. Dichgans's and Brandt's model explains the puzzling fact that during CV small contrasts that are stationary on the retina, such as a fixation point connected to the subject's frame of reference, seem to move with the observer. In this instance, the visual perception of stationarity would be reevaluated on the basis of a signal originating from the vestibular system.

The basic assumption of distinct cortical areas for object-motion perception and self-motion perception was taken up by Probst (1983) on the basis of his experiments and was further elaborated into a working model of visual-vestibular interaction (Fig. 2). Here simply structured object-motion and self-motion perception are mediated by adequate stimulation of the corresponding peripheral receptors "eye" and "labyrinth" via the well-known subcortical and cortical pathways. In addition, a large, moving visual environment stimulating the retinal periphery with low spatial frequencies leads first of all to adequate object-motion perception via retino-cortical visual structures. However, the special characteristics of this visual stimulus, which Wertheim calls "optokinetic," lead to a simultaneous stimulation of the central vestibular system in the brainstem, the vestibular nuclei, via both subcortical (eye–vestibular nuclei) and cortical (visual cortex–vestibular nuclei) pathways. The consequence of this gradually increasing ("$\int dt$," corresponding to Wertheim's "gating mechanism") stimulation of the vestibular nuclei is a dual one: it excites the cortical self-motion perception area and second, it leads to a concurrent reduction of the input signal to the object-motion perception area via a negative feedback loop. Thus, the interrelation between both centers mediates the perception of a gradual change, from pure object-motion perception, through a phase of apparent rotary body acceleration during which the surroundings seem to decelerate at the same rate, until there is exclusive self-rotation, with the visual surround appearing stationary. The working model has one restriction, however: eye movements are not taken into account. Nevertheless, a slight revision of this model also explains inhibitory vestibulo-visual interaction for orthogonal stimulus conditions with visual stimulation about the $x$-axes and simultaneous vestibular excitation in yaw, roll, and pitch (Hofstetter-Degen 1988). This slightly revised model has recently been published (Probst 1991).

The working model of Probst (1983) and Probst et al. (1986) had to be completed because of new clinical findings in patients with cortical hemianopia. According to the findings of Straube and Brandt (1987), information about full-field motion in the visual cortex must descend corticofugally to the vestibular nuclei where the perceptual decision between object motion and self-motion is reflected in latency-dependent frequency modulation of second-order neurons. The sensation of CV requires functional integrity of the vestibular cortex "switched on" by the vestibular nuclei. The velocity of self-motion, however, can be mediated by visual-vestibular cortical interaction where the decision about self-motion is made by vestibular brainstem structures (Straube & Brandt 1987). These important corticofugal pathways (e.g., from the visual cortex to the nucleus of the optic tract linked to the vestibular nuclei, partly via cerebellar structures) are missing in Wertheim's model (Schoppmann 1981).

In conclusion, despite several disagreements, it is undeniable that Wertheim has embarked on an exciting and promising line of research.

## Perception of motion with respect to multiple criteria

Gary E. Riccio

*griccio@ux1.cso.uiuc.edu; Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801*

Wertheim describes a paradigm for evaluating the relative influence of retinal and extraretinal "signals" on the phenomenology of ego motion. He uses data from such a paradigm to develop a model for the combination of retinal and various extraretinal signals in the "interfacing of percepts of object- and ego-motion" (Fig. 7). Phenomenological data can inform the study of perceived motion and Wertheim's model can help organize the data. Below I present some caveats that should be considered in the interpretation of such data and models. The caveats concern the multidimensionality of perception and phenomenal experience and, related to that, the multiple affordances of ego motion in the context of posture and locomotion.

Wertheim's paradigm allows a subject to indicate whether or

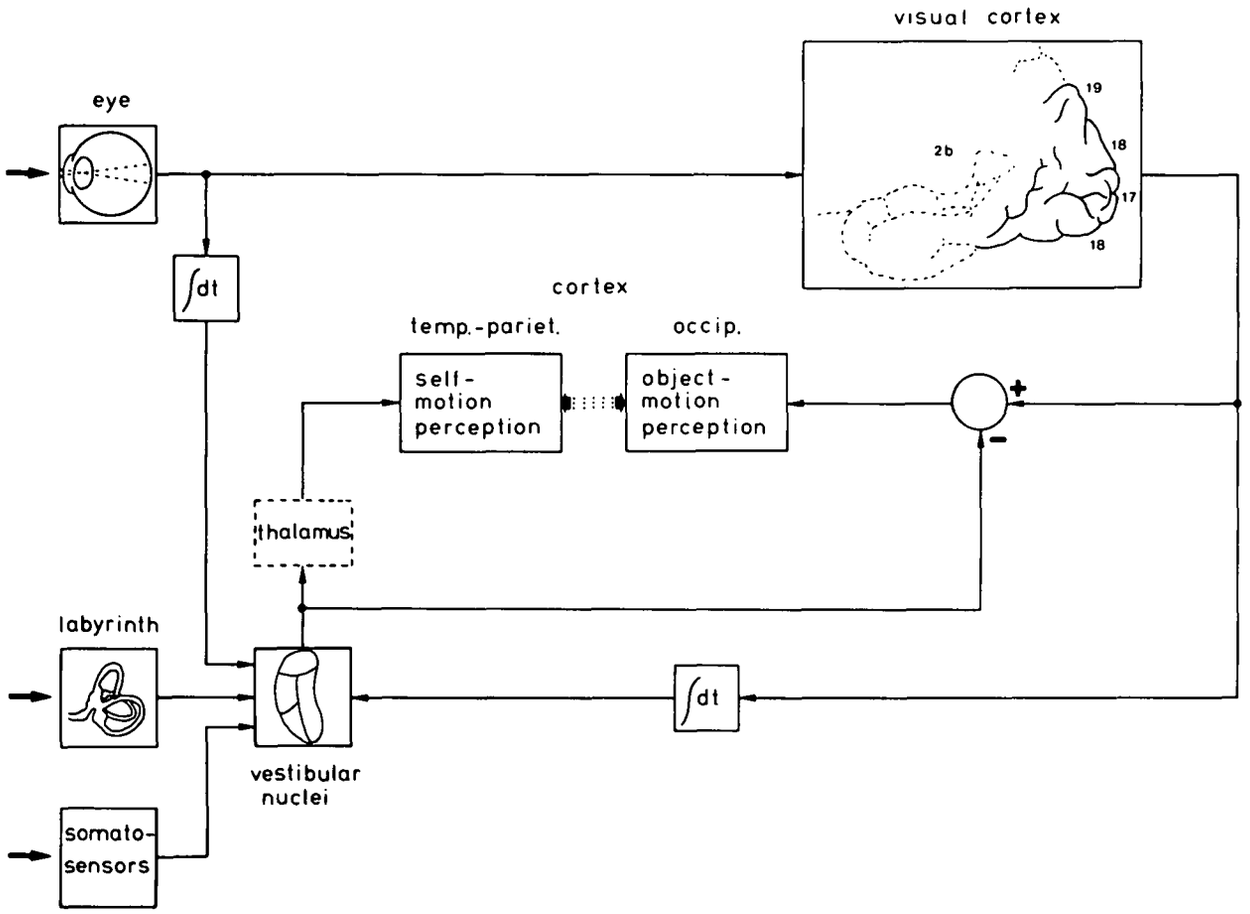


Figure 2 (Probst). Working model shows that large moving visual surroundings that stimulate the retinal periphery with low spatial frequencies lead first of all to an adequate *object-motion perception* via retino-cortical visual structures. The special characteristics of this stimulus, however, result in a simultaneous stimulation of the central vestibular system in the brainstem, the vestibular nuclei, via both subcortical (eye–vestib. nuclei) and cortical (visual cortex–vestib. nuclei) pathways. The consequence of this gradually increasing ("∫dt") stimulation is twofold: (1) it excites the cortical self-motion perception area and (2) it leads to a concurrent reduction of the input signal of the object-motion perception area via a negative feedback loop. Thus, the interrelation between both centers mediates the perception of a gradual change from pure object-motion perception over a period of apparent rotary body acceleration during which the surroundings seem to decelerate at the same rate up to *exclusive self-motion* with the visual surround appearing to be stationary.

not "relative motion" between objects is perceived and whether or not "absolute motion" with respect to an invisible inertial reference-frame is perceived. The perception of ego motion has a less certain status in Wertheim's model and this is revealing. In this and other models or theories, perception of ego motion can result from motion of the body relative to an inertial reference frame or from motion of the visible surroundings relative to the body even when the body is stationary in the inertial frame (the later is referred to as "vection"). It is assumed by Wertheim and many others that the phenomenology of vection is commensurate with the phenomenology of motion relative to an inertial reference frame. The phenomenology of these two situations may be similar, but rarely, if ever, identical. The primary difference is in the dynamical consequences of motion with respect to an inertial reference frame.

All motion involves speeding up and slowing down, at the very least for starting and stopping, and most motion involves change in direction. The linear and centripetal acceleration in these activities is resisted by linear and centrifugal inertial "forces," respectively. The forces on an accelerating observer stimulate vestibular and somatosensory mechanoreceptors and have consequences for any unrestrained parts of the body. Motion with respect to an inertial reference frame feels (literally) different from motion with respect to the visible surroundings, and these two kinds of motion have different meaning in the context of the ubiquitous postural activity of an active observer (Riccio 1993b). An active observer must work with and against the changing forces on the body to maintain balance and control movement. An inertially stationary observer may also have to adjust the configuration of the body to optimize the visibility of moving surroundings. Thus, both inertial and optical motion have potential consequences though different ones for postural control and for the perception and movement systems that depend on postural control (Riccio 1993a). The meaningful and perceivable consequences of an observer's interaction with the environment, or *affordances* (see J. J. Gibson 1979), differentiate motion of the observer from motion of the surroundings.

The motion of an observer is not a unitary phenomenon, nor is motion of the surroundings; and the experience of these phenomena is not unitary (cf. Neisser 1976, pp. 103–4). The closest they come to being unitary is constant-velocity relative-motion between a fully restrained observer and the visible surroundings. In this special case, motion of the observer and motion of the surroundings are nearly identical phenomena. These are the conditions to which Wertheim's paradigm currently applies, but they are vanishingly rare in nature. The difference between Wertheim's experimental conditions and the natural conditions for a moving observer cannot be overemphasized. Perhaps this is why Wertheim defines a distinction between ego motion and

"self-motion" (see his Note 1). Self-motion naturally requires multimodal perception and multicriterion control of the nested body segments involved in the coordination of posture and locomotion. The multiple criteria generally relate to the speed and trajectory of self-motion, the management of kinetic and potential energy, the changing direction of balance for the various body segments, and stability of the sensory and motor "platforms" (Riccio 1993b; see also J. J. Gibson 1966). There are presumably phenomenological dimensions for each of these criteria. The extension of Wertheim's model, or any model of *self*-motion, to natural situations requires the development of experimental paradigms that are sensitive to the multidimensionality of natural perceiving.

Wertheim's arguments are undermined by his extension of the physiological distinction between retinal and extraretinal signals to the philosophical distinction between "direct perception theory" and "inferential theory." This questionable conceptual leap distracts attention from his research and his model; Wertheim thereby does himself a disservice. He states, for example, that "the theory of direct perception, [which] originated from Gibson (1966; 1979) . . . assumes that the perception of motion derives exclusively from afferent retinal information" (sect. 1, para. 2). This is not an accurate statement about Gibsonian ecological psychology (as articulated, for example, in J. J. Gibson 1966; 1979); Wertheim must accordingly struggle with the obvious inconsistency in his reading of ecological psychology (see his Note 2). Furthermore, it is difficult to imagine any modern philosophical argument that would link direct perception to a single sensory modality.

A cause of Wertheim's confusion is revealed in his belief that "actual research in the tradition of direct perception theory has taken this line of thought and confined itself exclusively to the investigation of optic (or retinal [*sic*]) flow invariants" (Note 2). This statement is not accurate. The fact that J. J. Gibson inspired much more work in visual perception than in nonvisual perception is not relevant to Wertheim's contention about the Gibsonian view of direct perception. It is reasonable to look to the work of others who have been inspired by J. J. Gibson, but in doing so, one should consider the full diversity of such work and one should give special emphasis to the work that is widely recognized for its adherence to Gibsonian first principles (e.g., E. J. Gibson 1991; Shaw et al. 1992; Turvey 1992).

## Ego-centered and environment-centered perceptions of self-movement

John J. Rieser

*rieserjj@ctrvax.vanderbilt.edu; Department of Psychology and Human Development, Vanderbilt University, Nashville, TN 37203*

This is a rich paper. Especially helpful is Wertheim's framing of the inference-based versus direct perception debate and the research summary focused on perceptions of self-movement induced by optokinetic drum stimuli. Especially thought-provoking is the original "reference signal" theory and the method to investigate it with enough precision to build quantitative models. My commentary focuses on perceiving self-movement. In it I point out three ways that the issues raised in the target article could be broadened.

First, the data reviewed are about the ego-centered perception of self-movement; subjects were asked to judge whether they were moving and how fast. What is left out is discussion of environment-centered perception in which subjects are asked to judge how fast or how far they have moved relative to features of their surroundings. Second, the data reviewed relate to the visually induced perception of self-movement and the role played by eye movements in the perception. What is left out is the context of visual perception set by biomechanical information specifying head and trunk position and action. Third, the evidence discussed is concentrated on the perception of simple rotations while viewing the stimulus field created by an optokinetic drum. What is left out is perceiving simple translations and complex paths combining rotations and translations while viewing stimulus fields that provide multiple depth planes. These are each discussed briefly below.

1. Wertheim's discussion of self-movement is mostly focused on findings of the inaccurate feeling of self-turning that occurs in optokinetic drums where the entire visual field rotates around an observer. Those who have experienced it know the resulting illusion of self-turning is compelling. The methods typically involve spinning a textured (usually striped), opaque, cylindrical wall around an observer who is centered within it. Subjects are typically asked to say whether they perceive self-movement or drum movement and to estimate the magnitude of the movement; in addition, the presence of forms of nystagmic eye movements are sometimes recorded.

All these dependent variables assess ego-centered perceiving (self-movement relative to one's earlier positions), not environment-centered perceiving (self-movement relative to features of the surroundings). This emphasis on ego-centered measures is consistent with the traditional methods of vestibular researchers, who wished to understand near-threshold events and the transduction properties of the end organs. To understand adaptive behavior, however, it seem important to investigate environment-centered perceiving in order to know how perceptions of self-movement are integrated with remembered features of the surroundings toward which actions might be directed.

But how might visually based environment-centered perception of self-movement be studied? A direct method would be to ask subjects to judge their position relative to features of their surroundings. The method makes sense as long as the to-be-judged features of the surroundings are occluded from view (e.g., Rieser et al. 1991; Rieser et al. 1994), but it does not make sense in test situations such as an optokinetic drum, because subjects could see the to-be-judged targets and localize them directly with vision whether or not they perceived self-movement. Optokinetic drum test situations could be adapted to assess environment-centered self-movement. This could be engineered in different ways. For example, the drum could be built with windows that could be switched from opaque (adding to the optical flow stimulus used to induce perceived self-movement) to transparent (providing a glimpse of the stable surroundings, in order both to situate the perceiver within the surroundings and to provide a brief view of a target object). In this way optokinetic stimuli might be used to induce self-movement perceived relative to the fixed surroundings.

2. There has been much research and computational theory about whether (and how) optical flow alone can specify the direction and rate of self-movement in situations when the eyes are aimed off-path or when the eyes move. Perceiving self-movement depends on information specifying whether retinal image movement specifies object movement or self-movement, and Wertheim points out that both visual and biomechanical information can serve as a reference signal for eye movements. In addition, it is clear that a perceiver's head and body posture need to be taken into account as well. Visual information specifies the direction of movement of the eye, not the trunk. Perceiving the trunk's direction of motion from dynamic visual information is possible only given information about the direction of the eyes and the head. Wertheim notes that vision alone can specify eye direction relative to the skull, informed by the shape of the visual field as it is bounded by the nose and eye socket.

Similarly, a view of one's shoulder and upper trunk provides information specifying the head-trunk relation. Whereas the eye-in-head relation is ubiquitous and not readily screened from view, the head-trunk relation is easily screened by wearing an

occluding collar or viewing a lighted scene in the dark. Perceivers are not fooled in such situations, however; for example, leftward optical flow with the head turned to the left is perceived as forward self-motion, and so on. The existence of self-movement is visually determined, as is the perceived rate of self-movement. However, the perceived direction of self-movement is based on visual information conditioned by its biomechanical context specifying body posture.

3. Wertheim suggests that the time course for perceiving self-movement based on optokinetic input is slower than that based on vestibular input. The suggestion is based on observations that people very rapidly detect the onset of turns that are well above threshold when they are specified by vestibular input whereas 2 to 10 seconds is needed for an optokinetic stimulus to give rise to perceive self-turning (instead of drum turning). The target article provides an elegant explanation for these different time courses, one that involves visual-vestibular interaction. I wonder whether the different time courses might in part reflect a methodological artifact, however, because although optokinetic drum situations result in relatively slow perceptions of self-movement, stimuli give rise to rapid-onset feelings of self-movement in other situations. A commonplace example involves stopping an automobile on a hill for a traffic signal when a truck pulls alongside, giving rise to the panicked illusion that one's automobile is rolling backwards. A laboratory example is the swaying room used by Lee and others (e.g., Lee & Aronson 1974) to investigate optically induced body sway by infants as well as adults.

Driving situations and the swaying room differ from optokinetic drum situations in many ways, and the different time courses might reflect either different underlying processes or stimulus differences. Consider three possibilities. First, it is well known that the perception of actual self-turning is rapid when it is specified vestibularly, for example, while being turned with eyes closed in a rotating chair. The illusory perception of self-turning necessarily conflicts with a vestibular reference signal consistent with maintaining a fixed heading, and perhaps this conflict results in generically slow onset illusions of self-turning. Unpublished observations in our laboratory seem to contradict this, however, indicating a rapid onset for illusory self-movements that are biomechanically specified. Following earlier observations (Bles 1984; Lackner & DiZio 1985), Daniel Ashmead, Anne Garing, Herbert Pick, and I built a turntable consisting of a fixed T-bar centered within a rotating platform. Subjects stood centered in the turntable, grasping the stationary T-bar to maintain a fixed heading. When the platform rotated, subjects stepped to compensate for its rotation while maintaining their fixed heading. When they closed their eyes, subjects experienced compelling illusions of self-turning relative to their remembered surroundings within 2 seconds after the onset of the biomechanical stimulus, averaging less than one second delay (Rieser et al., submitted). Thus, although the onset of perceived self-turning induced by an optokinetic drum has a relatively slow onset, a slow onset is not generic to all illusory perceptions of self-turning.

As the second possibility, consider whether optical information for self-turning may have a slow onset, whereas optical information for self-translation has a rapid onset. This can be tested easily within an optokinetic drum stimulus situation by simply asking subjects to move from the center of the drum to stand close to the wall. If they are facing the wall, the rotating wall stimulus would specify a rotation combined with a translation – the optical information would be the same as if subjects stepped sideways along the inside circumference of the drum. This is easy to try. Our informal observations are that the onset of perceived rotation combined with translation is also slow, similar to the onset for simple rotation.

Finally, the third possibility is that the slow onset might occur in optokinetic-drum-like stimulus situations where self-movement information is specified mainly at the single depth plane defined by the cylindrical wall of the drum whereas the swaying room and everyday driving situations both involve differentiated depth planes. Perhaps a more "optimal" optico-kinetic stimulus, arranged by providing multiple depth planes or in other ways, would result in faster vection. Whether true or not, this points out a logical limit to understanding perceptual systems by investigating purely artificial stimulus conditions.

## Active and passive head and body movements

Helen E. Ross

*h.e.ross@stirling.ac.uk; Department of Psychology, University of Stirling, Stirling, Scotland FK9 4LA, United Kingdom*

Wertheim must be broadly right in asserting that visual movement perception involves a multicomponent reference signal and that the signal is related to how the eyes move in space rather than in their sockets. He states that reference signals are compound signals that may include any combination of an efference copy, a vestibular and a visual component (sect. 3, para. 4). In his model, knowledge of eye movements within the head is (probably) derived from efference copy, whereas knowledge of head movements is derived from vestibular afferent information caused by head movements. The theory should be generalised, however, to cover all types of information about how the head moves in space, including efference copy about intended head or body movements and afferent information from the vestibular system and tactile and kinaesthetic receptors.

Wertheim does not devote much space to the active/passive distinction. Vestibular information is essentially passive, as it normally signals that an accelerative force is being applied; however, it may or may not be accompanied by efference copy (for corresponding self-initiated movement) and by confirmatory information from the skin, joint, and muscle receptors.

An extreme case of unaccompanied passive vestibular stimulation occurs in divers who suffer from pressure (or "alternobaric") vertigo: high pressure air trapped in the middle ear stimulates the semicircular canals, thus causing a strong sensation of circular vection. The whole body seems to rotate and any object fixed opposite the diver appears to rotate around him at the speed needed to keep pace (Ross 1976). The lack of tactile cues to countermand the sensation of bodily rotation produces very powerful sensations of both visual and bodily motion. Circular vection is also very strong in zero gravity, when the vection is induced visually by a rotating optokinetic drum: the reduction on contradictory tactile cues makes it stronger than under normal gravity (Young & Shelhamer 1990, pp. 523–38). In the case of self-produced rotary and other types of movement, a glowing light fixed opposite a diver's eye appears to move with him, but with a slight lag (Ross 1990, pp. 480–81; Ross & Lennie 1968). This effect may be similar to lags noted by Graybiel and Brown (1951) and Gregory (1958). Lags or leads may occur when there is other information besides vestibular or visual information (e.g., passive tactile or active movement information) that may subtract from or add to the vestibular or visual effect.

Wertheim thinks it unlikely that the gain of the reference signal is close to 1.0 during fast body movements but that the discrepancy fails to cause apparent movement of the visual world, perhaps because the JNDs for the reference signal are too large (sect. 5.4, para. 5). Yet body movement can cause illusory visual movement under ambiguous circumstances. For example, Thompson (1879) described walking up a path to the Clifton suspension bridge and watching it move up and down in relation to his steps when observed intermittently through trees. Hill walkers often report that rocks appear to move in a mist, resembling moving climbers (Ross 1974, pp. 128–29). This

phenomenon may be an example of autokinetic movement without a predictable direction, but the effect usually occurs only when the observer is walking (in which case the rocks appear to move in the direction opposite to the motion of the walker). The excessive apparent movement may be caused by exaggerated apparent distance in a mist, a given angular movement corresponding to a greater linear movement at a greater distance; but it may equally well be caused by an undersized visual contribution to the reference signal in such circumstances. Kinaesthetic information may be inadequate for a veridical reference signal without a normal visual scene component.

## What does linear vection tell us about the optokinetic pathway?

Xavier M. Sauvan

*k364070@czhrzu1a.bitnet; Department of Neurology, University Hospital, CH-8091 Zurich, Switzerland*

Wertheim states that an optokinetic (visual) pattern must be large, must have relatively low spatial frequency characteristics, must not move too fast across the retina, and must remain visible for more than a very brief period. It should be possible to get a more accurate understanding of the optokinetic pathway by studying linear ego motion. This has not been done exhaustively in the target article.

Regarding the low spatial frequency characteristics of an optokinetic stimulation, it has been shown that rectilinear and curvilinear vection is perceived when the spatial frequency of the optokinetic stimulation is less than 1 c/deg (Sauvan & Bonnet 1989). Rectilinear vection corresponds to the sensation of moving in a straight line (Berthoz et al. 1975), and curvilinear vection to the sensation of making a turn. The latter is similar to the actual observer ego motion in a curved path and is induced by asymmetrical visual stimuli, especially with respect to spatial frequency (Sauvan & Bonnet 1993). Moreover, the perceived velocity of curvilinear vection varies (decreases) to spatial frequency (Sauvan & Bonnet 1993). The experienced velocity of curvilinear vection therefore, depends on the spatial characteristics of the optokinetic stimulation, as in circular vection (de Graaf et al. 1990). This should be related to the low spatial bandpass characteristic of gating mechanism defined here by Wertheim.

