

Chapter 2

On the Perception of Position

A. H. C. Van der Heijden¹, Jochen Müsseler², & Bruce Bridgeman³

¹*University of Leiden, The Netherlands*

²*Max Planck Institute for Psychological Research, Munich, Germany*

³*University of California, Santa Cruz, USA*

The visual field – the experienced visual space with objects on positions – is, in one way or another, the 'product' of calculations in retina and brain. It is often assumed, however, that for position perception no real calculations are required. The alternative is that the topographic location of an object in the outer world is represented geometrically by the location of a set of neurons in a topographic map in the brain. This position-as-a-code-for-position assumption introduces severe theoretical problems. One problem, for instance, is that even the best topographic map in the visual cortex, V1, is not geometrically congruent with the topography in the visual field. Research with partial-report bar-probe tasks and with position judgment tasks strongly suggests that in the calculation of perceived position two processes are involved, a locally operating one and a globally operating one. A tentative functional model, indicating how these two processes co-operate, is briefly described and it is shown how the model deals with the theoretical problems. Some implications for the neuroanatomical implementation of the model and its consequences for the status of the philosophical psycho-neural identity theory are indicated.

1 Introduction

With Bridgeman (1992) and many others, we assume that there are at least two independent representations of visual space in humans and primates, a cognitive or 'what' one, subserving visual experience and a sensorimotor or 'how' one, controlling visually guided behavior. The present contribution is about the cognitive representation, that is, about the visual field, the experienced visual space with objects at positions. The exact problem we address is: How is it brought about that we see or experience objects in a scene at what appear to be their correct positions?

Our interest in issues related to the perception of position arose from our previous studies of selective attention in vision (see, *e.g.*, Van der Heijden, 1992, 1993, 1996; Müsseler, 1987; Müsseler & Aschersleben, 1998; Müsseler & Neumann, 1992). Those studies strongly suggested to us that visual selective attention is closely connected with visual perception of position. One of us even concluded that visual selective attention is position information used in a special way (see Van der Heijden,

1992). Unfortunately, or fortunately, these conclusions did not bring our curiosity to an end. A quick look at the literature revealed that the topic 'perception of visual position' is at least as problematic as a theoretical issue as the topic 'selective attention in vision' is (see, *e.g.*, Bridgeman, Van der Heijden, & Velichkovsky, 1994). So, in the recent past we presented the solution of a dark issue, selective attention, in terms of a still darker issue, position perception. Now the real problems can no longer be evaded.

2 Position as a Code for Position

The visual field, the experienced visual space with objects on positions, is in one or another way the 'product' of operations or calculations in retina and brain. It is often assumed, however, that for the visual perception of position no operations or calculations are required. That assumption is based on the knowledge that in the visual pathways there are a number of topographic maps; orderly topographic maps in the retina, the lateral geniculate, V1 and V2, and less orderly maps higher on in the visual system. The assumption states that the location of an object, X, in the outer stimulus field is represented geometrically by the location of a set of neurons firing in one, or possibly more, of these topographic maps (see Smythies, 1994). Thus, anatomical location in one or another map is taken to be the code for location of an object in the visual field; anatomical position is taken as the code for position.

That there is something weird with the view that anatomical position serves as the code for perceived position becomes apparent when we look at how other object attributes are coded in the brain. Assume, for example, that the visual system is presented with a patch of light that we perceive as bright and red and at our right. A neurosurgeon or a neurophysiologist looking for the representation of that patch in the brain is a priori convinced that he will never find something bright or something red there. What she/he will look for are stimulation-dependent patterns of activation in smaller or larger groups of neurons, that is, for the outcomes of the stimulus-induced calculations or operations. The assumption that anatomical location serves as the code for location entails that the brain deals with position in an essentially different way. No operations or transformations are assumed to be required because an inert anatomical position is assumed to suffice.

Problems with position as a code for position. The assumption that anatomical position serves as, and is sufficient as, the code for position has repeatedly led and still often leads to theoretical problems, most of which are not unambiguously solved even today. The assumption creates philosophical and psychological problems. To the philosophical problems, circling around the psycho-neural identity theory, we turn at the end of this contribution; with regard to this issue, philosophy without facts is like words without meanings.

The psychological problems can be classified into two groups, problems for visual perception that arise when the eyes are stationary and problems for visual perception that arise when the eyes move. Because most readers are familiar with these problems we introduce them briefly.

Among the problems that arise when the eyes are stationary¹ are:

- the inverted (and left-right reversed) image on the retina and upstream in the visual system while we nevertheless perceive a correctly oriented world (the inverted image problem);
- the different sizes of the retinal projections of an object with different object distances while we nevertheless see the object as of approximately the same size (the size constancy problem);
- the distorted, squeezed, stretched and split topographic map in V1 and the even worse topography upstream in the visual brain while we nevertheless see a topographically correct visual world (the topography problem);
- the representation of different object features in different anatomical maps while we nevertheless see integrated color-form objects on their appropriate places (the feature integration problem).

Among the problems for visual perception that arise when the eyes are moving are, for instance,

- the moving image in retina and brain when the eyes saccade (or move otherwise) while we nevertheless see a stable visual world (the stable perceptual world problem);
- the fixed afterimage in retina and brain when the eyes saccade that we nevertheless see as moving in visual space (the moving afterimage problem);
- the fixed image in retina and brain of an object followed with pursuit movements and the moving background image while we nevertheless see the object moving and the background as standing still (the moving object-stable background problem);
- the fixed series of afterimages in retina and brain when the eyes saccade over a rapidly blanking small light and the moving background image while we nevertheless see the afterimages move and the background standing still (the 'phantom array' problem of Hershberger & Jordan, 1996).

Moreover, there are a number of related problems that one can easily observe when sitting in a train and looking at different distances in the landscape (what is apparently moving and what is apparently standing still?). These psychological problems, for which several principled and ad hoc solutions can be found in the literature (see, *e.g.*, Bridgeman *et al.*, 1994), make, in our view, abundantly clear that a search for alternatives for 'position as a code for position' is in order.

3 The Codes for Position

3.1 Partial-Report Bar-Probe Tasks

Our search for the codes for position started with our selective attention research. An important paradigm in that attention research is the partial-report bar-probe task

¹ A first problem is that the eyes are never really stationary. Even during steady fixation small eye movements, micro-saccades, drifts and tremors, occur (*e.g.*, Bridgeman *et al.*, 1994).

introduced by Averbach and Coriell (1961). There are two variants of the task that produce results that differ in a highly interesting way (Figure 1). In the first variant a circular array of letters and in the second variant a linear array of letters is briefly presented. One of the letters is indicated by a bar marker or an arrow. The visual position indicator appears either before, simultaneous with, or after the array. The exposure conditions are such that observers cannot make useful eye movements. The observers are instructed to name the letter indicated by the bar or arrow.

