

# Gain Modulation in the Central Nervous System: Where Behavior, Neurophysiology, and Computation Meet

EMILIO SALINAS and TERRENCE J. SEJNOWSKI

Gain modulation is a nonlinear way in which neurons combine information from two (or more) sources, which may be of sensory, motor, or cognitive origin. Gain modulation is revealed when one input, the modulatory one, affects the gain or the sensitivity of the neuron to the other input, without modifying its selectivity or receptive field properties. This type of modulatory interaction is important for two reasons. First, it is an extremely widespread integration mechanism; it is found in a plethora of cortical areas and in some subcortical structures as well, and as a consequence it seems to play an important role in a striking variety of functions, including eye and limb movements, navigation, spatial perception, attentional processing, and object recognition. Second, there is a theoretical foundation indicating that gain-modulated neurons may serve as a basis for a general class of computations, namely, coordinate transformations and the generation of invariant responses, which indeed may underlie all the brain functions just mentioned. This article describes the relationships between computational models, the physiological properties of a variety of gain-modulated neurons, and some of the behavioral consequences of damage to gain-modulated neural representations. *NEUROSCIENTIST* 7(5):430–440, 2001

**KEY WORDS** *Gain fields, Computational neuroscience, Computer model, Parietal cortex, Neglect, Coordinate transformations*

Neurons in the visual system represent the visual world; the deeper into the nervous system (in the sense of hierarchical processing), the more abstract these representations are. For instance, retinal ganglion cells and thalamic visual neurons respond to small, simple illumination patterns, such as a bright spot surrounded by a dark ring, presented at specific locations in the visual field (Shapley and Perry 1986). Neurons in the primary visual cortex, or V1, respond to somewhat larger and somewhat more complex patterns, such as a bright or dark bar of a certain orientation surrounded by contrasting edges (Hubel 1982). Neurons in area V4 still respond to large oriented bars to some degree, but they may be better activated by more complicated images such as spirals or star-like patterns (Gallant and others 1996). At the next stage, in the inferotemporal cortex (IT), neurons respond to even larger and much more complex images, such as faces, and often the responses are insensitive or invariant to the location, size, or orientation of the image (Desimone and others 1984; Desimone 1991; Gross 1992; Tovee and others 1994; Logothetis and others 1995).

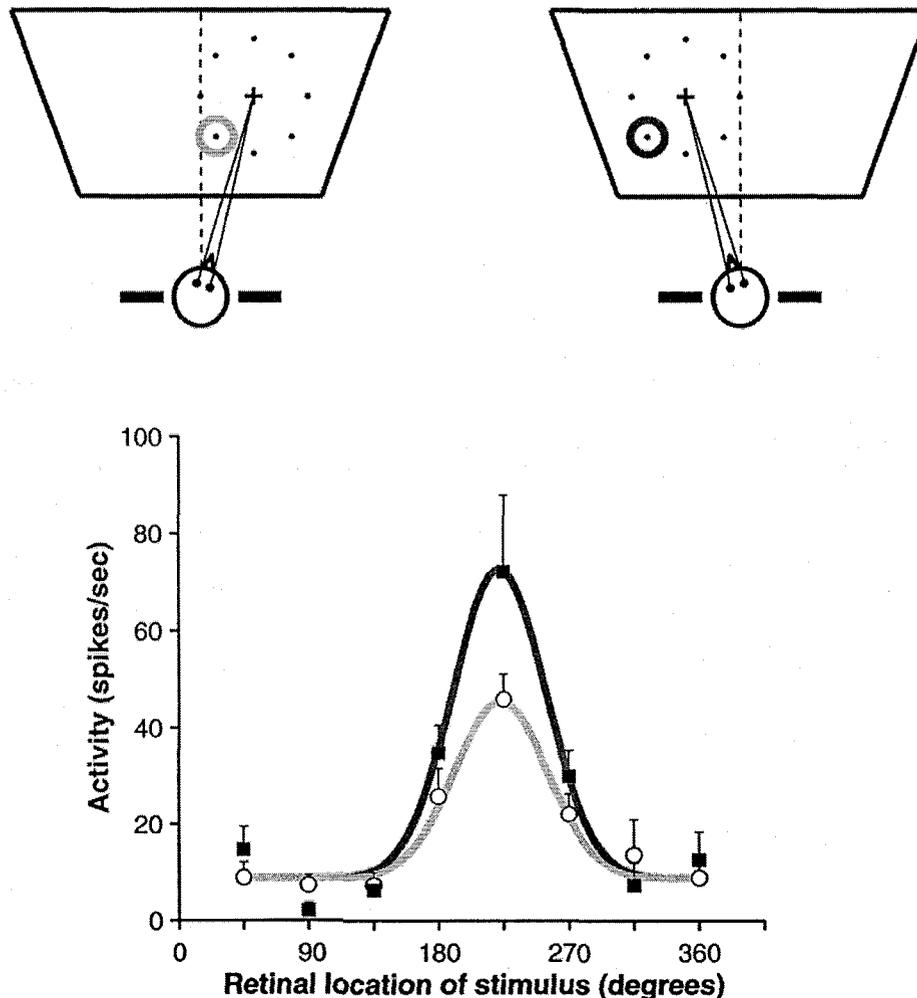
Research was supported by the Howard Hughes Medical Institute. We thank Paul Tiesinga for helpful comments. Computational Neurobiology Laboratory, Howard Hughes Medical Institute, The Salk Institute for Biological Studies, La Jolla, California (ES, TJS). Department of Biology, University of California at San Diego, La Jolla, California (TJS).