There is also a zone of maximal sensitivity to the contrast for rectilinear vection centered on the low spatial frequencies and the middle temporal frequencies (Sauvan & Bonnet 1988). It has also been shown that the perceived velocity of curvilinear vection varies (increases) with contrast (Sauvan & Bonnet 1993). Further investigation should indicate how this contrast input is involved in the gating and/or comparator mechanism.

How the optokinetic pathway responds to angular velocity seems to depend on the kind of ego motion. Indeed, linear vection is perceived when the angular velocity of the stimulation is slower than about 40 deg/sec (Sauvan & Bonnet 1989); circular vection can be induced with faster angular velocity up to more than 100 deg/sec (Büttner & Henn 1981).

Linear ego motion can be induced with central or peripheral visual stimulation of small size. Rectilinear vection can be induced with central radially expanding depth patterns as small as 7.5 deg (Andersen & Braunstein 1985). Also, curvilinear vection can be generated with drifting sinewave gratings displayed peripherally behind an aperture subtending a visual angle of 23 deg (Sauvan & Bonnet 1989; 1993). Moreover, it has been shown that circular vection can be obtained with central stimulations as small as 13.5 deg (Howard & Heckmann 1989). Consequently, the optokinetic pathway actually responds not only to small and large but also to central and peripheral optokinetic patterns.

It is known that complex visual information can characterize optokinetic stimulation such as depth cues. Andersen and Braunstein (1985) used radial flow patterns that simulated linear movement through a three-dimensional cloud of dots for inducing rectilinear vection. It was found that apparent depth (introduced by using kinetic occlusion information) influences horizontal linear vection strength (Telford et al. 1992). Therefore, the parietal dorsal medial superior temporal (MSTd) area should be a main part of the optokinetic pathway at the cortical level. Indeed, neurons have been found in this area which code motion in depth (Goodale & Milner 1992), respond to optic flow stimuli (Wurtz & Duffy 1992, or project to the accessory optic system (Maioli et al. 1989).

In short, it is worthwhile and challenging to compare linear ego motion with circular ego motion. A better understanding not only of complex self-motion should thereby be reached but also of visuo-vestibular interactions. Indeed, there is doubt as to the extent to which visuo-otolith and visuo-semicircular canal interaction pathways are segregated. We think it is now important to propose specific and testable hypotheses about the physiological mechanisms and anatomical pathways involved in visuo-vestibular interactions.

ACKNOWLEDGMENT

This work was supported by ESPRIT 6615 Mucom II.

## Ecological efference mediation theory and motion perception during self-motion

Wayne L. Shebilske

*wls9357@zeus.tamu.edu; Department of Psychology, Texas A&M University, College Station, TX 77843-4235*

Wertheim's model of motion perception during self-motion and Shebilske's (1984; 1987a; 1987b; 1990) ecological efference mediation (EEM) theory have mutual implications that highlight important parts of Wertheim's contribution. First, Wertheim makes a strong case that motion perception during self-motion is determined by an interaction of information originating in light patterns and information originating inside the observer. Second, although Wertheim's quantitative model of this interaction stems from inferential theory, the model with respect to relative motion is consistent with direct perception theory, as is the model with respect to absolute motion, though with one exception: The mainstream of direct perception theory has restricted its analysis of visual perception to the investigation of optic flow invariants as documented in Wertheim's Note 2. As documented in the same footnote, however, Gibson's (1966, pp. 283–84) earlier systems approach recognized that sensory information originating in light and sensory information originating in the observer could enter into a unique invariant relation that could directly determine perceptions.

EEM theory extended this reasoning to include special invariant relations between light-based and efference-based information. Wertheim's quantitative model can be thought of in terms of this special kind of invariant relation. The impressive predictive ability of Wertheim's model suggests that direct perception theorists made a false step when they confined their analysis of visual perception to information originating in light patterns. Direct perception theory was on the right track when it encouraged the consideration of invariant relationships between sensory information originating in light and sensory information originating in the observer. Furthermore, the success of Wertheim's model suggests that EEM theory is on the right track.

EEM theory holds, with other ecological theories, that perception-action relationships are shaped by the interaction of an organism with its environment and that operations for encoding sensory information approach optimal efficiency in the environment in which a species evolved. In contrast with the dominant ecological theories, however, it also holds that (a) efference-based information (e.g., from the oculomotor system) interacts with higher-order light-based information (e.g., from optical flow patterns) to determine performance during natural events and that (b) fallibility in both visual and efference-based information function synergistically to shape both the phylogeny and ontogeny of the visual system (Shebilske et al. 1984).

EEM theory is supported by research on minor motor anomalies (MMAs) – dysfunctional states of slight misalignment or misregistration of body part positions (Shebilske 1984; 1987a). These motor states are abnormal in the sense that they are dysfunctional rather than rare, and minor as opposed to major errors (such as paralysis). For example, after people maintain an eccentric direction of gaze for about 30 sec or longer, the biased gaze direction causes a misregistration of eye position, which in turn causes observers to misjudge visual direction. This MMA influences pointing and dart throwing in reduced and full viewing conditions (Shebilske 1977; 1984). Similar MMAs influence pointing to different distances (Shebilske et al. 1984) and baseball batting (Shebilske 1987b). These experiments have induced MMAs during natural events and have measured MMA-illusions during natural events. Thus, the perceived direction of objects in a given light pattern, including those occurring during natural events, depends on the state of oculomotor information; and the visual consequences of change in oculomotor information depend upon the state of visual information. This pattern of results is exactly what would be expected if visual and oculomotor information interact.

A gap in the previous support for EEM theory, however, was the lack of a quantitative model of the interaction. Wertheim's model fills that gap. Accordingly, if MMAs can be found for the estimators of eye velocity in the orbit or head velocity in space, Wertheim's model will be able to predict precisely the influence of these MMAs on motion perception during self-motion in natural events.

Even without MMAs, Wertheim's quantitative predictions should allow precise specifications of boundary conditions that are predicted by EEM theory. The theory distinguishes between conscious perceptions and sensorily guided actions that are not mediated by conscious perceptions. It predicts that practice can bring a sensorily guided skill under the control of an unconscious representation generated by a unique input operation. EEM theory hence rejects the linear models that have guided diverse scientific theories according to which sensorily guided performance is based on perception (cf. Goodale 1988). Advocates of linear models would not deny that perceptions with and without action require different processes, but they would insist on the assumption that the separate sensorimotor modules diverge after perceptual constancy mechanisms yield the stable representations needed for skilled performance. Linear theorists must insist on this assumption because it is a necessary consequence of a model in which action is based on perceptual representations. This rationale could account for the emphasis on appearance as opposed to performance in the constancy literature. Diverging theories, constructs, and operations in constancy research suggest, however, that linear models cannot account for constancy data (Shebilske & Peters, in press), and that progress toward synthesis will be made when unilinear approaches that ask either-or questions about perception are replaced by a multilinear approach that investigates multiple spatial representations for perception and performance in specific integrated sensorimotor modules.

The distinction between inferential and direct is accordingly lost in the domain studied by Wertheim, but the distinction is not resolved for other constancies (Shebilske & Peters, in press), and we cannot assume that it is resolved for all sensorily guided actions during self-motion either.

At this early stage in its development, the ecological efference mediation theory provides only a loose packaging for the constancy data. Quantitative models such as Wertheim's should help articulate the integration and coordination of diverse operations (e.g., unconscious inference or direct perception), by specifying the boundary conditions for the operation of specific multiple representations that emerge as a result of interactions between an organism and its environment.

## The idea that space perception involves more than eye movement signals and the position of the retinal image has come up before

Alexander A. Skavenski

*skavenski@neu.edu; Department of Psychology, Northeastern University, Boston, MA 02115*

The main thrust of Wertheim's target article is that controversy about how visual perception of motion takes place can be resolved if it is realized that global retinal stimulation such as an optokinetic drum actually contributes information that supplements vestibular signals about how the body is moving in space. This suggestion has been raised by numerous investigators of the oculomotor compensation for body motion who noted that the dynamic ranges of the vestibular and optokinetic systems were complementary (for review, see Collewijn 1985; Simpson & Graf 1985; Young 1985). As Wertheim notes, a problem in motion perception is that information about head motion in space must be used in addition to retinal image motion and extraretinal eye movement signals referenced to the head. Earlier, Skavenski (1990) made a similar suggestion, and noted that if both retinal stimulation by large scenes and vestibular signals contributed to the head motion signal this would account for a number of illusory situations often taken to support a limited role for nonretinal signals in space perception.

Skavenski made this suggestion to counter the Post and Leibowitz (1985) suggestion that only the phylogenetically newer pursuit system contributes eye movement signals to motion perception while the older optokinetic system moves the eye without signaling perception of that motion. Post & Leibowitz had explained the Dunker illusion in which a physically stationary spot is seen to move when it is enclosed by a moving frame in the following way. The frame produced retinal image motion that stimulated optokinetic nystagmus (OKN) without extraretinal signs. This OKN was countered by smooth pursuit to keep the eye on the small spot. The pursuit-generated extraretinal signals thus caused perceived motion in the spot. Skavenski noted that it was equally plausible that the frame motion generated information that the head was moving in space. Since the eye did not move in the head and the small spot did not move on the retina, it must be moving with the person. Skavenski also argued that the latter idea is more efficient because it, unlike the Post & Leibowitz suggestion, does not require an explanation of why nature would have taken the trouble to evolve relatively sophisticated compensatory oculomotor systems to output behavior that would confuse the animal about object location and motion.

In sum, Wertheim's target article mainly offers additional evidence to support the idea that information beyond that contained in the retinal image is used in perceiving object motion. In this light, it is a surprise that Wertheim does not also include the results of Hansen (1979), who reported a high

quality of extraretinal eye position signal during smooth pursuit eye movement or of Hansen and Skavenski (1985), who showed that a good quality extraretinal signal (both in space and time) was available for localization during saccades. Both of these reports were concerned with judgments of the position of objects, but other such observations which dealt with position judgments were included (e.g., Matin et al. 1969).

## "Sensory" reference frames and the information for self-motion versus object motion

Thomas A. Stoffregen

*stoffreg@ucbeh.san.uc.edu; Department of Psychology, University of Cincinnati, Cincinnati, OH 45221-0376*

Wertheim describes a situation in which passive, restrained observer's experience themselves as rotating while sitting in a chair that is stationary with respect to a rotating optokinetic drum. He describes their experience as "an illusory sensation of ego rotation." He does not explain why this percept is interpreted as an illusion. There is, however, a clear implication that "illusory" self-motion is not real, that it constitutes a perceptual error. There is an implied assumption (common in the self-motion literature) that motion of the self relative to the substratum and gravity is "real," whereas motion relative to the optical (or acoustic) surround is "illusory." Wertheim defines motion relative to earth gravity as "absolute" motion. Similar interpretations are made in the vestibular community in the case of the so-called oculogravic and audiogravic illusions (Graybiel 1952; Graybiel & Niven 1951). In the literature on perceived orientation, gravity is assumed to be primary (Schöne 1984). In none of these cases is any a priori (theoretical) justification offered for the primacy of gravity. In all of these cases, subjective reports that deviate from gravity are illusory only if we assume that observers are basing their perception of self-motion on earth gravity.

In fact, the distinction between absolute and relative motion is inaccurate and misleading; there is nothing absolute about motion relative to earth's gravity, as Einstein showed (Einstein & Infeld 1938). Moreover, there is reason to believe that gravity is not a fundamental referent for animals. Recent theory and experimentation have shown that gravity is not the primary referent for the perception or control of orientation. Rather, orientation is perceived and controlled with respect to the direction of balance, which often differs from the direction of gravity (Riccio et al. 1992; Riccio & Stoffregen 1990; Stoffregen & Riccio 1988). Thus, the assumption that orientation and motion are perceived and controlled relative to gravity is not only theoretically problematic but contradicted by recent data.

The discussion of absolute and relative motion has importance for Wertheim beyond the interpretation of subjective reports. It supports one of his major conclusions: "Direct perception theory is concerned with the perception of relative motion and inferential theory with the perception of absolute motion." This amounts to a claim that direct perception addresses perception of motion relative to objects, whereas inferential theory addresses perception of motion relative to the earth. I doubt whether either camp would accept this characterization. Shorn of its incorrect terminology the claim becomes both less interesting and less defensible.

Perception can be understood as a form of measurement and, like other forms of measurement, it must be scaled relative to a frame of reference (Bingham 1987). Wertheim's analysis is consistent with traditional assumptions about scales or referents for perception. One assumption of the traditional view is that there is a separate "reference frame" for each perceptual system. The hypothetical referents are defined external to the animal, in terms of the types of energy that stimulate different perceptual systems. For example, the referent for the vestibular system is believed to be specific force, for the auditory system, airborne mechanical vibrations, and for vision anisotropic optical structure. An important property of these sense-specific reference frames is that they are mutually exclusive. This leads to the conclusion that each perceptual system can indicate object or self-motion independently and that they often indicate different (incompatible) motion relationships between the self and all or part of the environment (cf. Stoffregen & Riccio 1988). These assumptions underlie the common belief that an animal can have a perception of orientation or motion that is peculiar to a particular perceptual system, as when Wertheim, without any prelude or subsequent justification, defines his area of study as "the *visual* percept" of "motion or stationarity," (emphasis added).[1] The assertion that we can have a visual perception of self-motion or stasis (physical events that stimulate multiple perceptual systems; Stoffregen 1990) entails strong assumptions about relations among perceptual systems. These assumptions are rarely explicitly presented or defended. The belief in sense-specific perception of self-motion is common in the literature on vection (e.g., Sauvan & Bonnet 1993; Warren & Kurtz 1992), but it is not universal (e.g., Stoffregen & Riccio 1990; 1991; cf. DiZio & Lackner 1986).

Recent theoretical developments have questioned traditional assumptions about sensory reference frames. Stoffregen and Riccio (1988; 1991) have argued that the perceptual systems are not sensitive to extrinsic reference frames defined in terms of stimulus energy. Instead, the referents for perceptual systems are kinematic events and states of the interaction between the animal and the environment (and, hence, intrinsic rather than extrinsic; cf. Mark 1987). For example, rather than being sensitive to specific force, as has traditionally been assumed, the vestibular system appears to be sensitive to kinematics that have consequences for the control of posture, such as dynamic orientation relative to the direction of balance (Riccio et al. 1992). The perception of these kinematics can be achieved without prior or concurrent sensitivity to specific force. Information about dynamic orientation relative to the direction of balance is available to other perceptual systems in addition to the vestibule. For example, postural sway gives rise to optical flow and the parameters of this optical flow (e.g., patterns of changing acceleration) are specific to dynamic orientation relative to the direction of balance. Sensitivity to this information does not depend on sensitivity to anisotropic optical structure. These examples illustrate the argument that orientation can be perceived without appeal to sense-specific reference frames (Stoffregen & Riccio 1988; 1991). These postural kinematics are determined by dynamical properties of the body in interaction with the environment rather than by the dynamics of receptors or "estimators" that are central to Wertheim's approach.

How does this analysis apply to the perception of self-motion (translation) and its differentiation from object motion? Stoffregen and Riccio (1990; Stoffregen 1985; 1986) analyzed kinematic stimulation that is created by motion of the self (postural sway) and motion of objects (looming, or impending collision). Rather than relying on subjective reports of self- and object motion (which have the problems of interpretation discussed above), this research relies on adaptive behavioral responses to different events (compensatory sway, dodging in avoidance). Our data and analysis of the literature indicate that self- and object motion cannot be differentiated reliably on the basis of stimulus variables such as optical velocity and the angular extent of stimulation, the kinds of variables that are central to Wertheim's model. Optical velocity does not provide reliable information about parameters of physical motion; it can function as no more than a probabalistic cue (Stoffregen 1986). Its use as a cue for estimating these parameters would be unnecessary if the differing dynamics of self- and object motion were specified directly in perceptual stimulation. Stoffregen

and Riccio (1990) argued that this is the case: that self-motion and object motion, being different events, give rise to different patterns of sensory stimulation, and that these patterns of stimulation are lawfully (and uniquely) related to the underlying events, so that they provide information for the events.

Hence, the differing physical dynamics of object motion and motion of the body are available in perceptual stimulation. To the extent that perceptual systems are sensitive to these dynamical differences, object and self-motion can be differentiated directly, that is, without the need for internally generated (mediated) estimates. This is similar to Wertheim's characterization of the direct perception view as one in which veridical perception arises exclusively from afferent information. However, Wertheim's understanding of afferent information is in terms of sense-specific probabilistic dues in retinal flow, while in my view afferent information is in terms of amodal or cross-modal patterns of stimulation that are deterministically specific to object-motion and self-motion events.

NOTE

1. The power of this assumption can be extraordinary. Wertheim retains the terminology of "the visual perception of self-motion" despite the fact that he discusses nonvisual (vestibular) influences on the percept. Apparently, vestibular stimulation influences the perception of self-motion without disturbing or compromising its "visual" character.

## Does the reference signal cancel visual field motion?

Arnold E. Stoper

*astoper@seq.csuhayward.edu; Psychology Department, California State University, Hayward, CA 94542*

Wertheim's attempt to explain a wide variety of motion phenomena in terms of a simple reference signal that adds vectorially to the retinal signal has considerable appeal. However, I believe that a serious problem in his presentation is a failure to distinguish between "visual field" motion and "visual world" motion (Gibson 1950).

In the original Filehne (1922) illusion, the observer tracks a moving point in normal room illumination and observes retrograde motion of the background. As I described it (Stoper 1967; 1973), this is a paradoxical motion – the background moves, but it doesn't go anywhere. This can be characterized as a motion of the "visual field" rather than "visual world" (see Mack 1978 and Rock 1977, for sophisticated discussions of this distinction). According to Gibson, one must take a "pictorial attitude," that is, view the world as a picture, in order to experience the visual field, and in this case, visual field motion. On a first demonstration, only about 50% of observers will spontaneously report such motion, but with appropriate guidance (and patience!) I have found that nearly everyone will report it. All observers will agree that the background "smears" somewhat when one is following a moving target. Once attention is called to this "smearing," observers will agree that the image of the background does not just smear, it also moves, in some sense. At the same time, in another sense, the background remains stable, that is, there is no motion of the "visual world." Readers can easily demonstrate these phenomena for themselves simply by tracking a pencil point moving over a page of text.

Wertheim concluded (as did Mack 1978) that the Filehne illusion is caused by an underregistration of pursuit velocity, hence a reference signal too small to cancel the background retinal signal. However, the Filehne illusion as measured by Wertheim seems to differ in character from the one originally noticed by Filehne. Wertheim measured the illusion by nulling until the observer reported "stability." For a large, continuously present background (sect. 5.1 and Fig. 4), he finds no Filehne illusion at all, but these are precisely the conditions in which Filehne first noticed his illusion, and the ones I describe above. If Wertheim's observers were nulling visual world motion instead of visual field motion they would, of course, have seen the large background stable with no nulling necessary, and this would account for the lack of any measured illusion; it seems likely that this was in fact the case. Even though the large background was reported stable, it would still have had visual field motion. If this account is correct, then the illusion Wertheim measured was something other than the original Filehne illusion. Wertheim's experiments thus fail to show that the reference signal has any capacity at all to cancel retinal image motion once it results in visual field motion and his explanation of the Filehne illusion in terms of the undersized reference signal is not applicable to the illusion originally described by Filehne.

The theoretical importance of the visual field motion of the background seen during pursuit arises from its absence in the case of a saccade. In a normally illuminated environment, there is no hint of the field motion seen during pursuit. I showed (Stoper 1967; 1973) that this difference in appearance is not due to the smooth "dragging" of the image during pursuit as opposed to the "hopping" of the image during the saccade, as had been suggested by Gregory (1958). I believe that this difference between pursuit and saccade presents problems for Wertheim's assumption (Note 7) that the same reference signal and essentially the same JND (just noticeable difference) raising process are responsible for both stability during pursuit and stability during the saccade. My conclusion was, and remains, that some central process must take place during the saccade which does not occur during pursuit. One candidate for such a "saccade only" central process is the simple suppression of the retinal signal, as suggested by Wallach and Lewis (1965).

In fact, it is quite possible, as I suggested (Stoper 1967), that once generated the retinal signal is never cancelled by vectorial addition of a reference signal, no matter how large. If the retinal signal does not add vectorially to the retinal image signal, how do they interact? One possibility is illustrated by the famous "motion analysis" demonstration of Johannson (1982), consisting of a wheel rolling in the dark. If a point on the rim is illuminated, it is seen to describe a cycloid, which is its actual trajectory in space. If a point at the center of the wheel is illuminated, it will have only linear translational motion. If both points are illuminated and the point at the center is pursued, the reference signal will be translational and the retinal signal of the rim point will be circular. If these two signals were to add vectorially, the "true" motion of the rim point in space would be seen; it would be the sum of the circular and translational components, that is, a cycloidal motion. As is well known, this is not the percept at all. Instead, what is seen is a wheel rolling – the point at the rim is seen to describe a circular motion about the center and the entire wheel is seen to translate. The point at the rim has both circular and translational components – it goes around the center, but it is part of the wheel and translates with the wheel – but these two components do not add vectorially. They remain independent, and can be perceived separately.

Returning to the explanation of stability, perhaps it really is visual world motion that is important, and as Gibson would say, visual field motion is just an artifact produced by unnatural viewing conditions. Can stability of the visual world be brought about by Wertheim's proposed mechanism; that is, by vectorially adding a reference signal to the retinal signal? This is at least plausible in the cases of eye movement or of purely rotational movement of the observer, when the image motion has one constant velocity at any one time over the entire retina. It would be at least theoretically possible for a single valued reference signal to "cancel" this retinal motion by subtraction. However, if there is any translational component to the motion of the observer, the optic flow and the corresponding image flow on the retina become much more complex, with a wide range of

velocities at any one instant in time. The image movement would be a function of both distance and direction to each particular surface element in the world (see, e.g., Gibson 1954). Any reference signal which is subtracted from the retinal signal must have a range of values similar to that of the retinal signal at one instant of time in order to produce stability. A single valued reference signal such as the one proposed by Wertheim cannot fail to produce cancellation of the retinal signal and stability of the visual world during translational observer movement.

Finally, consider the appearance of the visual field in the case of translational observer movement. Despite the apparent stability of the visual world, the "optic flow" of the field remains visible, that is, nearby objects appear to move back faster than distant ones. This is, of course, visual field motion, and can be taken as showing that the retinal image motion is not cancelled. A better description of what happens to this field motion would seem to be that of Duncker (1929): the field motion is attributed to the motion of the observer rather than motion of the world.

## Spatial motion perception requires the perception of distance

Michael Swanston

*bstmts@cluster.cc.dct.ac.uk; Dundee Institute of Technology, Bell Street, Dundee DD1 1HG, Scotland*

Wertheim presents a model of the process by which perceived motion is derived from retinal and extraretinal information. The model provides an account of a wide range of motion phenomena, including veridical and nonveridical perceptions, the motion aftereffect, induced movement, and vection. Most crucially, the model is expressed in terms of motion in space (of the eyes, of the observer, and of stimuli), and thus appears to provide a complete system of explanation. This impression may be misleading, however, despite the use of the subscript "space" to indicate the frame of reference for velocity terms in various formulae.

The problem is apparent from a consideration of the situation discussed in section 5.4. Wertheim describes the case of a train driver looking forward as the train accelerates from rest. Initially, the vestibular system provides information for ego motion. This is said to provide a reference signal which will be "approximately equal to the retinal signal evoked by the moving image of the visual world," and the driver's world is therefore perceptually stable, but this is clearly unworkable. The reference signal from the vestibular system defines an acceleration along the *z*-axis. The retinal motion information is multidirectional, and variable in magnitude across the retina. There is no single retinal signal from which the vestibular reference signal can be subtracted. What is required for this is a transformation of the retinal information to express the spatial motion of objects relative to the observer. Such a signal, which is egocentric rather than retinocentric, could be combined additively with a reference signal for self-motion to give a perceptually stable visual world. Gogel and Tietz (1992) have recently described experiments which show that the perceived sagittal motion of objects during self-motion is determined by the product of retinal motion and perceived distance and not just by the subtraction of perceived self-motion from retinal motion.