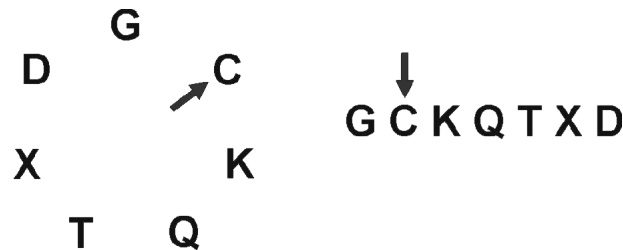


Figure 1: Examples of circular (left) and linear (right) arrays in partial-report bar-probe tasks.

In Figure 2 the main results obtained with the two variants are approximately sketched (the figure is based on Figure 1 in Eriksen & Collins, 1969, and Figure 2a in Hagenaar & Van der Heijden, 1997). The figure presents the proportions of letters correctly reported with pre-exposure, simultaneous exposure and post-exposure of the indicator for the circular variant (solid line) and the linear variant (broken line). The figure clearly shows the important difference in results obtained with the two variants. With the circular variant performance decreases monotonously as a function of moment of presentation of the indicator relative to the display (see Eriksen & Collins, 1969; Eriksen & Rohrbaugh, 1970). With the linear variant, however, performance first increases and then decreases, yielding an inverted U-function (see Hagenaar & Van der Heijden, 1997). So, the main difference in performance between the two variants of the task is found with pre-exposure of the indicator (there are other differences as well to which we turn further on).

A further analysis of the errors observed with the circular variant never revealed anything really spectacular. The bulk of the errors consisted of letter identification errors while virtually no errors in perceiving the position of the indicator (Eriksen & Collins, 1969) and a very modest increase in the number of these errors with delay of the indicator (Eriksen & Rohrbaugh, 1970) was observed. A further analysis of the errors observed with the linear variant revealed a number of spectacular results. Mewhort, Campbell, Marchetti and Campbell (1981) observed already that the bulk of the errors consisted of localization errors, not of identification errors, and that the bulk of these localization errors consisted of near localization errors, that is, consisted of report of the name of a letter that was in a position adjacent to the target letter. Further experiments and analyses by Hagenaar and Van der Heijden (1997) revealed that the majority of the near location errors consisted of central near location errors, that is, of the name of the letter adjacent to and at the foveal side of the target letter. Moreover, their analyses revealed that central near location errors occurred about equally frequently with pre-exposure as with post-exposure of the

visual indicator and appreciably less with simultaneous presentation of indicator and linear array.

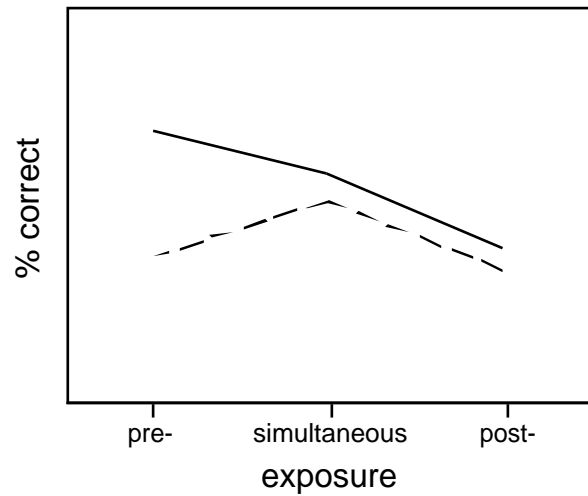


Figure 2: Percentage correct with pre-exposure, simultaneous exposure and post-exposure with circular (solid line) and linear (broken line) variant of the partial-report bar-probe task.

The total pattern of results just described is easily understood when we use one, seemingly ad hoc, assumption. That assumption is that a short-duration visual indicator, an arrow or a bar, when presented in relative isolation, is seen closer to the point of fixation than it really is or inspection at leisure would reveal, that is, that without visual 'anchors' a short duration arrow or bar appears as moved, shifted or jumped in the direction of the fixation point. With circular arrays and with an indicator moving in the direction of the fixation point nothing of relevance happens. With and without such a movement, with pre-exposure, simultaneous exposure and post-exposure, the indicator points in the same correct direction and neither an abundance of (central, near) location errors nor a dependence of proportion of location errors on moment of presentation of the indicator is to be expected. Only ease of letter identification determines the results depicted in Figure 2. With linear arrays, however, the situation is different. With pre- and post-exposure of the indicator, the indicator is presented in relative isolation, and, by ad hoc assumption, moves in the direction of the point of fixation. So, under these conditions, an abundance of central near location errors is to be expected.

With simultaneous presentation of array and indicator, the indicator is not presented in relative isolation but in the company of the array. By the ad hoc assumption, in this condition appreciably fewer central near location errors are expected. So, not ease of letter identification but localization problems with pre- and post-exposure of the indicator determine the results depicted in Figure 2.

3.2 Absolute Position Judgment Tasks

The assumption that a short duration visual indicator, an arrow or a bar, when presented in relative isolation, is seen closer to the point of fixation than it really is, is

not as ad hoc as it seems at first sight. In fact, a great number of investigators have reported position judgment studies that revealed exactly such an effect (see Van der Heijden, Van der Geest, De Leeuw, Krikke, & Müsseler, in press, for references). In none of these studies, however, were the exposure conditions exactly the same as those that are generally used in the partial-report bar-probe tasks. Therefore, we investigated the perception of position of a short duration small bar, presented on an otherwise empty screen, a situation mimicking the situation with pre- and post-exposure of the visual indicator in the partial-report bar-probe task (see Van der Heijden *et al.*, in press). In the first experiment, during practice, a fixation point appeared that was subsequently replaced by a horizontal array with seven equally-spaced positions, numbered 3, 2, and 1 on the left through 0 in the center to 1, 2, and 3 on the right (see Figure 3). Then a small, 30 ms, bar appeared below one of the digits and observers had to indicate its position by verbal labels 3 *on the left* through 0 to 3 *on the right*. In the main experiment first the fixation point appeared and then the bar in one of the seven positions on an otherwise empty screen. Observers had to indicate the position of the bar in the same way as during practice.