**Address correspondence to:** Emilio Salinas, Department of Neurobiology and Anatomy, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157 (e-mail: esalinas@wfubmc.edu).

As these examples show, a lot is known about how neurons encode the world; however, our understanding of how or why these higher-order representations are generated is rather limited. Many neurons in the visual pathway respond not only to visual stimuli but also to extraretinal signals, such as eye position. They do so by modifying the magnitude or “gain” of the visual response. Gain modulation is one of the few mechanisms for integrating information that is sufficiently widespread and sufficiently well understood from a theoretical perspective that may serve as a guiding principle for understanding how the nervous system builds such representations (Salinas and Thier 2000; Salinas and Abbott 2001).

## Gain Modulation by Eye and Head Position in the Parietal Cortex

The term *gain field* was coined to describe the interaction between visually evoked responses and eye position in the parietal cortex (Andersen and Mountcastle 1983). It parallels the concept of receptive field, which is central to the study of sensory systems. The receptive field of a visual neuron refers to a combination of two things: the location in the visual field from which the neuron may be activated and the kind of pattern that it responds to (its selectivity). To appreciate the coordinate transformation problems that the nervous system faces, it is important to keep in mind that the receptive fields of early visual neurons are defined in retinal or eye-centered coordinates, that is, the neural response depends on



**Fig. 1.** Visual responses that are gain-modulated by gaze angle. The response of a parietal neuron as a function of stimulus location was measured in two conditions, with the head turned to the right or to the left, as indicated in the upper diagrams. In these diagrams, the cross corresponds to the location where gaze was directed, called the fixation point; the eight dots indicate the locations where a visual stimulus was presented, one location at a time; and the colored circles show the position of the recorded neuron's receptive field. This was centered down and to the left of the fixation point. In the diagrams, the rightmost stimulus corresponds to 0 degrees, the topmost one to 90 degrees, and so forth. The dashed line indicates the direction straight ahead. The graph below plots the neural responses in the two conditions, indicated by the corresponding colors. The continuous lines are Gaussian fits to the data points. When the head is turned, the response function changes its amplitude, or gain, but not its preferred location or its shape. (Data redrawn from Brotschie and others 1995.)

where in the retina the illumination pattern falls. This is equivalent to saying that the neural response depends on the location of the stimulus relative to the fixation point (see Fig. 1). Early visual neurons, such as retinal ganglion cells, behave as if their receptive fields were attached to the eyes; when the eyes or head move, their receptive field locations stay in the same place relative to the fixation point and their selectivities are not altered. But this is not what happens in the parietal cortex.