In general, the operation of subtracting a reference signal from a sensory signal can be considered a conversion of the frame of reference with respect to which the sensory signal is expressed (Swanston et al. 1987). Thus, subtraction of a reference signal representing movements of the eyes with respect to the head from a sensory signal for image motion over the retinae yields a value which represents image motion with respect to the head. Such a signal will be altered by movements of the head in space and by perceived distance (Swanston & Wade 1988; 1992a). This may not seem to be the case, however, if an observer is stationary, perceived distance is constant, and retinal motion is entirely due to object or eye movements. Under such conditions, which obtain in the experiments described by Wertheim in section 5.1, only one frame of reference (egocentric) will appear to be operating and the influence of perceived distance will not be apparent.

Even if physical distance is held constant, the influence of perceived distance on perceived motion can readily be demonstrated. Gogel's extensive work on this issue (see Gogel 1990 for an overview and synthesis of his theoretical and empirical investigations) can hardly be ignored by any theory of motion perception that aims to describe object and observer motions in three dimensions. Wertheim does not refer to this work, which cannot be accounted for by the model presented in the target article. Consider the case of an observer making lateral head movements while fixating an object at a fixed distance. There will be no motion on the retina, and the rotation of the eyes in the head will be such as to cancel out the effects of the lateral head movements on the retinal position of the object. If, as Wertheim makes clear, registered values for image motion, eye movements, and head movements are veridical, then the object will appear to be stationary. However, this will not be so if there is any mismatch between the object's physical and perceived distance. If the object is seen to be nearer than its physical distance, it will appear to be moving with the head. If it is seen as farther away, it will appear to move against the head. These effects are indistinguishable from, and additive with, real motions, and they occur equally in a structured visual field (Gogel 1982; Gogel et al. 1985). Thus, a physical motion can be perceptually increased, cancelled, or reversed, solely by the manipulation of perceived distance. In Wertheim's model, only a single value of the reference signal representing eye movement "in space" is available, regardless of the perceived distance of the stimulus object, and thus no influence of perceived distance on perceived motion would be predicted.

A similar problem arises when there is relative motion between visible objects. In section 5.3, Wertheim states that such motion signals are independent of eye movements; this is only the case if the head is stationary, however, and the eye movements are rotations in the head (Wade & Swanston 1987). If the head moves, and the retinae move in space, then the objects' physical distances will determine their relative retinal motions and give rise to motion parallax. The perceived outcome is then some combination of relative depth (Rogers & Graham 1979) and relative motion, the balance between the two being influenced by the perceived egocentric distance of the objects (Ono et al. 1986). We have discussed the role of relative motion information in induced movement (Wade & Swanston 1987) and the motion aftereffect (Swanston & Wade 1992b) in the context of a model of motion perception which explicitly incorporates perceived distance and covers much of the same ground as that proposed here by Wertheim. In addition, we have argued that the issues of binocular combination and visual direction from the cyclopean eye should be addressed by any model of spatial motion perception (Swanston et al. 1990).

From our point of view, Wertheim has provided a detailed analysis of the early stages in motion perception, concerning the derivation of an egocentric signal from motions of and over a retina, where the former are due to equal rotations of both eyes. Much valuable information regarding the psychophysics of this process is summarised, but the model itself does not appear to be able to support a truly spatial account of motion perception.

## A cortical substrate for motion perception during self-motion

Peter Thier,[a] Roger G. Erickson[b] and Johannes Dichgans[a]

*[a]thier@mailserv.zdv.uni-tuebingen.de; [a]Department of Neurology, University of Tübingen, 72076 Tübingen, Germany and [b]Laboratory of Neurophysiology, National Institute of Mental Health, Poolesville, MD 20837*

Wertheim's target article provides a model of motion perception during self-motion that rests on the assumption that percepts of object motion are derived by comparing retinal slip of the object image with an internal reference signal. The reference signal suggested by Wertheim describes how the eyes move in space and relies on both visual and nonvisual information. His theory is largely based on psychophysical observations. Quite understandably, the question of where and how in the brain the mechanism he proposes might be implemented is accordingly not the focus of his attention. The results of our recent experiments using monkeys shed some light on this as well as on some of the other issues discussed by Wertheim.

In our experiments, single-unit electrophysiological recordings from cortical visual areas were used to compare the responses of visual neurons to slow retinal image motion caused either by object motion or pursuit eye movement. Three of our results are directly relevant to the target article. First, directionally selective cortical visual neurons that respond selectively only to externally induced visual motion were found (we named these "passive-only" cells). Second, these neurons were localized in a single visual area (the dorsal part of medial superior temporal area MST). The overwhelming majority of neurons sampled from other cortical visual areas (including medial temporal area, MT) could not discriminate the source of retinal image motion. Third, our results clearly showed that both visual and apparently nonvisual signals contribute to the passive-only properties of cells in MST. We therefore support Wertheim's suggestion that both visual and nonvisual signals contribute to evaluation of object- and self-motion.

The passive-only neurons found in MST respond vigorously to retinal image motion resulting from object motion ("passive" image slip) but weakly or not at all if the same retinal image slip results from smooth pursuit eye movements ("active" image slip). Assuming that monkeys perceive a stable visual world during ego motion much as we do, these MST passive-only cells are currently the only known candidate substrate for spatial stability during ego motion. Our finding that most other parts of the monkey visual system simply ignore the source of retinal image motion (object motion or ego motion) simply reaffirms the fact that visual motion information can contribute to several aspects of visual perception in addition to spatial orientation and ego motion. This is demonstrated clearly by our finding that most neurons in area MT, the prototypical cortical motion-processing area, do not discriminate object motion from ego motion (Erickson & Thier 1991). The same inability to discriminate object motion and ego motion also characterizes earlier stages of visual-motion processing. As early as 1969, Wurtz had shown that visual neurons in area 17, the primary visual cortex, could not discriminate high velocity retinal image motion caused by saccadic eye movements from those caused by object motion. Although one might argue that the neurons studied by Wurtz (1969) were mostly nondirectional visual neurons, that is, neurons not involved in the analysis of visual motion, the absence of passive-only properties in directionally selective MT neurons indicates that this information is not available to the direction-specific V1 neurons projecting to MT, a fact recently confirmed in our laboratory (Ilg & Thier 1993).

Our analysis of the mechanisms contributing to the passive-only property is directly relevant to evaluating the competing theories of motion perception discussed by Wertheim. Proponents of the direct perception theory have suggested that the ability to discriminate object- and self-motion might be derived from retinal information alone. One visual cue that might be used for this purpose is the coherent displacement of the entire visual background that occurs during eye movements. The work of Tanaka et al. (1986) has shown that full-field visual motion can activate otherwise silent directionally specific inhibitory receptive field surrounds in some MST neurons, thus suppressing responses that would otherwise result from the motion of contours across the receptive field center if contours outside the field were moving in the same direction. Our experiments showed that some of the passive-only cells did indeed appear to depend on this strictly visual mechanism. When possible, such cells could rely solely upon easily calibrated relative-motion cues to discriminate the visual conditions that usually occur when self-motion occurs in an environment with ample background visual contours.

These cells lose their passive-only property, however, during eye movements against a background that is relatively dark or featureless except for the stimulus crossing the receptive field. In this instance it is necessary to rely upon other, possibly nonvisual reference signals. Our results demonstrated that some of the passive-only cells did not have center-surround visual interactions of the type described by Tanaka et al. and instead relied upon apparently nonvisual inputs to cancel the normal effect of retinal image motion across the cells' receptive field (Erickson & Thier 1992). We therefore agree with Wertheim that both visual and nonvisual mechanisms are necessary to provide an accurate sense of ego motion under the entire range of natural conditions. Our analysis has not yet addressed the question of whether passive-only cells that use nonvisual reference signals are able to enhance their selectivity for object motion when additional visual information is available to help discriminate ego motion. At any rate, our results show that, at least at the population level, the integration of both visual and nonvisual signals referencing eye movements is present.

While the contribution of both visual and nonvisual components to the reference signal is in general agreement with Wertheim's model, an apparent discrepancy between the single-unit data and his model should be mentioned. According to Wertheim, the visual component of the reference signal basically corresponds to an optokinetic signal which, despite a short onset latency, is characterized by slow buildup and decay involving a reinterpretation of the source of the perceived visual motion, that is, whether the retinal image motion represents target or self-motion. On the other hand, the discrimination of self-induced retinal image slip represented in the discharge of MST passive-only cells was always prompt and stable. It would be interesting to determine whether visual conditions provoking optokinetic responses larger than those usually prevailing in our experiments also induce gradual changes in the responses of passive-only cells. We do not, however, feel that lack of this direct comparison calls into question our attempt to relate Wertheim's unified theory of motion perception to a cortical substrate.

ACKNOWLEDGMENT

This work is supported by DFG Grants SFB 307-A1 and KFG Zr 1/9-1.

## Two straw men stay silent when asked about the "direct" versus "inferential" controversy

J. R. Tresilian

*james.tresilian@mrc-apu.cam.ac.uk; Medical Research Council (MRC) Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, United Kingdom*

According to the "inferential" theory, visual information about image motion is compared to extravisual information about eye movement to determine whether it is motion of the eyes or of the environment that is giving rise to the image motion. This

simple theory is so obviously wrong that it hardly merits mentioning. It predicts that if the eyes are stationary in the head as the head rotates, the resulting image motion will be interpreted as motion of the environment, yet everyone knows that this does not happen. As pointed out in the target article, it is motion of the eyes with respect to the fixed environment that should be compared with image motion if the "inferential theory" is to work. Thus, the inferential theory is the hypothesis that image motion is interpreted as being due to eye movement or to environmental motion by a comparison with information about how the eyes are moving relative to the environment.

This inferential theory is most bizarre. Despite all the theoretical work which demonstrates that global image motion specifies how the eyes are moving relative to the environment (e.g., Bruss & Horn 1983; Gibson et al. 1955; Koenderink & van Doorn 1981; 1987; Lee 1974; Longuet-Higgins & Prazdny 1981), the inferential theory asserts a priori that this information is not used. Instead of using global image motions as information about how the eyes are moving, the brain has to interpret these image motions as due to eye movement or environmental motion by actually computing how the eyes are moving, using articular, vestibular, and efference copy information. The inferential theory makes this most extraordinary assertion without any justification. Hardly surprisingly, it is wrong; it has been known for a long time that global image flows give rise to the perception of ego motion.

The "direct theory" described in Wertheim's target article states that image motion is interpeted as being due to eye movement or environmental motion directly, that is, without any use of extravisual information. The idea is simple: global image flows specify movements of the eyes with respect to the environment and are thus perceived as such. This direct theory is also obviously wrong. It predicts that in the absence of global image flows the eyes will not be perceived as moving when, for example, pursuing a moving target. That this is, in general, false is amply demonstrated by people's ability to accurately perceive the motion of self-luminous objects in the dark (e.g., Rosengren et al. 1988), something many people have experienced and can easily verify.

The major distinction between the two theories is not, as asserted in the target article, that the indirect theory maintains that "information about how eyes move (in space) is always necessary to perceive object motion or stationarity" (sect. 3, para. 8) while the direct theory denies this. Both theories take this to be a fact. The two theories differ only in what sensory system they consider to be the source of information about eye movement in space (Table 1).

The two "theories" must be straw men, because they make unjustified, a priori assertions which are obviously wrong. Moreover, they do not bear on the "direct" versus "inferential" controversy which they misrepresent by reducing it to a debate over whether motion perception involves only retinal (visual) information or both retinal and extraretinal information. Regardless of Gibson's opinion about the source(s) of information involved in visual motion perception, his notion of directness centres on the idea that stimulus information (across all senses: there exist what Gibson called "intermodal invariants") is *sufficient* for veridical perception – nothing need be added. Associated with this is the notion that the processes of information extraction cannot be meaningfully decomposed into a sequence of subprocesses (discussed extensively in Ullman 1980). Theories which contrast with direct perception are those which propose either that stimulus information is not sufficient or that information extraction proceeds in a series of stages. Neither contrast is made in the target article, which cannot, therefore, be considered relevant to the controversy. There is some mention of a claim that extraretinal information about eye movement is insufficient but no reasons are given – though one can see that there will typically be no extraretinal information at all about translation of the eyes through space when moving at constant velocity; but this case is not discussed.

There appear to be a variety of hypotheses about the source(s) of information about how the eye(s) move in space that might be proposed to account for motion and no-motion perception. Six are listed in Table 1. The only hypotheses worth considering are "dual mode," MDM (modified dual mode), TA (hypothesis presented in the target article), and MTA (modified target article hypothesis) because these propose that all available sources of information are used (though not necessarily all at the same time). Dual mode theory is the hypothesis that when visual

Table 1 (Tresilian). *Six hypotheses about the source(s) of information used in generating percepts of motion and stationarity*

| *Direct* | *Inferential* | *Dual mode* | *MDM*[a] | *TA*[b] | *MTA*[c] |
|---|---|---|---|---|---|
| Visual only | Extravisual only | When "large-field" image motion is present only visual information is used. Extravisual otherwise | When "whole-field" image motion is present only visual information is used. When less than whole field image motion is present extravisual information is used as well. When no visual field information is present only extravisual information is used. | Both visual and extravisual may be used when available. Not clear exactly what influences whether a source is used and what contribution it makes to perception. Perhaps the extent of a wide-field image motion influences its contribution. | Both visual and extravisual may be used when available. Whether a source is used and the contribution it makes to perception depend not only on whether it is available but also on the task and the context. Thus, sometimes only visual information may be used despite extravisual information being available (and vice versa). |

[a]MDM is a modified version of the dual mode hypothesis.
[b]TA is the hypothesis presented in the target article.
[c]MTA is a modified version of TA which supposes that the use and contribution of an information source depends on several factors.

information about eye movement is available it is used and when it is not, other information is used. I doubt whether this theory can really be criticised on the grounds that motion is perceived when pressing on the eyeball (sect. 5.1, para. 4). When I press on my eyeball I see motion, but if someone asked me if I saw the world move I would say "no, I saw the image move" – I suppose that with a bit of effort you could persuade someone to *say* that they saw the world move, but this does not prove anything. What is needed is an objective measure of world-motion (as opposed to ego-motion) perception following eye-ball pressing.

I do not see that the material reviewed in the target article allows one to distinguish between the MDM, TA, and MTA theories as described in Table 1 (the TA theory is rather vague about what factors influence the sources of the eye-movement information used and their contribution to perception). The interesting empirical questions are not concerned with the distinction between direct and inferential perception (which is essentially philosophical). Some interesting questions are, "What sources of information are used? If more than one source can be used are these sources always used when they are available or does their use depend on the task and stimulus conditions? When two or more sources are used together, how are they combined? Are they differentially weighted according to task and context? Research described in the target article addresses some of these questions; I think it is unfortunate that it has been submerged in a battle between two straw men.

## Space as reference signal? Elaborate it in depth!

Boris M. Velichkovsky[a] and A. H. C. Van der Heijden[b]

*[a]velich@hrz.uni-bielefeld.de and [b]heijden@rulfsw.leidenuniv.nl; [a]Center for Interdisciplinary Research, University of Bielefeld, W-4800 Bielefeld, Germany; Faculty of Psychology, Moscow State University, 103009 Moscow, Russia and [b]Department of Psychology, Leiden University, 2333 AK Leiden, The Netherlands*

Early psychologists/philosophers distinguished two different problems in the field of visual spatial perception: the perception of two-dimensional position and size on the one hand and the perception of depth on the other (see Kaufman 1974, pp. 213 and 322). The perception of size and of two-dimensional position were not regarded as problematic issues. An image of an object covers a portion of the retina at a particular location. It can therefore be argued that the location on the retina determines the object's perceived location and that the amount of retinal surface covered determines the size of the object. In contrast, depth perception was indeed regarded as problematic. Because the retinal surface is two-dimensional, depth perception cannot be related in the same obvious way to one or another aspect of the retinal image. Thus, the perception of depth required a different scientific treatment from the perception of size and two-dimensional position. For this problem, two questions had to be answered. The first was: What cues are used in depth perception? The second: How are these cues used in producing veridical depth perception? The first question brought discoveries such as "pictorial cues" (detail perspective, image size, relative brightness, interposition, etc.) and "physiological cues" (accommodation, convergence, retinal disparity, etc.). The second question is conventionally answered by a family of models using weighted linear combinations of various depth cues (Cutting et al. 1992a), although, for instance, a fuzzy logical (Bayesian) model can handle the empirical data equally well (Massaro & Cohen 1993).

Nowadays there is sufficient reason to ask whether this theoretical encapsulation – one type of explanation for size and position perception and another, more liberal, type of explanation for depth perception – is forced upon us by external nature. One can doubt (as Gibson 1979 and Köhler 1947 did) whether retinal size and retinal position are indeed the fundamental givens that theories about perception must start with. Possibly other retinal and "extraretinal" parameters are important. One can even take the view that for visual perception what is pictured on the back of the retina is of no importance (see Gibson 1979, for compelling arguments). Then one is clearly in the position to argue that exactly the same two scientific problems must be solved for size and position perception as for depth perception: What cues are used and how are they used in producing veridical size perception and veridical position perception?

Wertheim's theory of motion perception during self-motion makes these questions even more pressing. On the one hand, Wertheim takes an important step forward in claiming that it is the representation of space that plays the role of final frame of reference ("reference signal") for motion and self-motion perception. On the other hand, he does not seem to recognize all the consequences of this conclusion. He explicitly follows only one line of analysis derived from the conclusion, namely, the idea of intermodal, visual-vestibular contribution to the motion and velocity perception.[1] In general, however, his theory still belongs to the traditional "size and position" category and not in the "depth perception" class.

In Wertheim's theory, information processing for motion perception and for velocity perception starts with one retinal given, a "retinal signal" reflecting image motion on the retina, and only results in a veridical perceptual motion and velocity interpretation after subtraction of the "reference signal," reflecting artefactual image motion due to movements of the observer. In a similar vein, in orthodox size and position theories the size and position of the retinal image was taken as the starting code: code size is taken as size code and code position is taken as position code. For motion and velocity such an easy equation is not possible, of course, and a more abstract code must be looked for. This code has to be derived from, or calculated on the basis of, the retinal information. In Wertheim's theory the retinal signal, $V_{ret}$, is this more abstract code. But again the magnitude of the signal in general corresponds to the retinal image velocity (even if coding errors can occur in this velocity as when, for instance, retinal signals underregister image velocity; see sect. 6.4).

In other respects, however, Wertheim's theory must be qualitatively more sophisticated than other members of this category. Thus, in most orthodox size and position theories the problem of how to undo the raw retinal code from its artefactual components is solved in one step: veridical size equals retinal size times distance and veridical position equals retinal position minus eye position. For absolute velocity perception such a simple operation will not do. The problem is not that the reference signal is a compound signal which includes an efference copy, a vestibular component and a visual component (see sect. 3) that all have to be "subtracted" from a retinal velocity. Rather, contrary to what Wertheim's theory seems to suggest, veridical absolute velocity perception also requires distance to be taken into account. In terms of the classical approach, the equation for absolute perceived velocity has to be as follows:

$$\text{Distance } x \text{ (total retinal velocity} - \text{artefactual retinal velocity)}$$

Indeed, there are rather dramatic demonstrations of the importance of depth information exactly in the domain of Wertheim's research. One of us (Velichkovsky 1982) has, for example, described the following effect of an excessive eye vergence in the situation of induced motion perception and vection. If the convergence is strong enough to warrant the binocular fusion of vertical drum stripes shifted one period then the stripes suddenly jump closer to the observer, with their width diminishing accordingly. Of course, all this is to be expected on the basis of the common size-distance relationship

considerations. But at the same time and to the same extent the perception of motion is changing: both perceived motion of fixated objects and self-motion become much slower although the physical stimulation (retinal motion) and hypothetical subtractive components of Wertheim's theory remain effectively unchanged.[2] Together with well-known facts about the dependence of thresholds and perceived parameters of real (Kano 1970; Tyler & Foley 1974) as well as apparent (Corbin 1942; Larsen et al. 1983) motion on the trajectory in 3-D space these observations – paradoxically – support the main idea of Wertheim's work ("space as the reference for visual motion"), but cast doubts on how he tries to elaborate this insight.

There are some additional complexities connected with the organization of perceptual space. Not only is perceptual space three-dimensional and intermodal (or amodal). In contrast to the space of Newtonian physics, it is also anisotropic and nonhomogeneous, so our spatial orientation relies on a multitude of simultaneously presented frames of reference. In fact, a fairly common experience in induced motion situations is that several frames of reference are simultaneously at work: strong vection can coexist with equally strong motion perception of the "optokinetic" background. A rather similar effect of multiple localization of an object is also known from investigations of spatial perception during saccadic eye movements (Bischof & Kramer 1969). The Gestalt school – being perhaps more sensitive to peculiar aspects of phenomenal experience – was very concerned with describing and theoretically interpreting such phenomena (see Duncker 1929). According to one of these interpretations, proposed by Metzger (1941), spatial frames of reference can build functional hierarchies restricting exactly the applicability of vector analysis to the problems of visual motion perception (cf. sect. 6.5).

Orthodox size and position theories were convincing largely because of their simplicity. As is now clear, in Wertheim's theory of motion and velocity perception this tempting simplicity is completely absent. In his theory, retinal movement and retinal velocity are not codes but must be calculated on the basis of retinal information. The reference signal needed for a correct perceptual interpretation of this code is not a simple signal but a compound one calculated on the basis of a diversity of independently calculated contributions. And the veridical perceptual interpretation is not arrived at after a single arithmetic operation but requires at least several independent calculations within different domains of processing. The multiple of indirect calculations makes one wonder whether this type of motion perception theory is really on the right track. Maybe the time has come to analyze two important questions in depth: What cues are used in motion and velocity perception and how are these cues used in producing veridical perception?

NOTES

1. The list of possible intermodal coordinations in the visual perception of motion can indeed be larger: there are data on the possibility of influencing visual autokinesis by providing a stable or moving acoustic landscape (Velichkovsky 1971).

2. These observations also present a strong argument against modern versions of "innervation" or "effort" theories of perceived motion (see e.g., Post et al. 1986), especially in view of individual variability in parameters of the fixational optokinetic nystagmus in this situation (Velichkovsky 1973).

## The illusion of self-motion in virtual reality environments

John Wann[a,b] and Simon Rushton[a]

*john1@edinburgh.ac.uk and s.rushton@edinburgh.ac.uk; [a]Department of Psychology, University of Edinburgh, Edinburgh EH8 9JZ, Scotland and [b]Department of Human Movement, University of Queensland, Australia*

The model presented by Wertheim provides a useful stimulus to reopen debate on a dichotomy that is particularly relevant for the rapidly emerging technology of virtual environments. A problem with much of the previous debate in this field is that, in most natural settings, perception is veridical and hence the mechanisms of perception remain ambiguous. By contrast, the principle underpinning the creation of a virtual environment (VE) is deception. It is technically impossible to present an observer with a VE that has coherence across the perceptual domains (e.g., for vision and vestibular stimulation), so the emphasis has to be on presenting visual displays that are salient enough to induce the required percept and to establish what other sensory conditions may be necessary to maintain that illusion. Research into perception in VEs also reflects back onto theory by providing a bridge between the ecological contexts that were favoured by Gibson (1966; 1979) and the unusual experimental conditions used by Wertheim and coworkers (e.g., rotating subjects sinusoidally in darkness with brief flashes of a striped display).

A typical VE application would be based upon a head-mounted display (HMD) with two liquid crystal screens (to allow a stereoscopic display), one positioned in front of each eye, with magnifying optics. Using the HMD is attractive because it can be made sensitive to head movement. Hence users can turn their heads and gain a new perspective or walk forward or back to explore detail within the computer display. The small exit pupil of most HMDs places severe limitations on the use of eye movements to sample the environment so our primary interest is in Wertheim's proposals on the percept of self-motion.