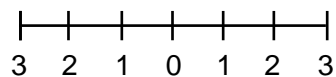


Figure 3: Horizontal array used during practice in absolute position judgment task.

The results of this experiment are presented in Figure 4. The figure shows that observers indeed underestimate the position of the bar or see it closer to the point of fixation than it actually is. A regression analysis showed that observers underestimate the eccentricity of the targets by about 15%. The experiment had a number of shortcomings, however. For instance only the coarse position responses '0', '1', '2' and '3' were allowed. Therefore, in subsequent experiments we tried to arrive at more precise estimates of the percentage underestimation of the position of the bar. In these experiments a moving cursor was used for indicating the perceived position. While these cursor experiments had their shortcomings – observers cannot really be prevented from following the cursor with eyes and head – we were nevertheless able to conclude that, under the exposure conditions used, observers underestimate the eccentricity of the bar by about 10% (see Van der Heijden *et al.*, in press).

The finding that observers underestimate the position of the bar by about 10% is of theoretical importance. The literature concerned with the precision of saccadic eye movements often reports a comparable 'undershoot' of about 10%.

"Most experiments with visually elicited saccades use the following paradigm: A subject fixates a stimulus, usually a small light dot in a dark field, that suddenly jumps to a new location. The subject is instructed to follow the target and does that by making one or more saccades. Under these conditions, and if the size and direction of the target jump is randomized, the saccade usually undershoots the target, and a correction saccade or smooth movement is necessary in the same direction. The amount of undershoot found in the literature varies considerably, but most authors agree now at an average value of about 10%." (De Bie, Van den Brink & Van Sonderen, 1987, p. 85; see, however, also Carpenter, 1988).

Neither for the saccadic undershoot nor for the perceptual 'undershoot' are presently good explanations available (but see Bridgeman & Stark, 1991, who found a gain of 0.87 for combined efference copy and oculomotor proprioception). The similarity in experimental conditions and the similar outcomes force us, however, to assume that the two phenomena are linked (see also Van der Heijden *et al.*, in press).

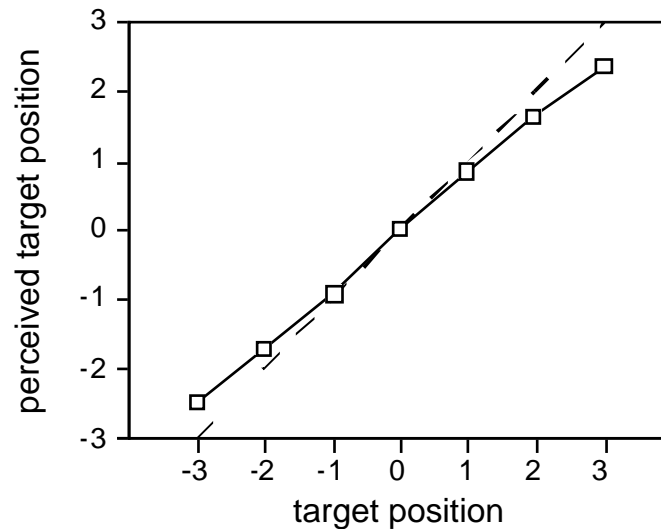


Figure 4: The perceived target positions. Veridical performance is indicated by the dashed line.

3.3 Relative Position Judgment Tasks

Before arriving at some conclusions about the codes for position in visual perception it is worthwhile to look at the results of a relative position judgment task that corroborates and summarizes what was said above. The task is a variant of the relative position judgment tasks used by Müsseler, Van der Heijden, Mahmud, Deubel, and Ertsey (in press). Figure 5 shows the two main conditions. In the unilateral condition (see upper panel) a row of 5 squares and 1 single square appear either to the left or to the right of a central fixation cross. In the bilateral condition (see lower panel) a row of 5 squares appears to one side of the fixation cross and a single square to the other side. In both main conditions the single square can precede the row of squares (pre-exposure, SOA -112 ms), appear simultaneously with the row (simultaneous exposure, SOA 0 ms) or follow it (post-exposure, SOA $+112$ ms). In both main conditions the single square is, averaged over trials, equally distant from the fixation cross as the square in the middle of the row. The position of the single square is, however, varied randomly over trials in small steps around its mean position. The task for the observer is to report the position of the small square relative to the position of the square in the middle of the row in terms of more peripheral (outer) or more central (inner).

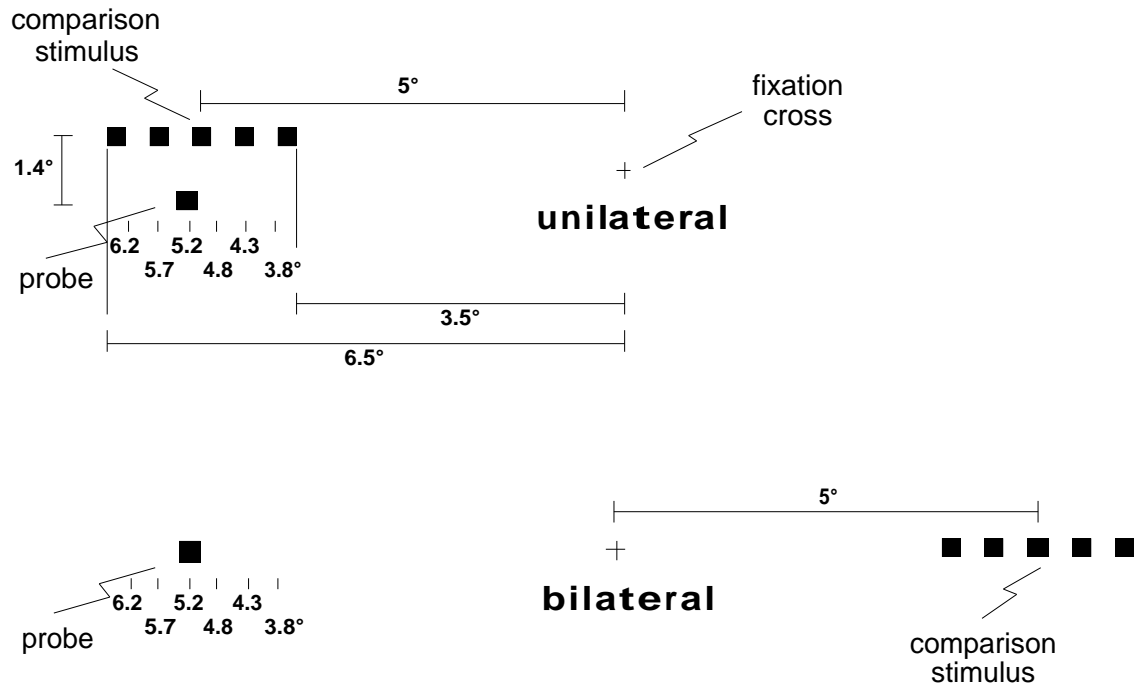


Figure 5: The relative position judgment task. Participants fixated a cross in the middle of a screen. In the unilateral presentation mode a stimulus configuration consisting of a single lower square (probe) and a spatially extended row of upper squares (comparison stimulus) appeared to the left or to the right of the fixation cross. Participants' task was to judge the position of the probe relative to the mid-position of the comparison stimulus. In the bilateral presentation mode participants' task was to judge which stimulus was more outer.