Wertheim presents the simple example of a train engineer, where the train accelerates and hence stimulates the vestibular apparatus, and produces a suitable motion percept ($V_{head,s}$: see target article). As the train velocity plateaus, however, the vestibular contribution to $V_{head,s}$ is gradually replaced with a visual reference signal, and the transition is smoothed by a conveniently large JND. The example of the train engineer tells us little, because in a natural environment it is most unusual to have vestibular conditions that conflict with the vection information, so the transition is assumed rather than confirmed. By contrast, the observer in a VE may experience visual-vestibular conflict. In our lab we have sat observers on a chair and sent them on a virtual (visual) roller-coaster ride where the visual experience should clearly be correlated with vestibular stimulation. Most naive users report quite a good impression of ego motion and while this is difficult to explain from an inferential perspective, it presents few problems for either direct theorists or Wertheim's model: $V_{head,s}$ builds up as a result of saturated vection in the absence of vestibular stimulation. On this point we make two observations.

(1) In recreating a virtual model of David Lee's (Lishman & Lee 1973) swinging room, we generally note that postural responses to expansion and contraction of a visual texture array are immediate and do not appear to build up over a saturation period, which argues for a more direct route than the proposed gating mechanism.

(2) Although some naive users can find our roller-coaster a moving experience, it is not compelling, and it is clear that vection is not sufficient to maintain a percept of ego motion beyond the transient perturbations of Lishman and Lee (1973). It is also well established that to produce an ego-motion experience strong enough to generate revenue from the public requires the addition of vestibular stimulation that is correlated, but not necessarily veridical (e.g., the type of public access

simulator built by Universal Studios). This in turn argues for some additive model for the perception of self-motion similar to that proposed by Wertheim.

So in what way might visual and vestibular inputs contribute to the $V_{head,s}$ reference? Well, further observations arise from the conditions under which the VE observer may not be *statically* rocketing around a roller-coaster. The observer may suddenly lean or walk forward, producing a conflicting vestibular signal rather than the convenient silence of the static observer or train engineer. Some VEs have even been designed on the basis of conflict, with observers locomoting around the VE by using small natural footsteps which are scaled to larger leaps in the visual world they see. The question facing the designer of VE concerns whether a stable percept of the world can be maintained in the face of visual-vestibular conflict. This then reflects on theoretical models of perception: How is this pattern of stability/nonstability to be explained in terms of a (nonlinear) additive model of ego-motion perception?

Imagine the following scenario: Participants sit on a playground swing and don an HMD through which they can observe a textured virtual world. We push them so that they feel motion through the environment and at the same time see an equivalent optic expansion. As they start to lose amplitude in their actual swing, however, we maintain the same expansion and contraction of the visual display. An inferential perspective might predict a breakdown in the percept of a stable visual world, but both Gibson (1979) and Wertheim could account for the subjects' perceiving that they were still swinging with the same amplitude and that visual information can overwhelm conflicting vestibular cues (e.g., Lishman & Lee 1973). Suppose, however, that we give them a second push, so their actual amplitude increases, but we rapidly decrease the visual amplitude of their swing. A point of departure must occur between a theory that stresses the primacy of one information source (vision) and a theory that suggests the previous percept resulted from strongly weighting a signal (vision) that is now diminishing and a lightly weighted signal that is now increasing (vestibular afference).

Unfortunately, we could not get a playground swing built in our laboratory before the deadline for this commentary, but we did translate this situation to the rotary equivalent. This has the advantage of being similar to the "vection drum" used by Wertheim, although we find it a less desirable paradigm because of the potential influence of eye movements (nystagmus) and also because uniform texture flows across the visual field very seldom occur in natural settings. Rotational movements of the head are also normally coupled closely with visual motion for many VE applications, but they serve to illustrate the paradigm.

Subjects knelt on a swivel platform while wearing an HMD. On the display they were presented with either a (stereoscopic) view of a drum with different texture stripes or a three-dimensional stereoscopic world with fields, roads, and buildings in view. The view then either rotated continuously at 60 deg/sec or it rotated sinusoidally (80 deg amplitude) to simulate the visual pattern that would result from the subjects' swivelling back and forth on the chair. In a second variant of the back-and-forth condition, an actual body rotation was introduced that was initially both in-phase with the visual rotation and of a similar amplitude (Fig. 1).

The subjects initially closed their eyes, then opened them and fixated a small rectangle presented at a constant (fusible) position on each screen. They then reported whether they had the impression they were moving and whether the world was stable. In the back-and-forth condition, the amplitude of actual movement was first steadily reduced and then increased as proposed for the playground swing experiment (above). Some further observations arise from these pilot experiments:

(3) It was much easier to induce the percept of self-motion in the back-and-forth (ecological) context than with constant unidirectional rotation, despite the fact that the back-and-forth motion would ordinarily have produced more variable vestibular stimulation. This in turn seems to emphasise the role of motion "plausibility" in such percepts.

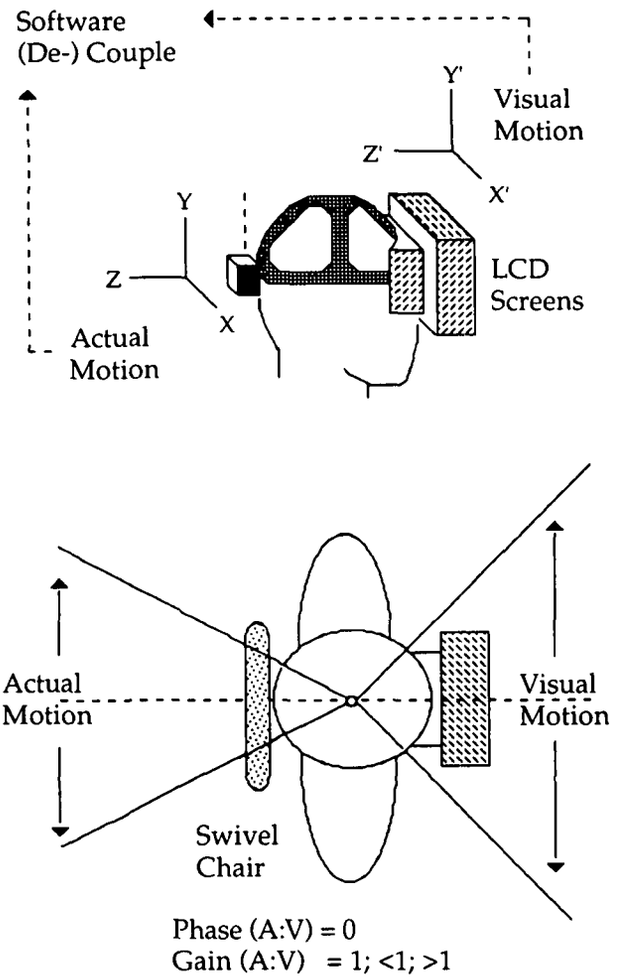


Figure 1 (Wann and Rushton). Upper: The two coordinate systems of a VE. The observer is presented with a visual array with orientation and flow components determined by the specification of a virtual viewpoint. Actual movement of the observer can be detected via a head tracking system. The crucial factor is that the coupling of actual (vestibular) motion to the motion of the visual array is done purely through software, and in many cases this coupling has to be weak to allow effective (visual) exploration in the absence of actual body motion.

Lower: Overhead schematic of the experiment reported here, where the observer is presented with visual rotation in the presence of differing degrees of actual body rotation. In all cases the two motion patterns are temporally synchronised (phase = 0), but can differ in amplitude.

(4) The illusion of self-motion was much stronger if the visual display was of a rotating world rather than a texture striped drum. Once again there seems to be some advantage to a more ecological context of objects and surfaces, which in turn produce a retinal flow that is not uniform (e.g., there is differential motion parallax).

(5) Although it was possible to reduce the amplitude of the actual (vestibular) motion without disturbing subjects' perception of their self-motion, this was not a robust effect. In particular, a breakdown in the perception of a stable world (e.g., in the fact that all global visual motion is the result of self-motion) was likely when there was a small phase difference between visual and vestibular information. This raises the question of how the optokinetic pathway can provide a strong self-motion reference signal to substitute for decreasing vestibular cues, while the

system is at the same time sensitive to small phase differences in visual and vestibular motion. These results are not in themselves sufficient to resolve any of the issues raised, but they suggest that in settings where visual and vestibular motion cues may be in conflict there is more to self-motion than meets the eye. They also emphasise that the target article makes a substantial contribution in restimulating research on the conditions supporting the percept of self-motion and environmental stability. We would stress our doubt, however, that a suitable theory of self-motion perception can be fashioned purely on data gleaned from highly constrained experimental conditions or unusual illusions such as the vection drum. What is required is a translation of such hypotheses to more naturalistic settings that can still be controlled experimentally. To take this model further there is a need to present subjects with natural optic arrays and flow characteristics while at the same time remaining able to manipulate the stability of the observer or visual world independently. Although current VE systems have a number of inadequacies, we believe they can provide a valuable window on human perception; the experiments quickly devised for this commentary provide example of this.

ACKNOWLEDGMENTS

*The production of this commentary was supported by a grant to J. Wann from the Joint Councils Initiative (U.K.) for Human Computer Interaction and by both of the Departments listed for the first author.*

## The significance of the active pick-up of information in ecological theories of motion perception

Lucy Yardley

*l.yardley@ucl.ac.uk; Department of Psychology, University College London, London WC1E 6BT, United Kingdom*

The target article presents a welcome reaffirmation of the contribution of vestibular information to the perception of ego motion relative to visual space. As Wertheim rightly notes, the contribution of vestibular information to visual perception has been underestimated or neglected by many investigators, and some of those who have adopted a "direct perception" approach have questioned the utility of the vestibular system as a source of information about absolute orientation (Stoffregen & Riccio 1988). Nevertheless, a number of studies have confirmed that information picked up by the vestibular system has a significant influence on judgments of visual orientation (e.g., Marendaz et al. 1993; Mittelstaedt 1983).

Wertheim also provides a valuable integration of research into ego and object motion, yielding many interesting and important analyses and questions for future research. His model of the processes involved in perception, however, appears to be an extension and refinement of the inferential approach; and his claim that the concept of a "reference signal" can incorporate, reconcile, and supersede both the inferential and direct theories of perception is consequently unconvincing. Wertheim has defined direct perception theory solely in terms of recent attempts to identify features of the optic flow that can visually specify object or ego motion. Although some proponents of Gibson's ecological theory of perception have been principally engaged in the search for visual kinaesthetic invariants, Gibson's original ecological theory of perception (as Wertheim himself acknowledges) contained many other vital elements, notably the idea that perception consists of the *active* detection of properties of the environment which are of significance in the context of the organism's activities; "The observer . . . explores the available fields of light, sound, odor and contact, selecting what is relevant and extracting the information" (Gibson 1966, p. 32). Gibson conceived of information pick-up as multimodal, emphasising that "information about the self is multiple and that all kinds are picked up concurrently" (1979, p. 115). Hence, although Gibson and his followers were interested in properties of the visual array which could offer the organism immediate veridical information about ego motion, the visual system was considered to be just one source of congruent information which could be picked up by many means and senses during active exploration of the natural environment.

The multimodal nature of the "reference signal" is therefore *not* incompatible with ecological theory, but the quasi-mathematical computation of ego motion from visual and vestibular "cues" posited by Wertheim *is* inconsistent with the ecological approach. Vection and other experimentally induced forms of misperception constitute perceptual anomalies, occurring only when the available perceptual information is impoverished or unusual, *and when exploration is prevented;* if head movement is permitted, the illusion of vection is delayed and attenuated (Lackner & Teixeira 1977), while more extensive exploration (e.g., moving around freely and touching the optokinetic drum) can dispel the illusion altogether. Artificial experimental conditions can provide a useful means of exposing constraints and biases inherent in our neurophysiology and perceptual habits, precisely because these conditions violate the naturally occurring environmental regularities to which we are attuned by evolution and experience (Runeson 1988). However, such conditions deliberately attempt to exclude most of the processes involved in natural perception, namely, the active, flexible utilisation of meaningful information derived from rich, veridical, multimodal sources.

Nevertheless, despite strict experimental control, the experience of researchers who have previously attempted to establish how visual and vestibular "stimuli" combine to "produce" specific sensations of ego motion has shown that inter- and intra-subject variability is typically very wide and seems to be more closely tied to perceived higher-order properties of the environment such as the foreground/background distinction (Ohmi et al. 1987) than to physical parameters such as the area and velocity of motion in the visual field. Moreover, perception of ego motion has already been shown to be influenced not only by visual and vestibular information but also by auditory and somatosensory information, cognitive evaluations, and sensorimotor activities, experience, and skills (see Yardley 1992, for a review). If Wertheim's conception of a reference signal is to provide a truly "unified framework," able to account for natural perception as well as for the various situations in which misperception occurs, it will be necessary to determine how the reference signal (perhaps more appropriately termed the "motion percept") is influenced by neurophysiological constraints, multimodal information, sensorimotor experience, and purposeful activity.

# Author's Response

## Motion perception: Rights, wrongs and further speculations

Alexander H. Wertheim

*wertheim@izf.tno.nl; TNO Institute for Human Factors, 3769 ZG Soesterberg, The Netherlands*

### R1. Ego-motion perception

The defining feature of inferential theory is its assumption that extraretinal information is used in the perception of

*object* motion. It is this feature which direct perception theory finds difficult to accept. My model provides an alternative to that debate, rather than an alternative to current theories of ego-motion perception. Many commentators do not seem to have understood this. **Yardley,** for example, calls the model a "welcome reaffirmation of the contribution of vestibular information to the perception of ego motion," and **Kim & Turvey** misrepresent the model by stating that it implies that when a retinal signal "is matched by the reference signal then it is ego movement that is taking place." However, as noted correctly by **Honda,** the model was proposed as a description of how we perceive object motion, not ego motion.

The description of object-motion perception should also be valid during ego motion of the observer. This is why the mechanism responsible for perceiving ego motion had to be included as a module in the model. The object-motion perception mechanism needs to "know" about ego motion, otherwise it cannot estimate how the eyes move in space (see **Becker & Mergner,** who consider this so "logical" as to call it "trivial").

In describing the subsystem for ego-motion perception, however, I did not hypothesize (as assumed by **Wann & Rushton**). Instead, I just borrowed from the existing body of (mainly neurophysiological) literature, in which the mechanism responsible for sensations of ego motion, vection, and visual-vestibular interactions is described in great detail. This literature – which rarely makes reference to mechanisms of object-motion perception – is often unfamiliar to researchers in the field of visual perception. This is illustrated by some of the commentaries. Thus, for example, **Coombs** provides some ideas about how we should investigate possible effects of confirmatory or conflicting vestibular stimulation on the salience and time course of vection sensations. **Rieser** (who gives a partial answer to Coombs) speculates about how exocentric ego-motion perception might be empirically measured. **Wann & Rushton** suspect that the complementary interaction between vestibular stimulation and vection in my example of the train engineer is "assumed rather than confirmed" (see also Rieser's doubts). But there is already a long research tradition and a rich literature on all these issues (for classic reviews see Cohen & Henn 1988; Dichgans & Brandt 1978; Guedry 1974; Henn et al. 1980).

This is why I deemed it necessary to summarize the literature to the extent that seemed relevant for present purposes. I may rightly be reprimanded for being incomplete. Thus, I discussed only those developmental time courses for visual vestibular interactions that were needed to illustrate certain theoretical arguments. This no doubt caused **Kim & Turvey** to remark that I seem to overstate the time course of development of circular vection as presented in section 2. Other objections (**Becker & Mergner, Belopolsky, Coombs, Da Vitoria Lobo, Previc, Rieser, Ross**) refer to my incomplete catalogue, or lack of a mathematical description, for additional factors contributing to the inputs of the mechanism that generates percepts of ego motion and vection (postural and tactile information, arthrokinetic feedback from the joints, afferents from the neck muscles, or specific cognitive inputs such as expectancies and depth perception, and whether ego motion and self-motion are active or passive). These objections are valid of course, and I am indebted to those commentators (e.g., **Büttner & Straube, Probst, Sauvan, Thier et al.**) who have taken the trouble to review in their commentaries much more of the relevant recent information from this field of research.

Although the present model can be used to add to this body of literature (see sect. 5.4), one might of course challenge these ego-motion theories from a different point of view and speculate about alternatives. Such challenges are found in the commentaries of **Riccio, Kim & Turvey, Coombs, Stoffregen, Yardley,** and **Wann & Rushton.** The main thrust of these arguments stems from direct perception theory, and is focused on the belief (expressed explicitly by Kim & Turvey) that much more attention should be devoted to the nature of the visual input to the system. This illustrates the point I made (perhaps a little too strongly) in Note 2, that the main line of research in the tradition of direct perception theory concerns the analysis of this visual input. Nevertheless, as the note states, I agree with **Stoffregen, Yardley, Shebilske,** and **Riccio** that even within that tradition some attention has also been given to other inputs, such as vestibular ones.

These challenges, however, (and whether or not they are sensible) are quite secondary to the model proposed in the target article as they do not relate to *object*-motion perception.

The erroneous belief that the model is intended to explain ego-motion perception may have led some commentators to become confused about the role of reference signals. The most obvious examples are **Kim & Turvey** (according to whom I supposedly hypothesize that "for an organism to perceive. . . self-motion. . . a special reference signal is needed"), **Previc** (who assumes that I use the concept of a "visual reference" to explain "whole-body percepts such as vection," or to establish "whether or not self-motion has occurred") and **Riccio,** (who assumes that my paradigm serves to describe the effects of extraretinal signals on "the phenomenology of ego motion"). These commentators apparently assume that the reference signal affects percepts of ego motion. It does not: in my Figures 1 and 7, reference signals feed only into the mechanism for object-motion perception.

Whatever the reason for these misunderstandings, they should serve as a warning not to use the term "motion perception" carelessly, without mentioning whether object or ego motion is meant. One can easily be misunderstood. For example, **Riccio** criticizes the first sentence of the second paragraph of my section 1, where I claim that according to direct perception theory "the perception of motion derives exclusively from afferent retinal information." Riccio erroneously assumes that I mean the perception of ego motion, but I meant the perception of object motion: up to this point in the target article the only topic discussed is object-motion perception.

## R2. The reciprocity assumption

The mixup of ego- and object-motion perception may well stem from what I would like to call the "reciprocity assumption." This is illustrated most explicitly in **Probst's** commentary. In his Figure 2, he proposes a model that

differs in an important way from the present one (apart from its failure to describe how eye movements affect object-motion perception): it postulates a reciprocal relationship between percepts of object motion and percepts of ego motion (see the dotted reciprocal arrows between the boxes which represent the mechanisms for self- and object-motion perception). A similar assumption is made by **Andersen,** who criticizes my model on the grounds that the percept of self-motion should also be an output of the object-motion perception system.

As **Belopolsky** correctly notes, there is no such interaction in my model (see Fig. 7). I postulate something else: that the object- and ego-motion perception systems share, to a large extent, the same inputs. Hence their outputs (the percepts) often develop more or less synchronously. This creates the impression that the two percepts are reciprocal, but it is more appropriate to speak of a complementarity (**Skavenski**) between percepts of ego and object motion, a complementarity I have called the interfacing of object- and ego-motion percepts. But no reciprocal two-way interaction, such as present in the **Probst** model, is implied.

The reciprocity assumption is easily associated with my example of circular vection in section 2, because here the time course of development of the percept of circular ego motion runs so nicely parallel to the development of the percept of (a gradually slowing down of) object motion. Similar parallel time courses between object- and ego-motion percepts occur in many normal nonlaboratory situations – such as when we look at a moving train next to our own stationary train, or, as **Rieser** mentions, when a car next to our own stationary car unexpectedly moves – creating both vection and a concurrent percept of object stationarity. Such instances may indeed give rise to the idea that object-motion percepts affect ego-motion percepts, or vice versa (see **Andersen,** who states that "stationarity resulting from vection saturation necessarily implies that the observer perceives self-motion").

The present model does not take the apparent reciprocity between percepts of ego and object motion as an a priori, self-evident premise, but as something that must be explained. In fact, the model was developed with this purpose explicitly in mind, as should be obvious from my introductory question in section 2: "Why, during saturated circular vection, is the drum perceived as stationary in space?" I might as well have phrased this question: "Why can we not perceive the drum as moving in space when vection has reached its maximum steady state?"

The reciprocity assumption sometimes takes a hidden form. This happens when we ask questions such as: "How does the system choose a combination of ego motion and object motion that accounts for the observations?" (**Coombs**), or: "How does an animal distinguish its own movements from those of objects?" (**Kim & Turvey**). Such questions presuppose acceptance of a reciprocity assumption. They imply that percepts can be ambiguous as to whether they refer to object or ego motion (**Held**). This implies that one can have percepts that do not mean anything until the brain "decides" what they mean, using a kind of trade-off rule in which part of the "percept" is attributed to object motion and the remaining part to ego motion or vice versa. **Rieser,** for example, seems to assume that the brain uses visual and vestibular information to take the decision, and **Probst** mentions that he has extended his model to accommodate the assumption that the "perceptual decision between object motion and self-motion" takes place at the level of the vestibular nuclei.

In my model, however, no such decisions are taken. It simply postulates the existence of two distinct mechanisms, which create percepts of different events (object motion and ego motion), often in temporal synchrony. Thus how such decisions are made becomes a nonquestion. A particularly interesting case of a hidden reciprocity assumption is the problem presented by **Kim & Turvey.** They claim that the model must be in error, because it cannot explain how we can have a single experience of ego motion in the presence of many objects moving at various speeds. If one embraces the reciprocity assumption, and believes that perceived object motion affects percepts of ego motion, Kim & Turvey would be right. This shows the problematic consequences of the reciprocity assumption. Without it Kim & Turvey's problem is nonexistent.

It should be noted here that, although some direct perception theorists appear to adhere to a reciprocity assumption, it is not a requirement of direct perception theory at all. According to that approach, there is no ambiguity in optic (or retinal) flow, because information about object and ego motion is present in separate invariants. Both **Tresilian** and **Stoffregen** make this quite clear. But traditional inferential theory does not need the reciprocity assumption either, because it postulates that percepts of object motion stem from retinal and extraretinal signals and inferential theory does not concern itself with percepts of ego motion (with the notable exception of the Post and Leibowitz (1985) theory, but that theory also implies no reciprocity assumption).

But if not from direct or inferential theory, where does the reciprocity assumption stem from? My guess is that it originated in the physiological literature on visual vestibular interactions and ego motion (see **Probst,** who mentions his own and Dichgans and Brandt's [1978] neurophysiological hypotheses to explain the reciprocity assumption, and Mergner and Becker [1990] for a summary review of discussions around the assumption). Could it be that perception theorists have borrowed the reciprocity assumption just to explain what happens in cases where it is difficult, if not impossible, to apply the traditional concepts of extraretinal signals or object-motion invariants?

## R3. The optokinetic component in the reference signal

The example of circular vection was brought up to illustrate problems for both direct and inferential theories if one does not accept the reciprocity assumption: how to explain the particular time course of development of the percept of object (i.e., drum) motion and stationarity in space, which is known to concur with the development of circular vection inside an optokinetic drum. This is where the optokinetic component in the reference signal comes in. It explains that particular time course.

The fact that it is a more or less new concept (but see my reply to **Skavenski** below) seems to have motivated **Held** to consider it an example of "contorted reasoning." As **Probst** correctly states, the name "optokinetic compo-

nent" might create some confusion (perhaps I should have called it just a "visual component"), but that is only a matter of semantics. Held's doubts about whether this "dubious" and "strange" postulate is "necessary at all" do not stem from the name I gave the component. They stem from his acceptance of the reciprocity assumption, which apparently eclipses my extensive discussions about the reasons for postulating the component and about the strong explanatory and predictive power of this concept, issues which are addressed throughout the target paper.

**Held**'s argument against the optokinetic component derives from his belief that during circular vection there simply is no reference signal, and that my model therefore does not apply to such circumstances. To bolster his view, **Held** refers to the phenomenon of "waxing and waning" of vection sensations in the optokinetic drum. He claims that this phenomenon "confirms the validity" of his belief that no reference signal is present during vection because it implies a perceptual ambiguity. According to **Kim & Turvey,** the ambiguity reflects what one would normally expect in such a situation of "ecological contradiction" (which is why **Stoffregen** would not call it an illusion), but Held explains it by referring to the reciprocity assumption: it presumably stems from "the trade-off between perception of self-movement and object movement."

These commentators present the "waxing and waning" phenomenon as a normal characteristic of vection. But in fact, it occurs only sporadically – mostly during circular vection induced with a fixed semicircular screen (on which moving stripes are projected) and during trials which last very long. Usually it does not happen at all inside a really rotating optokinetic drum. As **Rieser** notes, everyone who has ever experienced vection inside such a drum can testify to the overwhelmingly powerful and usually continuous sensation of ego motion. Neither **Held,** nor **Kim & Turvey,** nor **Stoffregen** seem to attribute any significance to this fact.

The "waxing and waning" phenomenon does not contradict my model as **Held** suggests. On the contrary, it can easily be explained as stemming from modulations of the reference signal: first, the "waxing and waning" phenomenon has, to my knowledge, never been reported with empirical evidence that the observer's eyes and head remain fixed in space, as was the case in my example of circular vection. Thus, when it happens, it might well be related to eye or head movements of the observer (for discussion see Mergner & Becker 1990), which by chance stabilize the image of the striped drum on the retinae. Vection should then stop. This is not my prediction, but a prediction of Mergner and Becker (1990), who showed this to happen during sinusoidal vection (see my sect. 6.1). There are in fact more reasons why the inputs to the mechanism responsible for perception of ego motion (i.e., vection) may be distorted, such as habituation (as Held himself notes) or attentional factors (see **Coombs**'s suggestions, and **Da Vitoria Lobo**'s observation that we do not seem to experience vection in the cinema). Thus, I do indeed recognize that vection may sometimes change or is occasionally lost entirely. Such features of the ego-motion perception mechanism are not at all incompatible with my model.

Second, the model uses precisely these features to explain the "waxing and waning" of perceived drum rotation in space: as mentioned before, inputs to the ego-motion perception system also serve to generate reference signals. Hence, when these inputs change they change reference signal magnitude, and thus percepts of object (i.e., drum) motion in space. In other words, whenever (and for whatever reason) vection "waxes and wanes," the perception of drum rotation in space "wanes and waxes."

## R4. Retinal image motion has no perceptual meaning

**Held**'s claim that drum velocity is perceived as "waxing and waning" because of an absence of reference signals implies that we can somehow perceive object motion without using reference signals. This runs counter to the main premise of the model – which is also the main premise of all inferential theories of object-motion perception: image motion across the retina cannot in itself serve as information about object motion. **Da Vitoria Lobo** also disagrees with this premise, but for another reason. He does not believe in an efference copy signal. Instead he assumes that all information in the reference signal about eye movements derives from computations made on the basis of retinal information. The concept of such a "visual efference copy" was discussed in section 6.3. I questioned it because when such visual information is not available, we can still perceive object motion (e.g., the motion of a single dot moving in total darkness; see the commentary of **Tresilian**). If Da Vitoria Lobo agrees that this is true and still maintains that eye movement information is purely visual, it follows that he believes there exist situations where object motion in space can be perceived in the absence of reference signals. Hence his opposition to my analysis of the graph presented in Figure 3 of the target article: he assumes that it is specific to experimental conditions that do not exclude visual information about how the eyes move (presumably stemming from the visibility of the borders of the screen). This is not true, however. Such a graph is always obtained, including with a very small stimulus visible in an absolutely darkened environment (e.g., De Graaf & Wertheim 1988), that is, when there is no visual information (screen borders, etc.) from which knowledge about eye movements might be abstracted.

Figure 3 is simply a plot of the velocity of the retinal image of a stimulus at the two opposite thresholds for motion ( of that stimulus) obtained while the eyes move across that stimulus. Other stimuli, or stimuli presented in absolute darkness, might yield different thresholds, but one always gets thresholds. The slope or width of the no-motion range might differ, but there will always be a certain no-motion range, an area above that range where the subject perceives stimulus motion against the eyes and an area below that range where the stimulus is perceived as moving with the eyes. Any horizontal line will always necessarily cross through these areas. Hence, whatever the conditions of the experiment, the argument that retinal image motion itself has no perceptual meaning remains valid.

The same can be said of the reference signal. This signal has in itself no perceptual meaning either. The perception of absolute object motion – that is, the awareness that

we are seeing something that moves in space – arises from a specific interaction between retinal and reference signals. Neither of these signals has any priority over the other. Some commentators have not understood this basic point. For example, **Yardley** proposes that I should have termed the reference signal "the motion percept," and **Andersen** incorrectly assumes that my model implies that retinal motion is "the primary perceptual stimulus." **Bridgeman & Blouin** also err when they claim that my model supports their assumption that "in a normal structured visual environment, retinal signals provide the dominant information. . . for perceiving the motion of a visual object," while "in a dark room. . . extraretinal signals necessarily dominate."

## R5. Frames of reference

Several commentators (e.g., see **Swanston, Velichkovsky & Van der Heijden**) suggest that the vector subtraction process between retinal and reference signals can or should be viewed as a process of redefining retinal image motion in terms of an exocentric frame of reference. This is true, but I avoided that terminology because it could mix levels of analysis, confusing cognition with physiology. **Belopolsky,** however, seems to believe that such translations are impossible (there being "no uniform rule for transforming one coordinate system into another"), and Swanston, in his final remark, doubts whether my model supports such an account of motion perception. Yet there is no reason for such doubts. In this section I will try to illustrate this, as it might elucidate several other issues.

Retinal image motion is encoded in the retinal signal as image slip across the retinal surface, that is, it is defined in terms of a retinal coordinate system. If image motion is to gain any other meaning, it must be redefined in coordinates of another frame of reference, one created by the brain itself.

It is possible to assume that this self-generated frame represents a head- body- or egocentric coordinate system, as mentioned by **Becker & Mergner** and by **Swanston.** Percepts are then generated of how objects move relative to the head or body. The model presented in my Figures 1 and 7, however, is not about such percepts at all. **Kim & Turvey**'s remark that the model is limited to an "object displacing relative to the observer" is incorrect. The model concerns percepts of absolute object motion, which is object motion defined in terms of the spatial coordinates of exocentric space. To generate such percepts, the brain must translate the retinal coordinates of image motion into coordinates which somehow correspond to the abstract "Newtonian" dimensions of environmental space. Those who like to use the terminology of cognitive psychology might refer to this self-generated frame of reference as a mental representation of the dimensions of external space. The signal that carries the building blocks for this mental representation – or, stated differently, the signal that enables the appropriate coordinate translations of image motion to be carried out – may be called the "information-for-an-exocentric-frame-of-reference signal." I prefer the shorter name of "reference signal."

The generation of a percept of object motion in space can thus be viewed as a recalibration of image motion with the help of the reference signal. **Shebilske** thus rightly suspects that my model may be classified as a calibration model in the sense intended in the Bridgeman, Velichkovsky, and Van der Heijden article in this *BBS* issue. It is this recalibration that I have described formally as a vectorial subtraction of retinal and reference signals. Hence, I agree with **Stoper** that the subtraction does not mean that a retinal signal is (partially) canceled, that is, annulled. The subtraction is needed to determine the difference between the two signals. It is this difference that provides the stimulus for percepts of object motion or stationarity in space.

One single reference signal can recalibrate the movement parameters of many image motions at the same time. Thus the solution to **Kim & Turvey's** problem (if there are many moving objects at various speeds, "what could it then mean to speak of perceived object velocity as due to the magnitude difference between retinal and reference signals. . . ?") is simple: the vectorial differences between all retinal signals on the one hand and that single reference signal on the other yield a multitude of percepts of how all these objects move in space.

It is external environmental space, defined by the direction of gravitation and the earth's surface, with which we, and many other organisms, must deal in an ecological sense. Hence the mental representation of this frame of reference better be a good one. No wonder the reference signal uses vestibular afferents as inputs. The vestibular apparatus is a sensory system which (among other things) "picks up" the direction of the earth's gravitational field and thus allows a kind of "anchoring" of the mental representation to physical reality.

I hence strongly disagree with **Stoffregen,** who claims that our percepts (or the internal estimates involved in their generation) are not linked to such external physical parameters as gravitational force. He refers to the fact that there is "nothing absolute" about Newtonian space, because Einstein's theory of relativity has shown it wrong. But that is irrelevant. The dimensions of Newtonian and Einsteinian space differ indeed when our subjects, or their retinal images, move relative to us (the investigators) with speeds approaching the speed of light. But usually they don't. Thus Newtonian and Einsteinian dimensions are indistinguishable for all practical purposes. I concede that my model does not apply to Superman.

Although this vestibular "anchoring" no doubt improves the veridicality of the internal representation of exocentric space, it is no guarantee against errors. They may occur for a variety of reasons, such as errors made by the vestibular apparatus. In the target article much attention is devoted to the perceptual consequences (illusions) of such errors, and to how the perceptual system defends itself against them (e.g., by masking them with the JND).

It should by now be clear why I disagree strongly with **Held,** who claims that reference signals can be absent, but not zero. They cannot be absent, and sometimes they are zero. They are always present, because without them the translation of image motion into percepts of object motion or stationarity in space would be impossible. Zero is just a scale value: reference signals can have positive or negative values, depending on the direction of eye velocity in space. A zero reference signal simply means that the message encoded in the reference signal is that the eyes

do not move in space. The same can be said about retinal signals (the sign of which expresses the direction of image slip across the retina) and about the difference between retinal and reference signals (the sign of which determines the direction of perceived object motion). Hence, when we perceive an object as stationary, it means we perceive it as having a zero velocity. It is wrong to assume that in such cases the object-motion perception mechanism is inactive and generates no percept.

## R6. Perception of relative and absolute motion

It should now also be clear why I consider it crucial that we distinguish between percepts of absolute and relative motion. Absolute motion refers to an absolute "Newtonian" frame of reference. Relative motion between two objects does not need any frame of reference external to the objects themselves: the relative motion between two cars is defined by the spatiotemporal relationship between only the two cars themselves. If they move relative to each other with, say, a velocity of 50 mph, that velocity is the same, irrespective of how fast they move in Newtonian space (they may be moving at 30 and 80 mph, or any other combination of absolute velocities which differ by 50 mph) or in any other frame of reference. For example, the two cars will still move relative to each other with a velocity of 50 mph in an egocentric frame of reference, for example, when they move on a road parallel to the track of a train traveling at 100 mph, from which I look at them (they may then move relative to me with velocities of −70 and −20 mph). These examples illustrate that we need not take frames of reference into consideration when analyzing percepts of relative motion. This is why I have stated that reference signals have no effect on relative motion. Equation 12 is just shorthand for this argument. This is not to say that relative motion is always perceived veridically. When the two objects do not move in a plane parallel to the one in which the observer moves, their retinal image velocities are affected differentially (see below), and we will misjudge their relative velocity.

This is why I disagree with the assumption (see **Bridgeman & Blouin** and **Tresilian**) that reference signals are used only in dark environments. The involvement of reference signals in perception does not depend on circumstances but on the kind of percept we are talking about: percepts of absolute motion or percepts of relative motion. These are separate percepts. They do not reflect different interpretations of the same event but different aspects of it. The brain does not decide how much image motion should be attributed to absolute and how much to relative motion (as claimed by Wallach 1959). To assume that it does is like assuming that the brain has to decide how much of an object's image must be attributed to the perception of its size and how much to the perception of its shape. These are distinct percepts; they refer to different stimulus aspects, even though the aspects may have some physical (e.g., perspective) relation to each other.

I accordingly also object to the assumption that relative motion distorts percepts of absolute motion. **Mateeff & Hohnsbein** criticize my objection by referring to the "cycloid phenomenon" (see also **Stoper**): a single light spot on the rim of a rolling wheel in total darkness is perceived as absolute motion (its path is correctly seen as cycloidal). If another light is attached to the hub of the wheel, the perception of this cycloid disappears. The dot on the rim of the wheel now appears to rotate, like a real wheel, around the hub light. This presumably illustrates the influence of relative motion on percepts of absolute motion.

I disagree: in both conditions we perceive absolute motion on the basis of the difference between a retinal and a reference signal. In the condition without the hub light, where the eyes track the light on the wheel, the retinal signal is zero, but the reference signal is not. Its magnitude modulates in a manner that corresponds to how the eyes move in space, which is cycloidally. Hence the difference between retinal and reference signal, which corresponds to the perceived absolute motion of the rim light, is cycloidal as well. When the eyes pursue an added light positioned on the translatory moving hub, the reference signal encodes a unidirectional eye movement (it represents a velocity vector with a fixed direction). However, the image of the light on the rim now moves in a circular path on the retinae (around the fovea). The vectorial combination of this rotating retinal velocity vector and the unidirectional reference vector is a velocity vector rotating around a center, which itself moves unidirectionally in space. This is what we perceive, although the percept is not veridical. This explanation corresponds to the one given by Stern and Emelity (1978) and does not involve (presumptions of dominance of) relative motion.

This explanation of the "cycloid phenomenon" can be tested, because it yields a prediction: if we pursue the rim light in the presence of the hub light, we will again see it as moving cycloidally, because its retinal signal and the reference signal are the same as in the condition without the hub light. Their difference (i.e., the perception of absolute motion of the rim light) should not be affected by the presence of an additional image from the hub light. I tried this on my computer, because I could find no reference in the literature which includes this condition, and I can recommend it to my readers: one easily perceives the cycloid again.

The "cycloid phenomenon" is interesting for other reasons as well: it runs counter to two basic postulates of direct perception theory. First, it shows that our percepts may be radically different, depending on how we move our eyes. This contradicts the assumption that eye movements only serve to improve perceptual veridicality (see, e.g., **Kim & Turvey**). Second, the percept we have when fixating the hub light is one we are so accustomed to that we "recognize" it as that of a rotating wheel and therefore think it is veridical. In fact, it is not. This shows that the addition of visual information does not necessarily lead to improved perceptual veridicality. Even worse, the restricted laboratory condition with only a single light in total darkness (the condition without hub light) yields a more veridical percept than visually rich everyday situations. If, for example, we look at the hub of a bicycle wheel (as we normally do if asked to look at a wheel), we see the air valve of the tire as rotating in space, that is, we perceive an illusion. Hence it is not necessarily true that perception research carried out in "artificial conditions" runs into "logical difficulties" (**Rieser**), and needs to be done with "more natural visual stimulation" (**Wann & Rushton, Riccio**).

**Mateeff & Hohnsbein** ask what would be the outcome of an experiment on relative motion similar to the one that underlies my Figure 5, but with stimulus patterns that move not only relative to each other, but also relative to space. Actually, in the experiment of Figure 5 this was already the case. In each condition we used a criterion stimulus with a fixed absolute velocity, while the absolute velocity of the other stimulus was variable (it had to be adjusted by the subject until no relative motion between the two stimuli was perceived). In each condition, however, the absolute velocity of the criterion stimulus was not zero but slightly below its threshold for absolute motion in that condition (determined prior to the experiment, with only the criterion stimulus and the moving fixation point). This was done for methodological reasons: if the criterion stimulus had been perceived as moving at various absolute velocities, we might have been open to the criticism that this could have affected the results (as depicted in Fig. 5). I agree with Mateeff & Hohnsbein, however, that an experiment in which the criterion stimulus moves at higher (above threshold) absolute velocities, would be interesting. The reason is that, although relative motion cannot affect perceived absolute motion, the converse is not necessarily true: when two stimuli do not have the same distance from the observer, their absolute motions in space affect how we perceive the relative motion between them, because distance affects image velocity. But this does not alter my prediction that their thresholds for motion relative to each other will still be a function only of their retinal image velocities and the JND between them.

## R7. The distance parameter

This brings us to the role of distance in motion perception. Several commentators (e.g., **Belopolsky, Boothe, Hadani & Julesz, Rieser, Sauvan, Swanston, Velichkovsky & Van der Heijden**) criticize my model because it does not take distance into account. This is true. The model has been developed to explain and test the effects of other factors on object-motion perception, distance being constant between conditions. However, the model can be extended to include a distance parameter. In this section I will make a first attempt in this direction.

Image velocity, if expressed as linear velocity across the retinae, is reduced in proportion to the distance between the object and the eyes. The proportionality may be very complex, because it depends not on distance alone, but also on the particular combinations of ego and object motions, that is, their directional, angular, and linear components (see, e.g., **Stoper**).

The reduced image velocity with distance has implications for the reference signal. At the point of subjective stationarity (PSS), where the stimulus is perceived as stationary in space, retinal and reference signals are by definition equal irrespective of how far away the object is. It thus follows that the reference signal too is scaled down in proportion to distance. Hence, the definition of the reference signal as the brain's estimate of eye velocity in space must be changed to: the estimate of eye velocity in space scaled down in proportion to an estimate of distance. In the terminology of frames of reference, this estimated distance parameter in the reference signal could be said to correspond to the depth axis of the internal representation of external space.

**Honda**'s suggestion may be right, that "the visual system knows that distance is mathematically related to velocity" (see also Gogel 1981; Van de Grind et al. 1992; Wist et al. 1976). The function according to which the reference signal is scaled down might well derive from such knowledge, although it will always include the observer's subjective estimate of distance.

In a formal sense we have a case here in which a cognition (of distance) – whatever its source – is an added input to the reference signal. That agrees with the finding that perceived distance affects vection (see **Previc** and **Velichkovsky & Van der Heijden**), because any stimulus that serves as an input to the ego-motion perception mechanism should also contribute to the reference signal.

This reasoning yields the prediction that when distance is misperceived, the reference signal is in error and illusions may occur, for example, those described by **Ross** (e.g., illusory motion of objects during walking, which happen when the distance to the object cannot easily be perceived) and **Swanston.**

It is intriguing to speculate on the possibility that this also explains (part of) the stationarity tendency of large objects. The point is that size and perceived distance are Gestalt related: small objects may appear to be further away than large ones, especially when other distance cues are absent (e.g., in a darkened environment). Hence, reference signal magnitude may be reduced with smaller stimuli. Consequently, the Filehne illusion (which is caused by the fact that the reference signal is already too small) should become stronger with smaller stimuli, even if they are only briefly visible, as suggested by **Mateeff & Hohnsbein** (and Ehrenstein, see sect. 6.3). We are currently investigating this issue in our laboratory by comparing PSS measurements with large and small stimuli placed at various distances from the observer. Comparing the magnitudes of reference signals between conditions might yield an empirical measure of the distance parameter in the reference signal.

## R8. Motion parallax

Referring to Gogel's (1990) work, **Swanston** mentions correctly that during head and ego motion percepts of illusory object motion (sometimes called motion parallax) may occur. However, if the model includes a perceived distance parameter, it easily explains this phenomenon. Let us start by restating that, at a certain moment in time, there can be only one "single valued" (**Stoper**) reference signal and by assuming that its magnitude depends on the distance $D$ to a certain plane $P$. Let us further assume that a stationary object $O$, which lies in plane $P$ is indeed perceived as stationary in space. Hence the magnitude of the reference signal equals the velocity of $V$ of $O$'s image on the retinae. Now assume that at various other distances there are also some stationary objects in the environment. Their retinal image velocities, that is, their retinal signals, are thus smaller or larger than $V$, that is, different from the singular reference signal. If these differences become larger than one JND, those objects are seen to move in space.

Imagine an object farther away than $D$. Its retinal signal will be smaller than the reference signal. If that difference

exceeds one JND, the object will appear to move in space in the same direction as the eyes of the observer (see sect. 5.2). Conversely, an object closer to the observer than *D* will have a retinal signal larger than the reference signal. Thus it will, if this difference is larger than one JND, be perceived as moving in space in the direction opposite to the eyes. This is Gogel's phenomenon of motion parallax as described by **Swanston.**

The JND for reference signals containing vestibular or optokinetic components may be rather large (see sect. 5.4). Hence we may normally perceive little illusory object motion if any during ego motion. Note that this description of motion parallax qualifies it as illusory *absolute* motion in space. It does not refer to how (images of) objects move relative to each other on the retinae (as suggested by **Swanston**).

## R9. Perceiving visual world versus visual field motion

This brings to mind the distinction between percepts of visual world and visual field motion (see **Stoper** and **Tresilian**). It seems that this distinction just reflects the distinction between perceiving object motion with a reference signal that takes proper account of distance and perceiving object motion with a reference signal in which the distance parameter is incorrect with respect to that stimulus. The difference in "salience" that seems to relate to the distinction between percepts of world and field motion may simply reflect an attentional correlate: illusory object motion in space caused by errors in distance estimation would remain unnoticed, not only because the errors yield differences between retinal and reference signals smaller than one JND. It is also quite likely that we pay little attention to such objects, although this does not mean we never perceive motion of unattended objects (e.g., when driving a car, or when specifically asked to, as in my experiments on the Filehne illusion, or in the experiments mentioned by **Becker & Mergner**).

This explains why situations such as those described in the lower right hand panel of Figure 2 in **Boothe**'s commentary do not usually yield such illusory motion percepts. Hence, I disagree with **Stoper,** who calls visual field motion "an artifact produced by unnatural viewing conditions" and who attributes it to "the motion of the observer rather than motion of the world" (which implies the reciprocity assumption).

Another reason to doubt the tenability of the distinction between percepts of visual world and field motion is that it yields curious contradictions, as illustrated by **Stoper**'s remarks. He claims that my experiments on the Filehne illusion (or on "paradoxical motion" as Stoper would call it) do not measure percepts of visual world motion because such percepts are veridical: if I had measured visual world motion, my subjects "would, of course, have seen the large background stable." Because sometimes they did not see the background as stationary (e.g., with a briefly visible stationary background; see Fig. 4b), they must have been reporting visual field motion. That cannot be correct, however, because in other cases (e.g., with a continuously visible background) the subjects correctly reported background stationarity (see Fig. 4a), although Stoper claims that in such cases visual field motion should still occur.

In my experiments on the Filehne illusion, the fixation point and the background stimulus move in the same plane. Hence, their distance to the observer is the same and corresponds to the distance parameter in the reference signal. Both the velocity of the background stimulus and that of the fixation point are thus perceived with the appropriate reference signal, even though the eyes focus only on the fixation point. I therefore disagree with **Becker & Mergner**'s suggestion that my model does not apply to the perception of motion of the fixation point. Just enter its image velocity (zero) and the velocity of the tracking eyes into my Equation 9 and the perceived velocity of the tracked fixation point will result. In fact, this is how I explained the Aubert-Fleischl paradox and center surround induced motion (sects. 5.2 and 5.3).

## R10. Mathematics

**Belopolsky** states that the model confuses dimensions because it implies the addition of linear and angular motion vectors. The use of different dimensions is just a mathematical convention, however. The brain could just as well use only one dimension. Take the example of an observer who fixes his gaze on a tree while pressing his nose to a window in a train that moves at a given velocity. The reference signal then consists of a linear ego-motion component induced by optic flow and a rotary efference copy component from the counter-rotation of the eyes. Let us further assume that the observer correctly perceives the tree as stationary in space. Since the retinal signal is zero, the reference signal must be zero too (approximately). In my model, this happens because the two reference signal components cancel each other. Mathematically we can express this as a requirement to translate the efference copy angular velocity component into a linearly dimensioned velocity component equal in magnitude, but opposite in sign, to the optokinetic component in the reference signal. Such a translation is simple, although the (estimated) distance to the tree is needed. At any given moment in time it would be:

$$V_{\text{efference copy}} = -V_{\text{ego}} = \frac{D}{\cos\alpha} \times \omega$$

where $V_{\text{efference copy}}$ is the translated linear velocity of eye rotation, $V_{\text{ego}}$ is linear ego velocity, D is the (estimated) distance on a straight line between eye and object, $\alpha$ is the visual angle of the eyes in the head, and $\omega$ is the angular velocity of the eye rotation.