Figure 6 presents the results obtained. The complete interpretation of the figure is, in the present context, not of real importance (see for this explanation in terms of perceived position Müsseler *et al.*, in press). What is of importance is that with unbiased performance the number of 'outer' responses should approximately equal the number of 'inner' responses. This level of performance is indicated in the figure with the dashed line. The figure shows that this level of performance is found only with unilateral and simultaneous presentation of the square and the row of squares. Systematic deviations are found with unilateral presentation and with pre- and post-exposure of the single square (see Müsseler *et al.*, in press, for replications and extensions). Clearly, this pattern of results mimics the pattern of results obtained with the partial-report bar-probe task with linear arrays: good performance with simultaneous presentation of bar and array and systematic (localization) errors with pre- and post exposure of the indicator. With bilateral presentations the systematic deviations are found with all three SOA's. For this part of the relative position judgment task no equivalent bar-probe task exists. The data obtained in this condition therefore add to the existing knowledge of the perception of position.

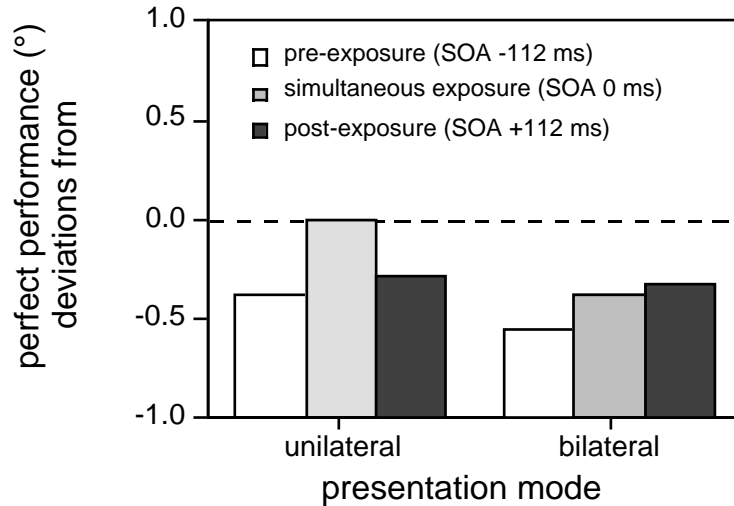


Figure 6: Deviations from perfect performance with pre-exposure, simultaneous exposure and post-exposure with unilateral and bilateral presentation mode in the relative position judgment task. Veridical performance is indicated by the dashed line.

4 Two Codes for Position

In the recent past a number of theorists have argued that in the perception of visual position two different 'sources,' mechanisms or codes are involved, a local 'visual' code and a global 'motor' code (see, *e.g.*, Koenderink, 1990, and Wolff, 1987, 1999). Wolff (1987, p. 259), for instance, states:

"Retinal stimulation merely provides the "substance" of which the spatial structure consists (see also the distinction between "space filling" and "space giving" sensations in Husserl's phenomenology, Scheerer, 1985)."

And Koenderink (1990, p. 126) distinguishes between a local 'simultaneous correlation structure or local receptive field structure' and a multi-local operation that 'requires a lot of additional structure.' Both theorists assume that (saccadic) eye movements are at the basis of the global or multi-local code (for earlier versions of this view see Von Helmholtz, 1866; Poincaré, 1902, 1905; Taylor, 1975).

Table 1: Good or optimal performance (+) and systematic deviations (-) for four different spatio-temporal exposure conditions.

	adjacent	separate
simultaneous	+	-
successive	-	-

In our view the data we presented are consistent with, and provide important further supportive evidence for, such a (double, dual) coding view. These data are summarized in Table 1. The table shows when good or optimal performance and when systematic deviations were found. Good or optimal performance was only found when a local operation sufficed, that is, when the assessment of spatial position could be based on two pieces of information that were close together in space and time, that is, adjacent and simultaneous (in the linear partial-report bar-probe task with simultaneous presentation of bar and array and in the unilateral relative position judgment task with simultaneous presentation of row of squares and single square). Systematic deviations, hinting at the involvement of the (saccadic) eye movement system, were found in all conditions in which a global operation was required because the relevant pieces of information, for the assessment of position, were separated in time or/and in space (in the linear partial-report bar-probe task with pre- and post-exposure of the bar, in the unilateral relative position judgment task with pre- and post-exposure of the single square and in all exposure conditions in the bilateral relative position judgment task) and in conditions in which only a single piece of information was presented in an otherwise empty field (in the absolute position judgment tasks). Under these conditions, only an (underregistered) extraretinal signal informs the brain about target eccentricity.

The results obtained in the bilateral relative judgment task with simultaneous (but spatially separated) presentation of the single square and the row of squares are of special importance. These data indicate that not only pre- or post-exposure, that is, temporal separation, invite the workings of the global operation but also distance in the visual field, that is, spatial separation. This outcome therefore strongly suggests that local and global operations work simultaneously in providing the spatial structure of the visual field.

5 The Perception of Position

5.1 The Static Eye

To show that the assumption that position is the code for position is not needed, Figures 7 and 8 present a tentative functional model that makes explicit what, in our view, can be at the basis of the visual perception of position (we neglect head and body movements and also exclude the third dimension [z-axis] and vergence movements in our analysis). The basic assumption of our model is that in the visual perception of spatial position, that is, in the construction of the visual field, two densely connected maps with different codes or representations are involved (further on we return to this assumption). In Figure 7 the situation in these maps is depicted for the situation when the eyes are straight ahead.

The map at the left, the visual sensory map (V), can be regarded as the 'substance' of which the spatial structure consists' (Wolff, 1987), as 'space filling' (Scheerer, 1985) and as a 'simultaneous correlation structure or local receptive field structure' (Koenderink, 1990). In the origin of this map the information coming from the middle of the fovea is represented. This map is not conceived as an orderly topographic

map. All kinds of deformations in terms of squeezing and stretching are allowed as long as neighborhood relations are preserved (we return to some properties of this map further on). The map can possibly be identified with V1 with all its deformations (*e.g.*, the cortical magnification factor).