**Becker & Mergner** seem to misunderstand the nature of a JND (and my Fig. 5), because they ask whether it also applies to retinal signals. Yes, it does. The JND can be conceptualized as noise in retinal and reference signals, that is, as temporal variability in the neural firing patterns of those cells whose output activity forms the physiological substrate of these signals (see, e.g., Wertheim et al. 1985). I don't know precisely what kind of noise this is. In Note 10 I assumed it to be Gaussian, but that is anyone's guess. It may well depend on the types of signals involved. This explains why, as Becker & Mergner note, the JND is different in Figures 2 and 5: in Figure 2 the JND represents noise in a retinal and a reference signal, whereas in Figure 5 it reflects noise in two retinal signals.

Thus, one cannot easily predict the magnitude of a JND from theory (apart from the indications that Weber's law

applies), especially with such compound signals as the reference signal, where each component contributes in its way to the noise level of the whole signal. I should have been more careful and warned the reader that one cannot just add or subtract noise components from each other without arriving at anomalies like those illustrated in **Gregson**'s Equation 17c (see also **Becker & Mergner**). In this respect, my equations are indeed "quasi-mathematical" (**Yardley**). However, whatever the answer to how JNDs add mathematically, this issue is itself irrelevant to the main thesis, which is that the JND can be measured and systematically and predictably affects perception. The inaccuracy implied by the concept of a JND is not so great that it prevents the perceptual system from using extraretinal signals, as suggested by **Hadani & Julesz.**

**Gregson** suggests that my vector equations describe or predict the outcome of the process, not its dynamics, and he doubts that my equations can be used to obtain a dynamic system analytical description of the process. I agree. The gain values of reference signals reported in the target article, with the exception of the one mentioned in Figure 6, should not be understood in system analytical terms. These values reflect the ratio between reference signal magnitude and physical eye velocity in space, but only at one point: where the PSS measurements took place. Thus, strictly speaking, the values may also reflect a phase shift in the modulation of reference signals, relative to eye velocity in space. This can only be figured out with experiments in which sinusoidal eye movements in space are made with many frequencies, while PSS measurements should take place at several phases of these sinusoidal eye movements in space (similar to the experimental setup that yielded Fig. 6). This may be the reason some reported gain values appear to differ from those reported by **Honda** and by **Becker & Mergner.**

Nor should **Becker & Mergner** interpret the fact that during relatively short duration of pursuit eye movements the optokinetic component is small, whereas it may grow much larger if there is no other component in the reference signal, as a difference in gain of the optokinetic component. In sections 3 and 6.5 I did mention the nonlinear contribution of the optokinetic component in reference signals. These nonlinearities are still largely unknown (which is why I applied to them no vector algebra, as **Belopolsky** notes, and why I agree with Becker & Mergner that my account of visual-vestibular interactions is insufficiently detailed). The nonlinearities may well contribute to variations in the time course of development of vection (as in **Wann & Rushton**'s example of David Lee's [1990] swinging room). Such nonlinearities might also have caused the increased "gain" of reference signals – consisting of an efference copy and an unusually large optokinetic component – with faster eye movements, in my 1987 study (as noticed by **Honda**). It is also possible that the optokinetic component always adds to the reference signal. If so, my explanation of induced motion would not be at variance with Becker & Mergner's unpublished observations about induced motion, which suggest that reference signals cannot be reduced to zero when they contain an optokinetic component.

My equations should rather be viewed as a kind of short-hand which saves space, summarizes ideas, and allows for precise quantitative predictions. I do not see why this should make **Velichkovsky & Van der Heijden** wonder whether the model "is really on the right track."

## R11. Priority

**Probst** suggests that there is no need for my model because he has already proposed a similar one. However, I considered that model (and others he and **Becker & Mergner** list) as part of the body of literature dealing mainly with ego-motion perception and visual-vestibular interactions, which I only needed to summarize briefly (see above). To discuss all existing models from this literature that may have something to say about object-motion perception, especially if this is not their main theme (Probst concedes that his is a "model of visual-vestibular interactions"), would have lengthened the target article beyond acceptable limits. A great many of these models can be found in the references I gave (see, e.g., Henn et al. 1980 for a comprehensive review). In addition, I did not list Probst's model among inferential theories, because it lacks their defining feature: an efference copy or extraretinal signal encoding eye movements.

**Skavenski** claims that my explanation of center/surround induced motion, as well as my claim that visual and vestibular signals contribute to the reference signal, is not new. He is right, of course, but not because he had already published some qualitatively similar ideas in 1990 (which I indeed overlooked, for which I apologize): he is right because I have been publishing my ideas ever since 1981, including a detailed description of my model and its explanation of induced motion (Wertheim 1987). Hints that these ideas have "come up before" are herewith gracefully beamed back to sender. Skavenski correctly suspects that I do not cite Hansen (1979) and Hansen and Skavenski (1985) because these papers do not concern motion, but position, perception. He is wrong, however, to suggest that this is unfair, given my attention to Matin's (1982; 1986; Matin et al. 1969) work on the perception of position during saccades. I referred to these three papers of Matin, not to cite his work, but because I wanted to mention the source of the term "extraretinal signal," and of the claim that it is generated from nonretinal information. Both the term and the claim had to be discussed in the target article.

## R12. The direct and inferential controversy

**Tresilian** calls it a "most bizarre" idea that the perceptual system would pick up invariants that specify eye movements from optic flow, and would, instead of using this information, generate efference copies to obtain it. I agree. The two assumptions are mutually exclusive. This is why adherents of inferential theory usually reject the "theoretical work which demonstrates" (Tresilian) that eye movement information is picked up from optic flow. Nobody accepts both assumptions at the same time.

Nevertheless, both **Shebilske** and **Yardley** state that my multidimensional concept of a reference signal is compatible with direct perception theory (Shebilske calls my approach an ecological efference mediation theory). The compatibility is attained by asserting that invariants specifying ego motion are not present *next* to reference signals, as **Tresilian** would have it, but are *used to generate* reference signals within the system responsible for object-motion perception (note the exception of invari-

ants that would specify eye movements – see the above discussion about the "visual efference copy"). This runs counter to the basic premise of direct perception theory, according to which that system only needs optic flow invariants that specify object motion. Questions such as whether the ego-motion invariants that generate reference signals are picked up exclusively from optic flow (**Kim & Turvey, Hadani & Julesz**), or consist of multidimensional combinations of sensory afferents (**Stoffregen, Tresilian**) are irrelevant. The point is that these invariants do not specify *object* motion. Even if they are encoded in reference signals, they would still have no perceptual significance because reference signals in themselves have no perceptual meaning (see above). The only way to save the direct perception assumption that object-motion perception derives exclusively from a invariant would be to assume that the interaction between retinal and reference signals yields something that could be called a "multidimensional" or "intermodal" invariant of object-motion perception within the mechanism responsible for object-motion perception (**Shebilske; Stoffregen** [?]). But that reduces the distinction between direct and inferential theory to a matter of semantics.

There are other comments on direct perception theory. Thus **Yardley, Da Vitoria Lobo,** and **Andersen** mention disagreements as to whether the perceptual system does or can perform computations on retinal or optical information (see also **Kim & Turvey,** and Ullman 1980), and one may disagree about whether or not such computations are mathematically possible (**Hadani & Julesz, Stoffregen**). These issues are irrelevant to the target article, however, as they do not concern the distinction between direct and inferential theory. This distinction is the topic of my paper: one either believes that object-motion percepts (not "motion percepts," as **Tresilian** states) derive only from visual information, or one does not believe this. I hope to have shown that, in terms of my model, these two opposite beliefs are no longer mutually exclusive.

## References


**Letters *a* and *r* appearing before authors' initials refer to target article and response respectively.**

Algom, D. & Cohen-Raz, L. (1984) Visual velocity input-output functions: The integration of distance and duration onto subjective velocity. *Journal of Experimental Psychology: Human Perception and Performance* 4:486–501. [aAHW]

Andersen, G. J. (1986) Perception of self-motion: Psychophysical and computational approaches. *Psychological Bulletin* 99:52–65. [GJA]

(1990) Segregation of optic flow into object and self-motion components: Foundations of a general theory. In: *Perception and control of self-motion*, ed. R. Warren & A. H. Wertheim. Erlbaum. [aAHW]

Andersen, G. J. & Braunstein, M. L. (1985). Induced self-motion in central vision. *Journal of Experimental Psychology: Human Perception and Performance* 11:122–32. [XMS]

Angel, R. W. & Malenka, R. C. (1982) Velocity-dependent suppression of cutaneous sensitivity during movement. *Experimental Neurology* 77:266–74. [aAHW]

Aubert, H. (1886) Die Bewegungsempfindung. *Pflügers Archiv* 39:347–70. [aAHW]

(1887) Die Bewegungsempfindung. Zweiter Mitteilung. *Pflügers Archives* 40:459–80. [aAHW]

Barthélémy, J., Xerri, L., Borel, L. & Lacour, M. (1988) Neuronal coding of linear motion in the vestibular nuclei of the alert cat. II: Response characteristics to vertical optokinetic stimulation. *Experimental Brain Research* 70:287–98. [aAHW]

Bedell, H., Klopfenstein, J. F. & Yuan, N. (1989) Extraretinal information about eye position during involuntary eye movement: Optokinetic afternystagmus. *Perception & Psychophysics* 46:579–86. [aAHW]

Belopolsky, V. I. (1978) Stability of visual world during reduction of visual field size. In: *Dvizhenie glaz i zritel'noe vospriiatie* [Eye movement and visual perception], ed. B. F. Lomov, A. A. Mit'kin & N. Yu. Vergiles. Nauka. [VIB]

(1985) Selective attention and eye movement control. *Psichologicheskii Zhurnal* [Soviet Journal of Psychology] 6(3):56–73. [VIB]

Benson, A. J. (1990) Sensory functions and limitations of the vestibular system. In: *Perception and control of self-motion*, ed. R. Warren & A. H. Wertheim. Erlbaum. [aAHW]

Berthoz, A. (1981) Intersensory interaction in motion perception. In: *Attention and performance IX*, ed. J. Long & A. Baddeley. Erlbaum. [TP]

Berthoz, A. & Droulez, J. (1982) Linear self-motion perception. In: *Tutorials on motion perception*, ed. A. H. Wertheim, W. A. Wagenaar & H. W. Leibowitz. Plenum Press. [aAHW]

Berthoz, A. & Melville-Jones, G., eds. (1985) *Adaptive mechanisms in gaze control.* Elsevier. [aAHW]

Berthoz, A., Pavard, B. & Young, L. R. (1975) Perception of linear horizontal self-motion induced by peripheral vision (linear vection): Basic characteristics and visual-vestibular interactions. *Experimental Brain Research* 23:471–89. [aAHW, XMS]

Berthoz, A., Yoshida, K. & Vidal, P. P. (1981) Horizontal eye movement signals in second-order vestibular nuclei neurons in the cat. In: Vestibular and oculomotor physiology: International meeting of the Barany Society, ed. B. Cohen. *Annals of the New York Academy of Sciences* 374:144–56. [aAHW]

Bingham, G. P. (1987) Kinematic form and scaling: Further investigations on the visual perception of lifted weight. *Journal of Experimental Psychology: Human Perception and Performance* 13:155–77. [TAS]

Bischof, N. (1974) Optic-vestibular orientation to the vertical. In: *Handbook of vestibular physiology.* Vol. 6(2):*Vestibular system* ed. H. H. Kornhuber. Springer-Verlag. [WB]

Bischof, N. & Kramer, E. (1969) Untersuchungen und Überlegungen zur Richtungswahrnehmung bei willkürlichen sakkadischen Augenbewegungen [Investigations and considerations on perception of direction during voluntary saccadic eye movements]. *Psychologische Forschung* 32(3):185–218. [BMV]

Bles, W. (1982) Stepping around: Circular vection and coriolis effects. In: *Attention and performance IX*, ed. J. Long & A. Baddeley. Erlbaum. [JJR]

Bles, W., Bos, J. E., Furrer, R., De Graaf, B., Hosman, R. J. A. W., Kortschot, H. W., Krol, J. R., Kuipers, A., Marcus, J. T., Messerschmid, E., Ockels, W. J., Oosterveld, W. J., Smit, J., Wertheim, A. H. & Wientjes, C. J. E. (1989) Space Adaptation Syndrome induced by a long duration +3Gx centrifuge run. *Institute for Perception Technical Report, IZF-1989-25.* TNO Institute for Perception, Soesterberg, The Netherlands. [aAHW]

Bles, W., Jelmorini, M., Bekkering, H. & De Graaf, B. (1994) Arthrokinetic information affects linear self-motion perception (submitted). [aAHW]

Blouin, J., Bridgeman, B., Teasdale, N., Bard, C. & Fleury, M. (submitted) Visual stability with goal-directed eye and arm movements toward a target displaced during saccadic suppression. [BB]

Bonnet, C. (1982) Thresholds of motion perception. In: *Tutorials on motion perception*, ed. A. H. Wertheim, W. A. Wagenaar & H. W. Leibowitz. Plenum Press. [aAHW]

Borah, J., Young, L. R. & Curry, R. E. (1988) Optimal estimator model for human spatial orientation. In: Representation of three-dimensional space in the vestibular, oculomotor, and visual systems, ed. B. Cohen & V. Henn. *Annals of the New York Academy of Sciences* 545. [aAHW]

Bötzel, K. & Grüsser, O. J. (1982) Horizontal and vertical circular vection and eye movements. *Neuroscience Letters Supplement* 10:86–87. [UB]

Brain, W. R. (1941) Visual disorientation with special reference to lesions of the right cerebral hemisphere. *Brain* 64:244–72. [FHP]

Brandt, T., Dichgans, J. & Büchele, W. (1974) Motion habituation: Inverted self-motion perception and optokinetic after-nystagmus. *Experimental Brain Research* 21:337–52. [UB]

Brandt, T., Dichgans, J. & Koenig, E. (1973) Differential effects of central versus peripheral vision on egocentric and exocentric motion perception. *Experiment Brain Research* 16:476–91. [aAHW, GJA, FHP]

Brandt, T., Wist, E. R. & Dichgans, J. (1975) Foreground and background in dynamic spatial orientation. *Perception & Psychophysics* 17:497–503. [FHP]

Braunstein, M. L. (1976) *Depth perception through motion.* Academic Press. [aAHW]

Bridgeman, B. & Graziano, J. A. (1989) Effect of context and efference copy on visual straight ahead. *Vision Research* 29(12):1729–36. [aAHW]

Bridgeman, B., Hendry, D. & Stark, L. (1975) Failure to detect displacement

of the visual world during saccadic eye movements. *Vision Research* 15:719–22. [aAHW]

Bridgeman, B. & Stark, L. (1991) Ocular proprioception and efference copy in registering visual direction. *Vision Research* 31:1903–13. [BB]

Bruss, A. R. & Horn, B. K. P. (1983) Passive navigation. *Computer Vision, Graphics and Image Processing* 21:3–20. [IH, JRT]

Büchele, W., Degner, D. & Brandt, T. (1980) Thresholds for object motion perception raised by concurrent head movements. *Pflügers Archives Supplement* 384:R33. [aAHW]

Buettner, U. W. & Büttner, U. (1979) Vestibular nuclei activity in the alert monkey during suppression of vestibular and optokinetic nystagmus. *Experimental Brain Research* 37:581–93. [UB]

Büttner, U. & Buettner, U. W. (1978) Parietal cortex (2v) neuronal activity in the alert monkey during natural vestibular and optokinetic stimulation. *Brain Research* 153:392–97. [aAHW, UB]

Büttner, U. & Henn, V. (1981) Circularvection: Psychophysics and single-unit recordings in the monkey. *Annals of the New York Academy of Sciences* 374:274–83. [aAHW, UB, XMS]

Cameron, E. L., Baker, C. L., Jr. & Boulton, J. C. (1992) Spatial frequency selective mechanisms underlying the motion aftereffect. *Vision Research* 32:561–68. [FHP]

Cohen, B., ed. (1981) Vestibular and oculomotor physiology: International meeting of the Barany Society. *Annals of the New York Academy of Sciences* 374. [aAHW]

Cohen, B. & Henn, V., eds. (1988) Representation of three-dimensional space in the vestibular, oculomotor, and visual systems. *Annals of the New York Academy of Sciences* 545. [arAHW]

Cohen, B., Matsuo, V. & Raphan, T. (1977) Quantitative analysis of the velocity characteristics of optokinetic nystagmus and optokinetic after-nystagmus. *Journal of Physiology* 270:321–44. [UB]

Cohen, R. L. (1965) Adaptation effects and aftereffects of moving patterns viewed in the periphery of the visual field. *Scandinavian Journal of Psychology* 6:257–64. [aAHW]

Collewijn, J. (1985) Integration of adaptive changes of the optokinetic reflex, pursuit and the vestibulo-ocular reflex. In: *Adaptive mechanisms in gaze control*, ed. A. Berthoz & G. Melvill-Jones. Elsevier. [AAS]

Collewijn, H. & Tamminga, E. P. (1984) Human smooth pursuit and saccadic eye movements during voluntary pursuit of different target motion on different background. *Journal of Physiology* 351:217–50. [HH]

Coquery, J. M. (1978) Role of active movement in control of afferent input from skin in cat and man. In: *Active touch*, ed. G. Gordon. Pergamon Press. [aAHW]

(1981) Changes in somaesthetic evoked potentials during movement. *Brain Research* 31:361–78. [aAHW]

Coquery, J. M. & Amblard, B. (1973) Backward and forward masking in the perception of cutaneous stimuli. *Perception & Psychophysics* 13(2):161–63. [aAHW]

Corbin, H. H. (1942) The perception of grouping and apparent movement in visual depth. *Archives of Psychology* 273:1–50. [BMV]

Cutting, J. E., Bruno, N., Brady, N. P. & Moore, C. (1992a) Selectivity, scope, and simplicity of models: A lesson from fitting judgments of perceived depth. *Journal of Experimental Psychology: General* 121(3):364–81. [BMV]

Cutting, J. E., Springer, K., Braren, P. A. & Johnson, S. H. (1992b) Wayfinding on foot from information in retinal, not optical, flow. *Journal of Experimental Psychology: General* 121(1):41–72. [aAHW]

da Vitoria Lobo, N. (1992) *Computation, egomotion, shape, and detecting independent motion from image motion.* Ph.D. dissertation, Department of Computer Science, University of Toronto. [NDVL]

da Vitoria Lobo, N. & Tsotsos, J. K. (1991) Telling where one is heading and where things move independently. *Proceedings of the Conference of the Cognitive Science Society*, August, Chicago, IL. [NDVL]

De Graaf, B. & Wertheim, A. H. (1988) The perception of object motion during smooth pursuit eye movements: Adjacency is not a factor contributing to the Filehne illusion. *Vision Research* 28:497–502. [arAHW]

De Graaf, B., Wertheim, A. H., Bles, W. & Kremers, J. J. M. (1990) Angular velocity and not temporal frequency determines circular vection. *Vision Research* 30(4):637–46. [aAHW, XMS]

Delorme, A. & Martin, C. (1986) Roles of retinal periphery and depth periphery in linear vection and visual control of standing in humans. *Canadian Journal of Psychology* 40:176–87. [FHP]

Denton, G. G. (1977) Visual motion after effect induced by simulated rectilinear motion. *Perception* 6:711–18. [aAHW]

Dichgans, J. & Brandt, T. (1972) Visual-vestibular interaction and motion perception. In: *Cerebral control of eye movements*, ed. J. Dichgans & E. Bizzi. Bibliotheca Opthalmologica, vol. 82. Kargel. [aAHW]

(1978) Visual-vestibular interaction: Effects on self-motion perception and postural control. In: *Handbook of sensory physiology.* Vol. 8: *Perception*, ed. R. Held, H. W. Leibowitz & H. -L. Teuber. Springer-Verlag. [arAHW, FHP, TP]

Dichgans, J., Korner, F. & Voigt, K. (1969) Vergleichende Skalierung des afferenten und efferenten Bewegungssehens beim Menschen: Linearen Funktionen mit verschiedenen Ausstiegssteilkeit. *Psychologische Forschung* 32:277–95. [aAHW]

Dichgans, J., Nauck, B. & Wolpert, E. (1973) The influence of attention, vigilance and stimulus area on optokinetic and vestibular nystagmus and voluntary saccades. In: *The oculomotor system and brain function*, ed. V. Zikmund. Butterworth. [TP]

Dichgans, J., Schmidt, C. L. & Graf, W. (1973) Visual input improves the speedometer function of the vestibular nuclei in the goldfish. *Experimental Brain Research* 18:319–22. [aAHW]

Dichgans, J., Wist, E., Diener, H. C. & Brandt, T. (1975) The Aubert-Fleischl phenomenon: A temporal frequency effect on perceived velocity in afferent motion perception. *Experimental Brain Research* 23:529–33. [aAHW]

Diener, H. C., Wist, E. R., Dichgans, J. & Brandt, T. (1976) The spatial frequency effect on perceived velocity. *Vision Research* 16:169–76. [aAHW]

DiZio, P. A. & Lackner, J. R. (1986) Perceived orientation, motion, and configuration of the body during viewing of an off-vertical, rotating surface. *Perception & Psychophysics* 39:39–46. [TAS]

Duncker, K. (1929) Über induzierte Bewegung [On induced motion]. *Psychologische Forschung* 12:180–259. [aAHW, WB, AES, BMV]

Dyhre-Poulsen, P. (1978) Perception of tactile stimuli before ballistic and during tracking movements. In: *Active touch*, ed. G. Gordon. Pergamon Press. [aAHW]

Dzhafarov, E. N. (1992) Visual kinematics. *Journal of Mathematical Psychology* 36:471–97; 498–523; 524–46. [RAMG]

Ehrenstein, W. H., Mateeff, S. & Hohnsbein, J. (1986a) Zeitliche Aspekte der Ortskonstanz bei Augenfolgbewegungen. Paper presented at the 63rd annual meeting of the German Physiological Society, March, Berlin. [aAHW]

(1986b) Temporal aspects of position constancy during ocular pursuit. *Pflügers Archives* 406:R15(no. 47). [aAHW]

(1987) Influence of exposure duration on the strength of the Filehne illusion. *Perception* 16:A29. [aAHW, SM]

Einstein, A. & Infeld, L. (1938) *The evolution of physics.* Simon & Schuster. [TAS]

Elsner, W. (1971) Power laws for the perception of rotation and the oculogyral illusion. *Perception & Psychophysics* 9(5):418–20. [aAHW]

Epstein, W. (1973) The process of "taking-into-accont" in visual perception. *Perception* 2:267–85. [SM]

Erickson, R. G. & Thier, P. (1991) A neuronal correlate of spatial stability during periods of self-induced visual motion. *Experimental Brain Research* 86:608–16. [PT]

(1992) Responses of direction-selective neurons in monkey cortex to self-induced visual motion. *Annals of the New York Academy of Sciences* 656:766–74. [PT]

Favreau, O. E. (1976) Motion after effects: Evidence for parallel processing in motion perception. *Vision Research* 16:181–86. [aAHW]

Filehne, W. (1922) Über das optische Wahrnehmen von Bewegungen. *Zeitschrift für Sinnephysiologie* 53:134–45. [aAHW, AES]

Fleischl, von E. (1882) Physiologisch-optische Notizen, 2. Mitteilung. *Sitzung Wiener Bereich der Akademie der Wissenschaften* 3(86):7–25. [aAHW]

Fletcher, W. A., Hain, T. C. & Zee, D. S. (1990) Optokinetic nystagmus and afternystagmus in human beings: Relationship to nonlinear processing of information about retinal slip. *Experimental Brain Research* 81:46–52. [aAHW]

Fuchs, A. F. & Kim, J. (1975) Unit activity in vestibular nucleus of the alert monkey during horizontal angular acceleration and eye movement. *Journal of Neurophysiology* 38:1140–61. [aAHW]

Gibson, E. J. (1991) *An odyssey in learning and perception.* MIT Press. [GER]

Gibson, J. J. (1950) *The perception of the visual world.* Houghton-Mifflin. [AES]

(1954) The visual perception of objective motion and subjective movement. *Psychological Review* 61(5):304–14. [AES]

(1966) *The senses considered as perceptual systems.* Houghton-Mifflin. [aAHW, N-GK, GER, WLS, JW, LY]

(1968) What gives rise to the perception of motion? *Psychological Review* 75:335–45. [aAHW]

(1973) Direct visual perception: A reply to Gyr. *Psychological Bulletin* 79(6):396–97. [aAHW]

(1979) *The ecological approach to visual perception.* Houghton Mifflin. [aAHW, VIB, N-GK, GER, BMV, JW, LY]

Gibson, J. J., Olum, P. & Rosenblatt, F. (1955) Parallax and perspective during aircraft landings. *American Journal of Psychology* 68:372–85. [JRT]

Gibson, J. J., Smit, O. W., Steinschneider, A. & Johnson, C. W. (1957) The relative accuracy of visual perception of motion during fixation and pursuit. *American Journal of Psychology* 70:64–68. [aAHW]

Gogel, W. C. (1981) Perceived depth is a necessary factor in apparent motion concomitant with head motion: A reply to Shebilske and Proffitt. *Perception & Psychophysics* 29:173–77. [rAHW]

(1982) Analysis of the perception of retinal motion concomitant with a lateral motion of the head. *Perception & Psychophysics* 32:241–50. [MS]

(1990) A theory of phenomenal geometry and its applications. *Perception & Psychophysics* 48:105–23. [rAHW, MS]

Gogel, W. C., Loomis, J. M., Newman, N. J. & Sharkey, T. J. (1985) Agreement between indirect measures of perceived distance. *Perception & Psychophysics* 37:17–27. [MS]

Gogel, W. C. & Sharkey, T. J. (1989) Measuring attention using induced motion. *Perception* 18:303–20. [SM]

Gogel, W. C. & Tietz, J. D. (1992) Determinants of the perception of sagittal motion. *Perception & Psychophysics* 52:75–96. [MS]

Goodale, M. A. (1988) Modularity in visuomotor control: From input to output. In: *Computational processes in human vision: An interdisciplinary perspective*, ed. Z. W. Pylyshyn. Ablex. [WLS]

Goodale, M. A. & Milner, A. D. (1992) Separate visual pathways for perception and action. *Trends in Neurosciences* 15:20–25. [XMS]

Grant, W. & Best, W. (1987) Otolith-organ mechanics: Lumped parameter model and dynamic response. *Aviation Space and Environmental Medicine* 58:970–76. [aAHW]

Graybiel, A. (1952) Oculogravic illusion. *Archives of Ophthalmology* 48:605. [TAS]

Graybiel, A. & Brown, R. H. (1951) The delay in visual reorientation following exposure to a change in direction of resultant force on a human centrifuge. *Journal of General Psychology* 45:143–50. [HER]

Graybiel, A. & Hupp, E. D. (1946) The oculogyral illusion: A form of apparent motion which may be observed following stimulation of the semi-circular canals. *Journal of Aviation Medicine* 17:3–27. [aAHW]

Graybiel, A. & Niven, J. (1951) The effect of a change in direction of resultant force on sound localization: The audiogravic illusion. *Journal of Experimental Psychology* 42:227–30. [TAS]

Gregory, R. L. (1958) Eye movements and the stability of the visual world. *Nature* 182:1214–16. [HER, AES]

Grüsser, O. -J. (1983) Multimodal structure of the extrapersonal space. In: *Spatially oriented behavior*, ed. A. Hein & M. Jeannerod. Springer-Verlag. [FHP]

Grüsser, O. -J., Pause, M. & Schreiter, U. (1990a) Localization and responses of neurons in the parieto-insular vestibular cortex of awake monkeys (*Macaca fascicularis*). *Journal of Physiology* 430:537–57. [UB]

(1990b) Vestibular neurons in the parieto-insular cortex of monkeys (*Macaca fascicularis*): Visual and neck receptor responses. *Journal of Physiology* 430:559–83. [UB]

Guedry, F. E. (1974) Psychophysics of vestibular sensation. In: *Handbook of sensory physiology*, vol. 6/2, ed. H. H. Kornhuber. Springer. [rAHW]

Guldin, W. O., Akbarian, S. & Grüsser, O. J. (1992) Cortico-cortical connections and cytoarchitectonics of the primate vestibular cortex: A study in squirrel monkeys (*Saimiri sciureus*). *Journal of Comparative Neurology* 326:375–401. [UB]

Guldin, W., Mirring, S. & Grüsser, O. J. (1992) Monosynaptic input from the cerebral cortex to the vestibular brainstem nuclei in the rat. *Society of Neuroscience Abstracts* 18:510. [UB]

Gyr, J. W. (1972) Is a theory of direct visual perception adequate? *Psychological Bulletin* 77:246–61. [aAHW]

Hadani, I., Ishai, G., Frisch, H. I. & Kononov, A. (1993) Two metric solutions to 3-D reconstruction for an eye in pure rotations. *Journal of the Optical Society of America* (in press). [IH]

Hadani, I., Ishai, G. & Gur, M. (1980) Visual stability and space perception in monocular vision: Mathematical model. *Journal of the Optical Society of America* 1:60–65. [IH]

Hadani, I. & Kononov, A. (1993) Passive navigation for an eye in six degrees of freedom. *Science* (in preparation). [IH]

Hansen, R. H. (1979) Spatial localization during pursuit eye movements. *Vision Research* 19:1213–21. [rAHW, AAS]

Hansen, R. H. & Skavenski, A. A. (1985) Accuracy of spatial localizations near the time of saccadic eye movements. *Vision Research* 25:1077–82. [rAHW, AAS]

Heckmann, T. & Howard, I. P. (1991) Induced motion: Isolation and dissociation of egocentric and vection-entrained components. *Perception* 20:285–305. [FHP]

Heckmann, T. & Post, R. B. (1988) Induced motion and optokinetic afternystagmus: Parallel response dynamics with prolonged stimulation. *Vision Research* 28:681–94. [FHP]

Helmholtz, H. (1910) *Handbuch der physiologischen Optik*, vol. 3. Voss. [aAHW]

Henderson, D. C. (1971) The relationship among time, distance, and intensity as determinants of motion discrimination. *Perception & Psychophysics* 10:310–20. [aAHW]

Henn, V., Cohen, B. & Young, L. R. (1980) Visual-vestibular interaction in motion perception and the generation of nystagmus. *Neurosciences Research Program Bulletin* 18(4). MIT Press. [arAHW, TP]

Henn, V., Young, L. R. & Finley, C. (1974) Vestibular nucleus units in alert monkeys are also influenced by moving visual fields. *Brain Research* 71:144–49. [aAHW]

Hofstadter, D. R. (1980) *Gödel, Escher Bach: An eternal golden braid.* Vintage Press. [aAHW]

Hofstetter-Degen, K. (1988) *Eine psychophysische Untersuchung zur visuell-vistibulären Interaktion. Beeinflussung der Objektbewegungswahrnehmung durch gleichzeitige Eigenbewegung.* Ph.D. dissertation, University of Mainz. (Author's transl.: A psychophysical investigation of visual-vestibular interaction. Influence of object-motion perception by simultaneous self-motion.) [TP]

Honda, H. (1990) The extraretinal signal from the pursuit-eye-movement system: Its role in the perceptual and the egocentric localization systems. *Perception & Psychophysics* 48:509–15. [HH]

Howard, I. P. (1982) *Human visual orientation.* Wiley. [aAHW, VIB, DC]

Howard, I. P. & Heckmann, T. (1989) Circular vection as a function of the relative sizes, distances, and positions of two competing visual displays. *Perception* 18:657–65. [XMS]

Howard, I. P. & Marton, C. (1992) Visual pursuit over textured backgrounds in different depth planes. *Experimental Brain Research* 90:625–29. [FHP]

Howard, I. P. & Templeton, W. B. (1966) *Human spatial orientation.* Wiley. [aAHW]

Hunzelmann, N. & Spillmann, L. (1984) Movement adaptation in the peripheral retina. *Vision Research* 24(12):1765–69. [aAHW]

Ilg, U. & Thier, P. (1993) Inability of visual area V1 of the awake rhesus monkey to discriminate between self-induced and externally-induced retinal image slip. *Society for Neuroscience Abstracts* 19:629. [PT]

Ioannou, P. A. & Kokotovic, P. V. (1983) Adaptive systems with reduced models. *Lecture notes in control and information sciences* 47. Springer-Verlag. [RAMG]

Ito, M. (1982) Cerebellar control of the vestibulo-ocular reflex: Around the flocculus hypothesis. *Annual Review of Neuroscience* 5:275–96. [aAHW]

(1987) Oculomotor system, mechanisms. In: *Encyclopedia of neuroscience*, vol. 2, ed. G. Adelman. Birkhäuser. [TP]

Jeannerod, M., Kennedy, H. & Magnin, M. (1979) Corollary discharge: Its possible implications in visual and oculomotor interactions. *Neuropsychologia* 17:241–58. [aAHW]

Johansson, G. (1982) Visual space perception through motion. In: *Tutorials on motion perception*, ed. A. H. Wertheim, W. A. Wagenaar & H. W. Leibowitz. Plenum Press. [AES]

Johnson, C. A. & Leibowitz, H. W. (1976) Velocity-time reciprocity in the perception of motion: Foveal and peripheral determinations. *Vision Research* 16:177–80. [aAHW]

Johnstone, J. R. & Mark, R. F. (1970) Two classes of eye movement and their perceptual consequences. *Proceedings of the Australian Physiological and Pharmacological Society* 1(2):46–47. [aAHW]

(1971) The efference copy neurone. *Journal of Experimental Biology* 54:403–14. [aAHW]

(1973) Corollary discharge. *Vision Research* 13:1621. [aAHW]

Jung, R. (1972) Neurophysiological and psychophysical correlates in visual research. In: *Brain and human behavior*, ed. A. G. Karczmar & J. C. Eccles. Springer. [VIB]

Kano, C. (1970) Changes of threshold of continuous visual movement by variation of phenomenal size with invariant retinal size. *Psychological Research* 33(3):242–53. [BMV]

Kaufman, L. (1974) *Sight and mind: An introduction to visual perception.* Oxford University Press. [BMV]

Kinchla, R. A. (1971) Visual movement perception: A comparison between absolute and relative movement discrimination. *Perception & Psychophysics* 9(2A):165–71. [aAHW]

Koenderink, J. J. (1990) Some theoretical aspects of optic flow. In: *Perception and control of self-motion*, ed. R. Warren & A. H. Wertheim. Erlbaum. [aAHW]

Koenderink, J. J. & van Doorn, A. J. (1976) Local structure of movement

parallax of the plane. *Journal of the Optical Society of America* 66:717–23. [GJA]
(1981) Exterospecific component of the motion parallax field. *Journal of the Optical Society of America* 71:953–57. [JRT]
(1987) Facts on optic flow. *Biological Cybernetics* 56:247–54. [aAHW, JRT]
(1991) Affine structure from motion. *Journal of the Optical Society of America* 8(2):377–85. [IH]
Köhler, W. (1947) *Gestalt psychology: An introduction to new concepts in modern psychology.* Liveright. [BMV]
Kornhuber, H. H. (1974) Nystagmus and related phenomena in man: An outline of otoneurology. In: *Handbook of sensory physiology.* Vol. 6/2, *Vestibular system,* part 2, *psychophysics. Applied aspects and general interpretations.* Springer Verlag. [aAHW]
Lackner, J. R. (1992) Spatial orientation in weightless environments. *Perception* 21:803–12. [FHP]
Lackner, J. & DiZio, P. (1984) Some efferent and somatosensory influences on body orientation and oculomotor control. In: *Sensory experience, adaptation and perception: Festschrift for Ivo Kohler,* ed. L. Spillman & B. Wooten. Erlbaum. [JJR]
Lackner, J. R. & Teixeira, R. A. (1977) Optokinetic motion sickness: Continuous head movements attenuate the visual induction of apparent self-rotation and symptoms of motion sickness. *Aviation, Space and Environmental Medicine* 48:248–53. [LY]
Lang, W., Büttner-Ennever, J. A. & Büttner, U. (1979) Vestibular projections to the monkey thalamus: An autoradiographic study. *Brain Research* 177:3–17. [UB]
Larsen, A., Farrell, J. E. & Bundesen, C. (1983) Short- and long-range processes in visual apparent movement. *Psychological Research* 45:11–18. [BMV]
Lee, D. & Aronson, E. (1974) Visual proprioceptive control of standing in human infants. *Perception & Psychophysics* 15:529–32. [JJR]
Lee, D. N. (1974) Visual information during locomotion. In: *Perception: Essays in honour of James J. Gibson,* ed. R. B. McLeod & H. Pick. Cornell University Press. [JRT]
(1990) Getting around with light and sound. In: Perception and control of self-motion, ed. R. Warren & A. H. Wertheim. Erlbaum. [rAHW]
Leibowitz, H. W., Post, R. B., Brandt, T. & Dichgans, J. (1982) Implications of recent developments in dynamic spatial orientation and visual resolution for vehicle guidance. In: *Tutorials on motion perception,* ed. A. H. Wertheim, W. A. Wagenaar & H. W. Leibowitz. Plenum Press. [aAHW]
Lisberger, S. G. & Fuchs, A. F. (1978a) Role of primate flocculus during rapid behavioral modification of vestibulo-ocular reflex I. Purkinje cell activity during visually guided horizontal smooth pursuit eye movements and passive head rotation. *Journal of Neurophysiology* 41(3):733–63. [aAHW]
(1978b) Role of primate flocculus during rapid behavioral modification of vestibulo-ocular reflex II. Mossy fiber firing patterns during horizontal head rotation and eye movement. *Journal of Neurophysiology* 41(3):764–77. [aAHW]
Lishman, J. R. & Lee, D. N. (1973) The autonomy of visual kinaesthesis. *Perception* 2:287–94. [N-GK, JW]
Lombardo, T. J. (1987) *The reciprocity of perceiver and environment: The evolution of J. J. Gibson's ecological psychology.* Erlbaum. [aAHW]
Longuet-Higgins, H. C. & Prazdny, K. (1980) The interpretation of a moving retinal image. *Proceedings of the Royal Society of London (B)* 208:385–97. [GJA, JRT]
Mack, A. (1978) Three modes of visual perception. In: *Models of perceiving and processing information,* ed. H. L. Pick & E. Saltzman. Erlbaum. [aAHW, AES]
(1986) Perceptual aspects of motion in the frontal plane. In: *Handbook of perception and human performance.* Vol. I: *Sensory processes and perception,* ed. K. R. Boff, L. Kaufman & J. P. Thomas. Wiley. [aAHW]
Mack, A. & Herman, E. (1972) A new illusion: The underestimation of distance during smooth pursuit eye movements. *Perception & Psychophysics* 12(6):471–73. [aAHW, HH]
(1973) Position constancy during pursuit eye movement: An investigation of the Filehne illusion. *Quarterly Journal of Experimental Psychology* 25:71–84. [aAHW]
(1978) The loss of position constancy during pursuit eye movements. *Vision Research* 18:55–62. [aAHW]
MacKay, D. M. (1972) Voluntary eye movements as questions. In: *Cerebral control of eye movements,* ed. J. Dichgans & E. Bizzi. Bibliotheca Opthalmologica, vol. 82. Kargel. [aAHW]
(1973) Visual stability and voluntary eye movements. In: *Handbook of sensory physiology,* vol. 7(3A), ed. R. Jung. Springer. [aAHW, VIB]
(1982) Anomalous perception of extrafoveal motion. *Perception* 11:359–60. [aAHW]
Maioli, M. G., Squatrito, S. & Domeniconi, R. (1989) Projections from visual cortical areas of the superior temporal sulcus to the lateral terminal nucleus of the accessory optic system in macaque monkeys. *Brain Research* 498:389–92. [XMS]
Marcus, J. T. (1992) *Vestibulo-ocular responses in man to gravito-inertial forces.* Ph.D. dissertation. TNO Institute for Perception, Soesterberg. [aAHW]
Marendaz, C., Stivalet, P., Barraclough, L. & Walkowiac, P. (1993) Effect of gravitational cues on visual search for orientation. *Journal of Experimental Psychology: Human Perception and Performance* 19(6):1266–77. [LY]
Mark, L. M. (1987) Eye-height-scaled information about affordances: A study of sitting and stair-climbing. *Journal of Experimental Psychology: Human Perception and Performance* 13:683–783. [TAS]
Massaro, D. & Cohen, M. M. (1993) The paradigm and the fuzzy logical model of perception are alive and well. *Journal of Experimental Psychology: General* 122(1):115–24. [BMV]
Mateeff, S. (1980) Visual perception of movement patterns during smooth eye tracking. *Acta Physiologica et Pharmacologica Bulgarica* 6:82–89. [SM]
Mateeff, S., Ehrenstein, W. H. & Hohnsbein, J. (1987) Constancy of visual direction requires time to develop. *Perception* 16:A29. [SM]
Mateeff, S., Yakimoff, N., Hohnsbein, J. & Ehrenstein, W. H. (1991) Perceptual constancy during ocular pursuit: A quantitative estimation procedure. *Perception & Psychophysics* 49(4):390–92.[aAHW]
Matin, L. (1982) Visual localization and eye movements. In: *Tutorials on motion perception,* ed. A. H. Wertheim, W. A. Wagenaar & H. W. Leibowitz. Plenum Press. [arAHW]
(1986) Visual localization and eye movements. In: *Handbook of perception and human performance.* Vol. I: *Sensory processes and perception,* ed. K. R. Boff, L. Kaufman & J. P. Thomas. Wiley. [arAHW]
Matin, L., Matin, E. & Pearce, D. (1969) Visual perception of direction when voluntary saccades occur. I. Relation of visual direction of a fixation target extinguished before a saccade to a subsequent test flash presented during the saccade. *Perception & Psychophysics* 5:65–80. [arAHW, AAS]
McConkie, A. B. & Farber, J. M. (1979) Relation between perceived depth and perceived motion in uniform flow fields. *Journal of Experimental Psychology: Human Perception and Performance* 5:501–8. [GJA]
McCrea, R. A., Yoshida, K., Evinger, C. & Berthoz, A. (1981) The location, axonal arborization and termination sites of eye movement-related secondary vestibular neurons demonstrated by intra-axonal HRP injection in the alert cat. In: *Progress in oculomotor research,* ed. A. F. Fuchs & W. Becker. Elsevier North-Holland. [aAHW]
Mergner, T. & Becker, W. (1990) Perception of horizontal self-rotation: Multisensory and cognitive aspects. In: *Perception and control of self-motion,* ed. R. Warren & A. H. Wertheim. Erlbaum. [arAHW, WB, N-GK, SM, FHP]
Mergner, T., Rottler, G., Kimmig, H. & Becker, W. (1992) Role of vestibular and neck inputs for the perception of object motion in space. *Experimental Brain Research* 89:655–68. [WB]
Mergner, T., Siebold, C., Schweigart, G. & Becker, W. (1991) Human perception of horizontal head and trunk rotation in space during vestibular and neck stimulation. *Experimental Brain Research* 85:389–404. [WB]
Metzger, W. (1941) *Psychologie* [Psychology]. Steinkopf. [BMV]
Miller, J. (1980) Information used by the perceptual and oculomotor systems regarding the amplitude of saccadic and pursuit eye movements. *Vision Research* 20:59–68. [HH]
Mittelstaedt, H. (1983) A new solution to the problem of the subjective vertical. *Naturwissenschaften* 70:272–81. [LY]
(1990) Basic solutions to the problem of head-centric visual localization. In: *Perception and control of self-motion,* ed. R. Warren & A. H. Wertheim. [aAHW]
Moulden, B. (1975) Eye movements and the movement after effect. *Vision Research* 15:1169–70. [aAHW]
Murphy, B. J. (1978) Pattern thresholds for moving and stationary gratings during smooth pursuit eye movement. *Vision Research* 18:521–30. [aAHW]
Nagao, S. (1988) Behavior of floccular Purkinje cells correlated with adaptation of horizontal optokinetic eye movement response in pigmented rabbits. *Experimental Brain Research* 73:489–97. [aAHW]
Nakayama, K. (1981) Differential motion hyperacuity under conditions of common image motion. *Vision Research* 21:1475–82. [aAHW]
Neisser, U. (1976) *Cognition and reality.* W. H. Freeman. [GER]
Noda, H. (1986) Mossy fibers sending retinal-slip, eye, and head velocity signals to the flocculus of the monkey. *Journal of Physiology* 379:39–60. [aAHW]
Ockels, W. J., Furrer, R. & Messerschmid, E. (1989) Space sickness on earth. *Nature* 340(August):681–82. [aAHW]

(1990) Space sickness on earth. *Experimental Brain Research* 79(3):661–63. [aAHW]

Ohmi, M., Howard, I. P. & Landolt, J. P. (1987) Circular vection as a function of foreground-background relationships. *Perception* 16:17–22. [N-GK, FHP, LY]

Oman, C. M. (1988) The role of static visual orientation cues in the etiology of space motion sickness. In: *Proceedings of the Symposium on Vestibular Organs and Altered Force Environment*, ed. M. Igarashi & K. G. Nute. National Aeronautics and Space Administration. [FHP]

Ono, M. E., Rivest, J. & Ono, H. (1986) Depth perception as a function of motion parallax and absolute-distance information. *Journal of Experimental Psychology: Human Perception & Performance* 12:331–37. [MS]

Owen, D. H. (1990) Perception and control of changes in self-motion: A functional approach to the study of information and skill. In: *Perception and control of self-motion*, ed. R. Warren & A. H. Wertheim. Erlbaum. [aAHW]

Paillard, J., Brouchon-Viton, M. & Jordan, P. (1978) Differential encoding of location cues by active and passive touch. In: *Active touch*, ed. G. Gordon. Pergamon Press. [aAHW]

Pavard, B. & Berthoz, A. (1977) Linear acceleration modifies the perception of a moving visual scene. *Perception* 6:529–40. [aAHW]

Post, R. B. & Leibowitz, H. W. (1985) A revised analysis of the role of efference in motion perception. *Perception* 14:631–43. [arAHW, SM, AAS]

Post, R. B., Leibowitz, H. W. & Sheehy, J. B. (1986) Efference, perceived movement, and illusory displacement. *Acta Psychologica* 63(1–3):23–34. [BMV]

Precht, W. (1982) Anatomical and functional organization of optokinetic pathways. In: *Functional basis of ocular motility disorders*, ed. G. Lennestrand, D. S. Lee & E. L. Keller. Pergamon Press. [aAHW]

Previc, F. H. & Donnelly, M. (1993) The effects of visual depth and eccentricity on manual bias, induced motion, and vection. *Perception* 22. [FHP]

Previc, F. H., Varner, D. C. & Gillingham, K. K. (1992) Visual scene effects on the somatogravic illusion. *Aviation, Space, and Environmental Medicine* 63:1060–64. [FHP]

Probst, T. (1983) *Beeinflussung der Objektbewegungswahrnehmung durch gleichzeitige Eigenbewegungsempfindung. Psychophysische Grundlagen und angewandte Aspekte bei der Fahrzeugsteuerung.* Ph.D. dissertation, University of Essen. (Author's transl.: Influence of object-motion detection by concurrent self-motion perception. Basic psychophysics and applied aspects for vehicle guidance.) [TP]

(1991) *Intersensorische Interaktion beim Menschen. Psychophysikalische und elektrophysiologische Untersuchungen.* Waxmann Münster. (Author's transl.: Intersensory interaction in humans. Psychophysical and electrophysiological investigations.) [TP]

Probst, T., Brandt, T. & Degner, D. (1986) Object-motion detection affected by concurrent self-motion perception: Psychophysics of a new phenomenon. *Behavioural Brain Research* 22:1–11. [aAHW, TP]