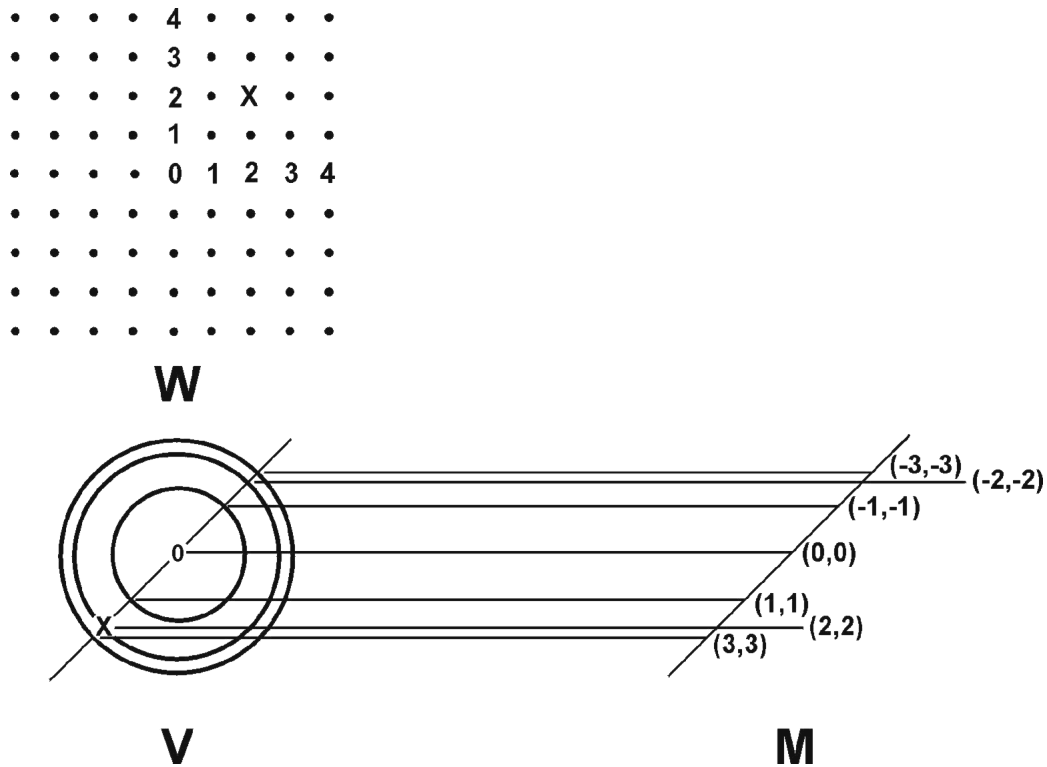


Figure 7: World (W), visual sensory map (V) and motor map (M). Eyes straight ahead and fixating 0 in W.

The map at the right, the motor map (M), has to be regarded as an eye position map, that is, a map that codes (eye) positions on (map) positions.² In the origin of this map the current eye position is coded ((0, 0) equals 'straight ahead'). For the moment, and for explanatory purposes only, this map can be regarded as an exact replica of map V (to this issue we return further on). Consequently in the other positions in this map the eye positions are coded that are required, or are to be realized, for bringing the spatially corresponding points in map V in the middle of the fovea (*e.g.*, for bringing the point marked X into the middle of the fovea the eye has to move to eye position (2, 2)). As stated, map M and map V are densely connected. Both maps determine what is seen. This can be taken to mean that the perceived positions result from map V 'enriched' by map M about the spatial positions in the visual field in terms of realized and required eye positions. Or, what is perceived results from map M 'enriched' by map V with identity information and local neighborhood relations.

² For evidence for the existence of this kind of map see, for example, Andersen, Snyder, Li, and Stricane (1993).

This conception of the visual perception of position evades the classical problems that result from the assumption that anatomical position serves as, and is sufficient as, the code for position. The inverted (and left-right reversed) image problem does not arise because the codes in map M tell, in terms of how the eye has to be positioned for foveating, what is up and down and left and right. Moreover, the local structure in map V may be of help in the perception of orientation (see Bridgeman *et al.*, 1994).

The size constancy problem is solved when the reasonable additional assumption is introduced that what the values of the codes in map M stand for is co-determined by cues for distance, vergence and accommodation, for instance. Moreover and again, the local structure in map V may be of help in the perception of size (see Gibson, 1979).

The topography problem, the problem of the distorted, squeezed, stretched and split topographic map in V1 and the even worse topographies upstream in the visual system, appears to be a pseudo problem in the present conceptualization. It is not the map positions in one or another weird topographic map, for example, in V1, that are the codes for perceived position but the eye position codes in map M.

The feature integration problem, the problem that is supposed to arise because different object features are represented in different anatomical maps, can be solved in exactly the same way as the position perception problem. Other maps, for example, a color map that codes colors in its map positions, when densely and appropriately connected with map V, can inform that map about other object features than position, for example, about color (see also Zeki, 1992, 1993). But, of course, 40 Hz will also always do.

5.2 The Moving Eye

Figure 8 shows how, in our functional model, the codes in map V and M (have to) change when a (saccadic) eye movement has been made. The figure shows the codes after an eye movement resulting in a new eye position (2, 2), that is, after a saccade of size and direction $(2, 2) - (0, 0) = (2, 2)$ (*i.e.*, intended position minus actual position) is made. Now in the origin of map V the visual information corresponding with eye position (2, 2), the X that is now in the middle of the fovea, is represented. The neighboring and further surrounding information is represented in neighboring and further surrounding regions in map V, subject to the fixed, inherent distortions in the map. In the origin of map M the current eye position (2, 2) is coded. All other eye position codes are incremented in the same way, that is, by (2, 2).