Probst, T., Degner, D. & Brandt, T. (1980) Object motion perception affected by concurrent self-motion. Paper presented at the third European Conference on Visual Perception. August, Brighton, UK. [aAHW]

Probst, T., Krafczyk, S., Brandt, T. & Wist, E. R. (1984) Interaction between perceived self-motion and object-motion impairs vehicle guidance. *Science* 225:536–38. [aAHW]

Probst, T., Straube, A. & Bles, W. (1985) Differential effects of ambient visual-vestibular-somatosensory stimulation on the perception of self-motion. *Behavioural Brain Research* 16:71–79. [aAHW]

Puckett, J. W. & Steinman, R. M. (1969) Tracking eye movements with and without saccadic correction. *Vision Research* 9:695–703. [HH]

Raphan, T., Cohen, D. & Matsuo, V. (1977) A velocity storage mechanism responsible for OKN, OKAN and vestibular nystagmus. In: *Control of gaze by brainstem neurons. Developments in neuroscience*, vol. 1, ed. R. Baker & A. Berthoz. Elsevier North Holland, Biomedical Press. [aAHW]

Rauch, R., Angel, R. W. & Boylls, C. C. (1985) Velocity-dependent suppression of somatosensory evoked potentials during movement. *Encephalography and Clinical Neurophysiology* 62:421–25. [aAHW]

Raymond, J. W., Shapiro, K. L. & Rose, D. J. (1984) Optokinetic backgrounds affect perceived velocity during ocular tracking. *Perception & Psychophysics* 36(3):221–24. [aAHW]

Reinhardt-Rutland, A. H. (1992) Does the type of eye motion determine whether induced motion is diminished or enhanced? *Perceptual and Motor Skills* 74:882. [aAHW]

Riccio, G. E. (1993a) Information in movement variability about the qualitative dynamics of posture and orientation. In: *Variability and motor control*, ed. K. M. Newell & D. M. Corcos. Human Kinetics. [GER]

(1993b) *Multimodal perception and multicriterion control of nested systems: Self-motion in real and virtual environments* (UIUC-BI-HPP-93-02). Beckman Institute for Advanced Science and Technology, University of Illinois. [GER]

Riccio, G. E., Martin, E. J. & Stoffregen, T. A. (1992) The role of balance dynamics in the active perception of orientation. *Journal of Experimental Psychology: Human Perception & Performance* 18:624–44. [TAS]

Riccio, G. E. & Stoffregen, T. A. (1990) Gravitoinertial force versus the direction of balance in the perception and control of orientation. *Psychological Review* 97:135–37. [TAS]

Rieser, J. J., Ashmead, D. A., Talor, C. & Youngquist, G. (1991) Visual perception and the guidance of locomotion without vision to previously seen targets. *Perception* 19:675–89. [JJR]

Rieser, J., Garing, A. & Young, M. (1994) Imagery, action and young children's spatial orientation: It's not being there that counts, it's what one has in mind. *Child Development* (in press). [JJR]

Rieser, J., Pick, H., Ashmead, D. & Garing, A. (submitted) The calibration of human locomotion and models of perceptual-motor organization. [JJR]

Rock, I. (1977) In defense of unconscious inference. In: *Stability and constancy in visual perception: Mechanisms and processes*, ed. W. Epstein. Wiley. [SM, AES]

(1983) *The logic of perception*. MIT Press. [SM]

Rogers, B. & Graham, M. (1979) Motion parallax as an independent cue for depth perception. *Perception* 8:125–34. [MS]

Rosengren, K. S., Pick, H. & von Hofsten, C. (1988) The role of visual information in ball catching. *Journal of Motor Behaviour* 20:150–64. [JRT]

Ross, H. E. (1974) *Behaviour and perception in strange environments*. Allen & Unwin. [aAHW, HER]

(1976) The direction of apparent movement during transient pressure vertigo. *Undersea Biomedical Research* 3:403–10. [HER]

(1990) Orientation and movement in divers. In: *Perception and control of self-motion*, ed. R. Warren & A. H. Wertheim. Erlbaum. [HER]

Ross, H. E. & Lennie, P. (1968) Visual stability during bodily movement underwater. *Underwater Association Report* 3:55–57. [HER]

Runeson, S. (1988) The distorted room illusion, equivalent configurations, and the specificity of static optic arrays. *Journal of Experimental Psychology: Human Perception and Performance* 14:295–304. [LY]

Sauvan, X. M. & Bonnet, C. (1988) Thresholds and variations of the forward rectilinear and curvilinear vections in man. In: *Proceedings of the European Brain and Behaviour Society Workshop "Visual processing of form and motion,"* March, Tübingen, Germany. [XMS]

(1989) Les sensations de déplacement curvilinéaire générées visuellement. *Psychologie Française* 34:19–24. [XMS]

(1993) Properties of curvilinear vection. *Perception & Psychophysics* 53:429–35. [TAS, XMS]

Schöne, H. (1984) *Spatial orientation: The spatial control of behavior in animals and man* (trans. C. Strausfeld). Princeton University Press. [TAS]

Schoppmann, A. (1981) Projections from areas 17 and 18 of the visual cortex to the nucleus of the optic tract. *Brain Research* 223:1–17. [TP]

Sekuler, A. B. (1990) Motion segregation from speed differences: Evidence for nonlinear processing. *Vision Research* 30(5):785–95. [aAHW]

Sekuler, R. R., Ball, K., Tynan, P. & Machamer, J. (1982) Psychophysics of motion perception. In: *Tutorials on motion perception*, ed. A. H. Wertheim, W. A. Wagenaar & H. W. Leibowitz. Plenum Press. [aAHW]

Shaffer, O. & Wallach, H. (1966) Extent-of-motion thresholds under subject-relative and object-relative conditions. *Perception & Psychophysics* 1:447–51. [aAHW]

Shaw, R. E., Kadar, E., Sim, M. & Repperger, D. W. (1992) The intentional spring: A strategy for modeling systems that learn to perform intentional acts. *Motor Behavior* 24:3–28. [GER]

Shebilske, W. L. (1977) Visuomotor coordination in visual direction and position constancies. In: *Stability and constancy in visual perception: Mechanisms and processes*, ed. W. Epstein. Wiley. [WLS]

(1984) Efferent factors in cognition and perception. In: *Cognition and motor processes*, ed. W. Prinz & A. F. Sanders. Plenum Press. [WLS]

(1987a) An ecological efference mediation theory of natural event perception. In: *Perspectives on perception and action*, ed. W. Prinz & A. F. Sanders. Erlbaum. [WLS]

(1987b) Baseball batters support an ecological efference mediation theory of natural event perception. In: *Sensorimotor interactions in space perception and action*, ed. D. G. Bouwhuis, B. Bridgeman, D. A. Owens, W. L. Shebilske & P. Wolff. North Holland. [WLS]

(1990) Visuomotor modularity, ontogeny, and training high-performance skills with spatial display instruments. In: *Spatial displays and spatial instruments*, ed. S. R. Ellis & M. K. Kaiser. Erlbaum. [WLS]

Shebilske, W. L. & Peters, P. (in press) Perception, action, and constancy. In: *Handbook of perception*, ed. W. Prinz & B. Bridgeman. Springer-Verlag. [WLS]

Shebilske, W. L., Proffitt, D. R. & Fisher, S. K. (1984) Efferent factors in natural event perception can be rationalized and verified: A reply to Turvey and Solomon. *Journal of Experimental Psychology: Human Perception and Performance* 10:455–60. [WLS]

Shulman, P. H. (1979) Eye movements do not cause induced motion. *Perception & Psychophysics* 26:381–83. [aAHW]

Simpson, J. & Graf, W. (1985) The selection of reference frames by nature and its investigators. In: *Adaptive mechanisms in gaze control*, ed. A. Berthoz & G. Melvill-Jones. Elsevier. [AAS]

Skavenski, A. A. (1972) Inflow as a source of extraretinal eye position information. *Vision Research* 12:221–29. [aAHW]

(1990) Eye movement and visual localization of objects in space. In: *Eye movements and their role in visual and cognitive processes*, ed. E. Kowler. Elsevier. [rAHW, AAS]

Sperry, R. W. (1950) Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology* 43:482–89. [aAHW]

Stark, L. & Bridgeman, B. (1983) Role of corollary discharge in space constancy. *Perception & Psychophysics* 34(4):371–80. [aAHW]

Steinbach, M. J. (1987) Proprioceptive knowledge of eye position. *Vision Research* 27(10):1737–44. [aAHW]

Stern, L. D. & Emelity, D. (1978) Evidence for frames of reference based on pursuit eye movements. *Perception & Psychophysics* 24:521–28. [rAHW, SM, AS]

Stoffregen, T. A. (1985) Flow structure versus retinal location in the optical control of stance. *Journal of Experimental Psychology: Human Perception and Performance* 11:554–65. [TAS]

(1986) The role of optical velocity in the control of stance. *Perception & Psychophysics* 39:355–60. [TAS]

(1990) Multiple sources of information: For what? *Newsletter of the International Society for Ecological Psychology* 4:5–8. [TAS]

Stoffregen, T. A. & Riccio, G. E. (1988) An ecological theory of orientation and the vestibular system. *Psychological Review* 95:3–14. [aAHW, TAS, LY]

(1990) Responses to optical looming in the retinal center and periphery. *Ecological Psychology* 2:251–74. [TAS]

(1991) An ecological critique of the sensory conflict theory of motion sickness. *Ecological Psychology* 3:159–94. [TAS]

Stone, L. S. & Lisberger, S. G. (1990a) Visual responses of Purkinje cells in the cerebellar flocculus during smooth pursuit eye movements in monkeys. I: Simple spikes; II: Complex spikes. *Journal of Neurophysiology* 63(5):1241–75. [aAHW]

(1990b) Synergistic action of complex and simple spikes in the monkey flocculus in the control of smooth pursuit eye movement. *Experimental Brain Research* 17: 299–312. [aAHW]

Stoper, A. (1967) *Vision during pursuit eye movement: The role of oculomotor information*. Ph.D. dissertation, Brandeis University. University Microfilms, No. 67-16, 579. [AES]

(1973) Apparent motion of stimuli presented stroboscopically during pursuit movement of the eye. *Perception & Psychophysics* 23:201–11. [AES]

Straube, A. & Brandt, T. (1987) Importance of the visual and vestibular cortex for self-motion perception in man (circularvection). *Human Neurobiology* 6:211–18. [aAHW, UB, TP]

Straube, A., Paulus, W. & Brandt T. (1990) Influence of visual blue on object-motion detection, self-motion detection and postural balance. *Behavioural Brain Research* 40;1–6. [UB]

Suzuki, D. A. & Keller, E. L. (1988) The role of the posterior vermis of monkey cerebellum in smooth pursuit eye movement control. II. Target velocity-related Purkinje cell activity. *Journal of Neurophysiology* 59:19–40. [UB]

Swanston, M. T. & Wade, N. J. (1988) The perception of visual motion during movements of the eyes and of the head. *Perception & Psychophysics* 43:559–66. [aAHW, MS]

(1992a) The interaction of perceived distance with the perceived direction of visual motion during movements of the eyes and of the head. *Perception & Psychophysics* 52:705–13. [MS]

(1992b) Motion over the retina and the motion aftereffect. *Perception* 21:569–82. [MS]

Swanston, M. T., Wade, N. J. & Day, R. H. (1987) The representation of uniform motion in vision. *Perception* 16:143–60. [aAHW, MS]

Swanston, M. T., Wade, N. J. & Ono, H. (1990) The binocular representation of uniform motion. *Perception* 19:29–34. [MS]

Tanaka, K., Hikosaka, K., Saito, H. -A., Yukie, M., Fukada, Y. & Iwai, E. (1986) Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *Journal of Neuroscience* 6:134–44. [PT]

Telford, L., Spratley, J. & Frost, B. J. (1992) Linear vection in the central visual field facilitated by kinetic depth cues. *Perception* 21:337–49. [XMS]

Thompson, S. P. (1879) Some new optical illusions. *Journal of Science* 9:373–74. [HER]

Tsai, R. Y. & Huang, T. S. (1985) Uniqueness and estimation of 3-D motion parameters and surface structures of rigid objects. In: *Image understanding*, ed. W. Richards & S. Ullman. Albex. [IH]

Turvey, M. T. (1992) Affordances and prospective control: An outline of the ontology. *Ecological Psychology* 4:173–88. [GER]

Tyler, C. W. & Foley, J. M. (1974) Stereomovement suppression for transient disparity changes. *Perception* 3(3):287–96. [BMV]

Ullman, S. (1980) Against direct perception. *Behavioral and Brain Sciences* 3:373–415. [arAHW, JRT]

Van de Grind, W. A., Koenderink, J. J. & van Doorn, A. J. (1992) Viewing-distance invariance of movement detection. *Experimental Brain Research* 91:135–50. [rAHW]

Velichkovsky, B. M. (1971) Autokineticheskaja illusia i intermodalnije otnoshenija v zritel'nom vosprijatii dvizhenija [Autokinetic illusion and intermodal relationships in visual perception of motion]. *Ergonomika* 2:182–96. [BMV]

(1973) O roli prostranstvennykh system otscheta v vosprijatii sobstvennogo i objectnogo dvizhenija [On the role of spatial frames of reference in perception of self-motion and object's motion]. *Voprosy Psykhologii* 18(2):15–26. [BMV]

(1982) Funktional'naja struktura perceptivnikh processov [Functional structure of perceptual processes]. In: *Osnovy psykhologii: Poznavatel'nije processy*, ed. A. A. Smirnov. Pedagogika. [BMV]

Von Holst, E. (1954) Relations between the central nervous system and peripheral organs. *British Journal of Animal Behaviour* 2:89–94. [aAHW]

Von Holst, E. & Mittelstaedt, H. (1950) Das Reafferenzprinzip (Wechselwirkungen zwischen Zentralnervensystem und Peripherie). *Naturwissenschaften* 37:464–76. [aAHW, TP]

Von Kries, J. (1910/1962) Notes on the perception of depth. In: H. von Helmholtz, *Treatise on physiological optics*, vol. 3, ed. and trans., J. P. C. Southall. Dover. (Original work published in German, 1867; in English, 1925.) [GJA]

Wade, N. J. & Swanston, M. T. (1987) The representation of non-uniform motion: Induced movement. *Perception* 16:143–60. [MS]

Waespe, W., Büttner, U. & Henn, V. (1981) Visual-vestibular interaction in the flocculus of the alert monkey I. Input activity. *Experimental Brain Research* 43:336–48. [aAHW]

Waespe, W. & Henn, V. (1979) The velocity response of vestibular nucleus neurons during vestibular, visual and combined angular acceleration. *Experimental Brain Research* 37:337–47. [aAHW]

(1981) Visual-vestibular interaction in the flocculus of the alert monkey II. Purkinje cell activity. *Experimental Brain Research* 43:349–60. [aAHW]

(1987) Gaze stabilization in the primate. The interaction of the vestibulo-ocular reflex, optokinetic nystagmus, and smooth pursuit. *Review of Physiology, Biochemistry and Pharmacology* 106:37–125. [UB]

Wallach, H. (1959) The perception of motion. *Scientific American* 201:56–60. [arAHW]

(1982) Eye movement and motion perception. In: *Tutorials on motion perception*, ed. A. H. Wertheim, W. A. Wagenaar & H. W. Leibowitz. Plenum Press. [aAHW]

(1985) Perceiving a stable environment. *Scientific American* 252(4):92–98. [aAHW]

(1987) Perceiving a stable environment when one moves. *Annual Review of Psychology* 38:1–27. [aAHW]

Wallach, H., Becklen, R. & Nitzberg, D. (1985) The perception of motion during collinear eye movements. *Perception & Psychophysics* 38:18–22. [aAHW]

Wallach, H. & Kravitz, J. H. (1965) The measurement of the constancy of visual direction and of its adaptation. *Psychonomic Science* 2:217–18. [aAHW]

Wallach, H. & Lewis, C. (1965) The effect of abnormal displacements of the retinal image during eye movements. *Perception & Psychophysics* 1:25–29. [AES]

Wallach, H. & O'Connell, D. N. (1953) The kinetic depth effect. *Journal of Experimental Psychology* 45:205–17. [aAHW]

Wallach, H., O'Leary, A. & McMahon, M. L. (1982) Three stimuli for visual motion perception compared. *Perception & Psychophysics* 32(1):1–6. [aAHW]

Walter, E. (1982) Identifiability of state space models. *Lecture Notes in Biomathematics* 46. Springer-Verlag. [RAMG]

Warren, R. (1990) Preliminary questions for the study of ego-motion. In: *Perception and control of self-motion*, ed. R. Warren & A. H. Wertheim. Erlbaum. [aAHW]

Warren, R. & Wertheim, A. H., eds. (1990) *Perception and control of self-motion*. Erlbaum. [TP]

Warren, W. H., Blackwell, A. W., Kurtz, K. J., Hatsopoulos, N. G. & Kalish, M. L. (1991) On the sufficiency of the velocity field for perception of heading. *Biological Cybernetics* 65:311–20. [N-GK]

Warren, W. H. & Hannon, D. J. (1988) Direction of self-motion is perceived from optical flow. *Nature* 336:162–63. [N-GK]

Warren, W. H., Jr. & Kurtz, K. J. (1992) The role of central and peripheral vision in perceiving the direction of self-motion. *Perception & Psychophysics* 51:443–54. [TAS]

Watanabe, E. (1984) Neuronal events correlated with long-term adaptation of the horizontal vestibulo-ocular reflex in the primate flocculus. *Brain Research* 297:169–74. [aAHW]

Welch, R. B. (1986) Adaptation of space perception. In: *Handbook of perception and human performance*. Vol. 1: *Sensory processes and perception*, ed. K. Boff, L. Kaufman & J. Thomas. Wiley. [RH]

Wertheim, A. H. (1981) On the relativity of perceived motion. *Acta Psychologica* 48 (special volume on the perception of motion): 97–110. [aAHW]

(1985) How extraretinal is extraretinal? *Perception* 14(1):A8. [aAHW]

(1987) Retinal and extraretinal information in movement perception: How to invert the Filehne illusion. *Perception* 16(3):277–414. [arAHW, HH]

(1990) Visual, vestibular, and oculomotor interactions in the perception of object motion during egomotion. In: *Perception and control of self-motion*, ed. R. Warren & A. H. Wertheim. Erlbaum. [TP]

(1992a) A psychophysical method to assess the gain of the otolith response. Paper presented at the satellite symposium of the 17th meeting of the Barany Society: "Vestibular-proprioceptive interaction for body orientation in space," June, Smolenice, Czechoslovakia. [aAHW]

(1992b) Motion perception during ego-motion: Measuring the otolith response. *Perception* 21(supp. 2):49–50. [aAHW]

(1993) Pilot studies on object motion perception during linear self-motion after long duration centrifugation of human subjects. *Institute for Perception Technical Report IZF-1993-B-3*. TNO Institute for Perception, Soesterberg, The Netherlands. [aAHW]

Wertheim, A. H. & Bekkering, H. (1991) The Filehne illusion is age dependent. *Perception* 20(1):85–86. [aAHW]

(1992) Motion thresholds of briefly visible stimuli increase asymmetrically with age. *Vision Research* 32(12):2379–84. [aAHW]

Wertheim, A. H. & Bles, W. (1984) A reevaluation of cancellation theory: Visual, vestibular and oculomotor contributions to perceived object motion. *Institute for Perception Technical Report IZF-1984-8*. TNO Institute for Perception, Soesterberg, The Netherlands. [aAHW, HH]

Wertheim, A. H., Hosman, R. J. A. W., de Graaf, B., Bles, W. & Krol, J. R. (1989) Visual motion perception during simulated space sickness on earth. *Proceedings of the 15th annual meeting of the European Undersea Biomedical Society*, Eilat, Israel. [aAHW]

Wertheim, A. H. & Mesland, B. (1993) Motion perception during linear ego-motion. *Institute for Perception Technical Report IZF-1993-3*. TNO Institute for Perception, Soesterberg, The Netherlands. [aAHW]

Wertheim, A. H. & Niessen, M. W. (1986) The perception of relative motion between objects during pursuit eye movements. *Perception* 15(1):49. [aAHW]

Wertheim, A. H. & Van Gelder, P. (1990) An acceleration illusion caused by underestimation of stimulus velocity during pursuit eye movements: The Aubert-Fleischl phenomenon revisited. *Perception* 19(4):471–82 (erratum in: *Perception* 19(5):700. [aAHW]

Wertheim, A. H., Van Gelder, P., Lautin, A., Peselow, E. & Cohen, N. (1985) High thresholds for motion perception in schizophrenia may indicate extraneous noise levels of central vestibular activity. *Biological Psychiatry* 20:1197–1210. [arAHW]

Whiteside, T. C. D., Graybiel, A. & Niven, J. I. (1965) Visual illusions of movement. *Brain* 88:13–210. [aAHW]

Wist, E. R., Diener, H. C. & Dichgans, J. (1976) Motion constancy dependent upon perceived distance and spatial frequency of the stimulus pattern. *Perception & Psychophysics* 19(6):485–91. [rAHW]

Wist, E. R., Diener, H. C., Dichgans, J. & Brandt. T. (1975) Perceived distance and the perceived speed of self-motion: Linear versus angular velocity? *Perception & Psychophysics* 17:545–54. [UB]

Wolpert, L. (1990) Field-of-view information for self-motion perception. In: *Perception and control of self-motion*, ed. R. Warren & A. H. Wertheim. Erlbaum. [aAHW]

Wong, S. C. P. & Frost, B. F. (1978) Subjective motion and acceleration induced by the movement of the observer's entire visual field. *Perception & Psychophysics* 24(2):115–20. [aAHW]

Wurtz, R. H. (1969) Comparison of effects of eye movements and stimulus movements on striate cortex neurons of the monkey. *Journal of Neurophysiology* 32:987–94. [PT]

Wurtz, R. H. & Duffy, C. J. (1992) Neuronal correlates of optic flow stimulation. *Annals of the New York Academy of Sciences* 656:205–19. [XMS]

Xerri, C., Barthélémy, F., Borel, L. & Lacour, M. (1988) Neuronal coding of linear motion in the vestibular nuclei of the alert cat. III: Dynamic characteristics of visual-otolith interactions. *Experimental Brain Research* 70:299–309. [aAHW]

Xerri, C., Barthélémy, F., Harlay, F., Borel, L. & Lacour, M. (1987) Neuronal coding of linear motion in the vestibular nuclei of the alert cat. I: Response characteristics to vertical otolith stimulation. *Experimental Brain Research* 65:569–81. [aAHW]

Yardley, L. (1992) Motion sickness and perception: A reappraisal of the sensory conflict approach. *British Journal of Psychology* 83:449–71. [LY]

Yoshida, K., Berthoz, A., Vidal, P. P. & McCrea, R. (1981) Eye movement-related activity of identified second order vestibular neurons in the cat. In: *Progress in oculomotor research*, ed. A. F. Fuchs & W. Becker. Elsevier, North-Holland. [aAHW]

Young, L. (1985) Adaptation to modified otolith input. In: *Adaptive mechanisms in gaze control*, ed. A. Berthoz & G. Melvill-Jones. Elsevier. [AAS]

Young, L. R. & Shelhamer, M. (1990) Weightlessness enhances the relative contribution of visually-induced self-motion. In: *Perception and control of self-motion*, ed. R. Warren & A. H. Wertheim. Erlbaum. [HER]

Zacharias, G. L. & Young, L. R. (1981) Influence of combined visual and vestibular cues on human perception and control of horizontal rotation. *Experimental Brain Research* 41:159–71. [WB]

Zeppenfeldt, P. (1991) *Bepaling van de drempelwaarden voor het waarnemen van verschillen tussen visuele en vestibulaire stimulatie tijdens eigenbeweging* [Determination of thresholds for the detection of differences between visual and vestibular stimulation]. Thesis, Technical University of Delft, The Netherlands. [aAHW]

Zinchenko, V. P. & Vergiles, N. Yu. (1972) *Formation of visual images: Studies of stabilized images*. Consultants Bureau, Plenum Press. [VIB]