With regard to the problem of when and how the eye position codes in map M are changed there are two answers. One possibility is that after the eye movement has been executed the new eye position (x, y) is assessed and that this 'inflow' parameter is entered into the origin of map M and used to update the other eye position values in the map. This solution is close to Bridgeman *et al.*'s (1994) proposal: assessment of spatial position anew within each fixation. Another possibility is that before (or during or after) the eye movement, information about the size and direction of the intended movement, for example, $(x, y) - (a, b)$, with (a, b) the actual eye position and

(x, y) the intended eye position, is used to increment all eye position values in map M. This solution comes close to Von Helmholtz's (1866) and Von Holst and Mittelstaedt's (1950, 1971) 'neural command' or 'efference copy' solution. In our view it is reasonable, especially in view of the opportunities for perceptual learning (see Wolff, 1987), that both solutions, during the movement the 'neural command' and after the movement the inflow, are used.³

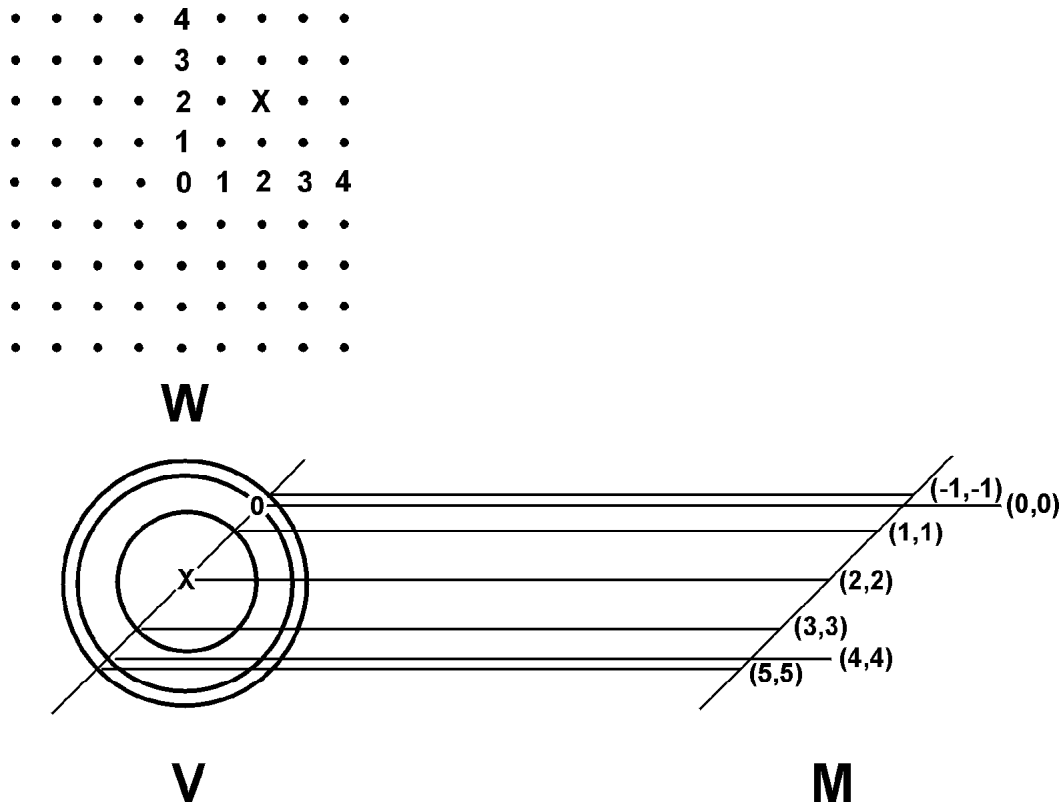


Figure 8: World (W), visual sensory map (V) and motor map (M). Eyes up and to the right fixating X in W.

This conception of the visual perception of position when the eyes move evades the classical problems that result from the assumption that anatomical position serves as, and is sufficient as, the code for position.⁴ Problems with the perception of a stable visual field with (saccadic) eye movements do not arise because the 'neural command' causes that during the (saccadic) eye movement the position codes in map M are adapted in correspondence with the displacement of the representation of the visual world in map V. After the movement 'inflow' is available to correct any errors in this adaptation process.

³ It is worthwhile to note that there are no space constant maps in the present proposal, that is, maps with neurons, with receptive fields, with constant locations in the world rather than constant locations on the retina (see Bridgeman, 1999, for evidence supporting this position).

⁴ In general, it is worthwhile to see that all 'moving eye' problems are 'secondary' problems. When correct position perception is accounted for, all 'moving eye' problems are accounted for. This goes not the other way around. When the stable visual world problem is solved (see, *e.g.*, Bridgeman *et al.*, 1994), the correct position perception problem needs not be also solved.

The moving afterimage problem, that is, the problem that an afterimage on a fixed retinal position is seen as moving with the eye during (saccadic) eye movements, is accounted for because, while the position of the afterimage in map V is indeed fixed, the corresponding eye position codes in map M change with each eye movement in proportion to the size and direction of the eye movement.

The moving object-stable background problem, that is, the problem that a moving object that is followed with pursuit movements (and that therefore is projected on a fixed position on the retina) is seen as moving and the problem that the background, whose projection is continuously moving on the retina, is seen as standing still, can be accounted for in a similar way. In map V the representation of the moving object is fixed and the representation of the background moves but in map M the position code corresponding to the object changes in exact correspondence with the position of the eyes while the position codes corresponding to the objects in the background remain unchanged.

The 'phantom array' problem and also the problems originating when looking out of a moving train can be accounted for in exactly the same way. We leave these problems for the interested reader.

6 Two Paradoxes

Niels Bohr once wrote that there is no hope for progress in science without a paradox. Fortunately, at this point in our theorizing we are confronted with two paradoxes, so, when we solve only one of these, there continues to be hope for progress in our science.

The first paradox is introduced with our assumption that in map M eye positions are coded and eye position codes are corrected using 'extra-retinal information.' Introspection, however, easily learns us that we are hardly aware, and if so then only very roughly and approximately, of the position of the eyes in their orbits. In his 'Remarks suggested by dr. Gordon Holmes's paper, "Disturbances of visual orientation"', Berry (1918, p. 604), a proponent of a retinal end-organ theory of space perception, writes

"When Dr. Holmes talks about impulses from the oculomotor muscles, as being provided by extra-retinal afferents, he obviously has in view something in connection with the muscles which he supposes exists apart from the retinal stimuli. But it has to be remembered that it is impossible with closed lids to tell for certain what is the position of the eyes. No doubt it is possible, without fixing any object, to rotate the eyes in the same direction, but apart from the consciousness of making this voluntary movement, there is no consciousness of the exact position of the eyes."

So, how can eye positions we are hardly aware of be coded in map M and be used for correcting codes in map M? The important issue is, of course, whether this argument based on 'direct' introspection is of any value. It is far from clear whether the information that the system uses for one or another purpose in creating the visual field, needs also to be available to us in introspection in 'raw' or 'unprocessed,' that is,

non-visual, form⁵. With auditory stimulation, for instance, there are tiny inter-aural arrival time differences that the system uses for working out the position of the sound source. These time differences, which are used for creating the auditory field, are not available to us in 'raw' form in introspection.

The second paradox that we can and will not solve is introduced with our, until now still implicit, assumption that the eye position codes in map M provide the metric and the topography in the visual field. The literature briefly referred to above tells us that in saccadic eye movement tasks and in position judgment tasks serious and systematic errors – 10% undershoot and 10 % underestimation – are made. So, why postulate a 'metric' and 'topography' for the visual field that seems to introduce serious and systematic distortions or deviations?

In the literature a great number of explanations and reasons for the undershoot with saccadic eye movements are listed. At present, however, even the exact conditions under which undershoot (and overshoot) will be found are far from clear (see, *e.g.*, De Bie *et al.*, 1987, and Carpenter, 1988). It cannot be excluded that, given some further conditions, undershoot will be mainly found with isolated targets in an otherwise empty field. In other words, it cannot be excluded that undershoot will be mainly found when nearly all cues for depth are eliminated. Further research with regard to saccadic undershoot and with regard to the relation between saccadic undershoot and underestimation in position judgment tasks is certainly in order⁶.

The paradox may be resolved with evidence that the combination of efference copy and proprioception also shows an undershoot in gain similar in magnitude to that required to account for the results reviewed above (Bridgeman & Stark, 1991). In structured fields the undershoot may not be important, as retinal information is available to calibrate visual direction.

7 Two Maps

At this point it is appropriate to move from the tentative functional model towards neuroanatomy to answer the question: What spatial properties are required of the possible neuroanatomical equivalents of map V and map M to perform the functions specified within the functional model?

We introduced the eye position map, M, as a map separate map from map V and stated that, for explanatory purposes only, map M could best be regarded as an exact replica of map V (see also Figures 6 and 7). Two remarks with regard to the properties of the neuroanatomical equivalent of map M are in order now. First, map M need not be anatomically separated from, and can as well be fully integrated with, map V for performing its required functions. What matters are only the appropriate dense connections between elements of map M and map V that deal with the same regions

⁵ It is worthwhile to note that it is generally assumed that the 'neural command' or the 'efference copy' is used for maintaining a stable visual field despite eye movements while it is far from clear whether, and if so to what extent, this information is available to introspection.

⁶ It is worthwhile to realize that an uncertain, peripheral, intended eye position attracts the fovea. Just because this intended position is uncertain, and because there is probably also some uncertainty about the actual eye position, undershoot and underestimation are possibly the result.

of the external world. Second, if separate, map M need not be an exact replica of map V, but can have any shape and internal order whatsoever. Map M codes positions on positions that are themselves meaningless (in the sense that the coordinates where the information is coded are arbitrary and have no meaning in themselves). Again, what matters is only the dense connectivity that allows elements in map M and in map V that deal with the same regions in the external world to communicate.

We introduced the 'space filling' map, V, as a spatial map in which all kinds of deformations in terms of squeezing and stretching were allowed as long as, in one way or another, neighborhood relations were preserved and we suggested that V1, with all its deformations, is possibly its neuroanatomical equivalent. It will by now be clear that, for performing its postulated function – creating a spatiotemporal extension – V1 needs neither to have a particular orientation nor to have a particular intrinsic metric, because the neuroanatomical equivalent of map M 'injects' orientation and distance. That V1 is cut into two halves seems to us to be of minor importance because the vertical meridian is represented in both hemispheres; seeing at exactly the same time at exactly the same place in exactly the same way cannot cause much confusion in visual perception.

8 To See

We already stated that we assume at least two independent representations of visual space in humans and primates, a cognitive or 'what' one, subserving visual experience, and a sensorimotor or 'how' one, controlling visually guided behavior (see also Bridgeman, 1992) and that our contribution is exclusively concerned with the cognitive representation. This forces us to consider the philosophical psychoneural identity theory, a theory that claims that a set of brain events, for example, the set of events in the neuroanatomical equivalents of map V and map M, is identical with a second set of subjectively experienced events, for example, the perception of a topographically ordered visual field, because it is the same set of events under a different mode of observation.

In 'Requiem for the identity theory,' Smythies (1994) argues with great insight and strength that the nearly generally held "Identity Theory is incompatible with the scientific evidence from an integrated approach to modern introspectionist psychology, clinical neurology, and neuroscience" (Smythies, 1994, p. 311). His "key point is that certain neural nets in the brain and in the visual field in consciousness contain the same information but coded in different ways – vectorially and topographically respectively" and that therefore the obstinate problem arises "how do you construct the purely topographic visual field ... out of this vectorially coded information ..." "Clearly a set of events ... can only be identical with a second set of events ... if the two codes are *congruent*." (Smythies, 1994, p. 326).

Smythies admits that Zeki (1992, 1993), knowing that V1 is the cortical map with the strongest topography, gets, to some extent, round this problem with his

"ingenious suggestion that the highest visual area of the cortex is actually V1, normally regarded as the lowest. This assumption is based on the observation that there are as many axons going 'down' from the visual centres in the temporal and

parietal lobes to V1 as there are going the other way. In Zeki's hypothesis information is fed by V1 to the specialized computational areas 'higher up' and then all this processed information is fed back to V1 and there put together to construct the visual field." (Smythies, 1994, p. 322).

In Smythies' view, however, even with this theory

"we are faced with the difficulty that the topographic code in V1 is geometrically and topologically incongruent with the topographic code in the visual field - as Lord Brain (1951) put this : 'When we perceive a two-dimensional circle we do so by means of an activity in the brain which is halved, reduplicated, transposed, inverted, distorted and three-dimensional. '" (Smythies, 1994, p. 326).

Without further provisions, all these observations are certainly correct. In our view, however, our introduction of a neuroanatomical equivalent of map M, that codes eye positions on otherwise meaningless positions and informs the neuroanatomical equivalent of map V, V1, about its results, extends and elaborates Zeki's (1992, 1993) proposal in the required way; a map that codes motor positions on meaningless anatomical positions can 'squeeze' and 'stretch,' that is, structure, the visual field in the desired way. With this proposal, however, the topographic coding – vectorial coding mystery is still not completely solved and further theoretical and empirical work with regard to this problem is needed.

9 Attention

As stated, our interest in questions related to the perception of position arose from our studies of selective attention and from our theoretical point of view that selective attention is position information, used in a special way. The highly problematic position-as-a-code-for-position view is incompatible with our observations with regard to properties and functions of selective attention. A bare, inert, anatomical position in one or another anatomic map can hardly play an active role in the selection of visual information. Actual and intended eye positions, represented or coded as patterns of activation in one or another map, however, can play an active role in the selection of visual information. Therefore, the position perception proposal here presented is, in principle, compatible with our views with regard to properties and functions of selective attention. Moreover, as a tentative model, the proposal provides an appropriate theoretical starting point for integrating issues with regard to the selection of positions for overt eye movements, the selection of positions for covert attending, and the relation between these two forms of selection.

References

- Andersen, R. A., Snyder, L. H., Li, C., & Stricanne, B. (1993). Coordinate transformations in the representation of spatial information. *Current Biology*, 3, 171-176.
- Averbach, E., & Coriell, A. S. (1961). Short-term memory in vision. *Bell System Technical Journal*, 40, 309-328.

- Berry, G. A. (1918). Remarks suggested by dr. Gordon Holmes's paper, "disturbances of visual orientation." *The British Journal of Ophthalmology*, 2, 597-607.
- Bridgeman, B. (1992). Conscious vs. unconscious processes. *Theory & Psychology*, 2, 73-88.
- Bridgeman, B. (1999). Strong and weak space constancy in receptive fields of striate cortex. *Psychological Research* (in press).
- Bridgeman, B. & Stark, L. (1991). Ocular proprioception and efference copy in registering visual direction. *Vision Research*, 31, 1903-1913.
- Bridgeman, B., Van der Heijden, A. H. C., & Velichkovsky, B. M. (1994). A theory of visual stability across saccadic eye movements. *Behavioral and Brain Sciences*, 17, 247-292.
- Carpenter, R. H. S. (1988). *Movements of the eyes* (2 ed). London: Pion.
- De Bie, J., Van den Brink, G., & Van Sonderen, J. F. (1987). The systematic undershoot of saccades: A localization or an oculomotor phenomenon? In J. K. O'Regan and A. Levy-Schoen (Eds.), *Eye movements: From physiology to cognition* (pp. 85-94). Amsterdam: Elsevier.
- Eriksen, C. W., & Collins, J. F. (1969). Temporal course of selective attention. *Journal of Experimental Psychology*, 80, 254-261.
- Eriksen, C. W., & Rohrbaugh, J. W. (1970). Some factors determining efficiency of selective attention. *American Journal of Psychology*, 83, 330-342.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Hagenaar, R., & Van der Heijden, A. H. C. (1997). Location errors in partial-report bar-probe experiments: In search of the origin of cue-alignment problems. *Memory & Cognition*, 25, 641-652.
- Hershberger, W. A., & Jordan, J. S. (1996). The phantom array. *Behavioral and Brain Sciences*, 19, 552-553.
- Koenderink, J. J. (1990). The brain a geometry engine. *Psychological Research*, 52, 122-127.
- Mewhort, D. J. K., Campbell, A. J., Marchetti, F. M., & Campbell, J. I. D. (1981). Identification, localization, and "iconic memory": An evaluation of the bar-probe task. *Memory & Cognition*, 9, 50-67.
- Müsseler, J. (1987). *Aufmerksamkeitsverlagerungen und Relativität. Ein experimenteller Beitrag zur Raum-Zeit-Wahrnehmung anhand eines Kontraktionsphänomens (Tandem-Effekt)* [Attention shifts and relativity. An experimental study to the perception of space and time with a contraction phenomenon (Tandem Effect)]. München: Minerva.
- Müsseler, J., & Aschersleben, G. (1998). Localising the first position of a moving stimulus: The Fröhlich Effect and an attention shifting explanation. *Perception & Psychophysics*, 60, 683-695.
- Müsseler, J., & Neumann, O. (1992). Apparent distance reduction with moving stimuli (Tandem Effect): Evidence for an attentional-shifting model. *Psychological Research*, 54, 246-266.
- Müsseler, J., Van der Heijden, A. H. C., Mahmud, S. H., Deubel, H., & Ertsey, S. (in press). Relative mislocalization of briefly presented stimuli in the retinal periphery. *Perception & Psychophysics*.
- Poincaré, H. (1902). *La science et l'hypothese*. Paris: Flammarion.
- Poincaré, H. (1905). *La valeur de la science*. Paris: Flammarion.
- Scheerer, E. (1985). The constitution of space perception: A phenomenological perspective. Paper to the symposium on "Sensorimotor interactions in space perception and action," February 18-22, 1985, Bielefeld.
- Smythies, J. R. (1994). Requiem for the identity theory. *Inquiry*, 37, 311-329.
- Taylor, J. G. (1975). *The behavioral basis of perception*. New Haven: Yale University Press. Reprint: Westport, CT: Greenwood Press.
- Van der Heijden, A. H. C. (1992). *Selective attention in vision*. London: Routledge.

- Van der Heijden, A. H. C. (1993). The role of position in object selection in vision. *Psychological Research*, 56, 44-58.
- Van der Heijden, A. H. C. (1996). Visual Attention. In O. Neumann & A. F. Sanders (Eds.), *Handbook of perception and action: Attention* (Vol. 3, pp. 5-42). London: Academic Press.
- Van der Heijden, A. H. C., Van der Geest, J. N., De Leeuw, F., Krikke, K., & Müsseler, J. (in press). Sources of position-perception error for small isolated targets. *Psychological Research*.
- Von Helmholtz, H. (1866). *Handbuch der physiologische Optik* (Vol. 3). Leipzig: Voss.
- Von Holst, E., & Mittelstaedt, H. (1950). Das Reafferenzprinzip (Wechselwirkungen zwischen Zentralnervensystem und Peripherie). *Naturwissenschaften*, 37, 464-476. English translation: (1980). The reafference principle. In C. R. Gallistel (Ed.), *The organization of action*. New York: Wiley.
- Von Holst, E., & Mittelstaedt, H. (1971). The principle of reafference: Interactions between the central nervous system and the peripheral organs. In P. C. Dodwell (Ed.), *Perceptual processing: Stimulus equivalence and pattern recognition*. New York: Appleton.
- Wolff, P. (1987). Perceptual learning by saccades: A cognitive approach. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action* (pp. 249-274). Hillsdale, NJ: Erlbaum.
- Wolff, P. (1999). Space perception and the intention of action. In G. Aschersleben, T. Bachmann, & J. Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events*. Amsterdam: Elsevier (this volume).
- Zeki, S. (1992). The visual image in mind and brain. *Scientific American*, 267, 68-77.
- Zeki, S. (1993). *A vision of the brain*. Oxford: Blackwell Scientific Publications.