Parietal neurons typically have very simple receptive fields—many cells essentially respond to spots of light—but their activity also depends on head and eye position. Figure 1 shows an example of the experiments performed initially by Andersen and Mountcastle (1983)

and later by Andersen and collaborators in parietal areas 7a and LIP (Andersen and others 1985, 1990; Andersen 1989; Brotschie and others 1995). The top diagrams show two conditions under which the receptive field of a parietal neuron was mapped: with the head turned to the right and with the head turned to the left. The response curves are shown below as functions of stimulus location relative to the fixation point. The curves peak at the same point, indicating that the receptive field did not shift when the head was turned. The shape of the response function did not change either, but its amplitude, or gain, did. This neuron has a visual receptive field that is gain-modulated by head position. In analogy with the receptive field, its gain field corresponds to the

full map showing how the response changes as a function of head position (not shown, but see Andersen and others 1985, 1990). Initially it was found that parietal neurons have gain fields that depend on eye position (Andersen and Mountcastle 1983; Andersen and others 1985, 1990), but an additional dependence on head position was described later (Brothie and others 1995), such that the combination of head and eye angles—the gaze angle—seems to be the relevant quantity.

Note that although gain-modulated parietal neurons carry information about stimulus location and gaze angle, the response of a single neuron, no matter how reliable, is not enough to determine these quantities; several neurons need to be combined. This is an example of a population code (Pouget and others 2000).

### Cardinal Issues

The interaction illustrated in Figure 1 is well described by a multiplication of two factors, one that depends exclusively on visual input, times another that depends on gaze angle only (see Box 1). The theoretical results mentioned below are based on this multiplicative description. Indeed, gain modulation is often very close to multiplicative (Brothie and others 1995; McAdams and Maunsell 1999; Treue and Martínez-Trujillo 1999), and there are some theoretical reasons why an exact multiplication may be advantageous; for instance, it may simplify the synaptic modification rules underlying the coupling between gain-modulated neurons and their postsynaptic targets (Salinas and Abbott 1995, 1997b). However, the essential feature of the gain interaction is that it is nonlinear (Pouget and Sejnowski 1997b; Salinas and Abbott 1997b); the influences from the two kinds of inputs are not simply added or subtracted. This is the key condition for most of the theoretical results reviewed below to be valid.

Two questions arise about this type of nonlinear interaction. First, what is its biophysical basis? In other words, what is it that allows a neuron to combine information from two sources such that its output behaves approximately as a product of two functions, each depending on a single source? This is not a trivial problem because, intuitively, the natural operation for a neuron to perform would be some kind of average or weighted sum, not a product. Four candidate mechanisms have been studied in some detail. First, if the dendrites of a neuron contain voltage-sensitive channels, neighboring synapses may interact nonlinearly, giving rise to an approximate multiplication at the output level (Mel 1993, 1999). Studies in invertebrates make this a viable possibility (Gabbiani and others 1999), but whether this mechanism actually applies to the cortex is uncertain. A second alternative is to have a network of recurrently connected neurons. According to theoretical studies (Salinas and Abbott 1996), if the recurrent connections are strong enough, they may give rise to gain fields that are very nearly multiplicative. This mechanism is based on ordinary properties of cortical circuits and is highly robust but requires network interactions. A third way in which the gain of a neuron may change is through

### Box 1: Computing with Gain Fields

Gain fields are a powerful computational tool; this can be seen as follows. First translate the result illustrated in Figure 1 into mathematical terms. Let  $x_{target}$  be the retinal location of the stimulus. The response of the neuron as a function of  $x_{target}$  is a curve with a single peak; let  $f(x_{target} - a)$  represent that function, where  $a$  is the location of the peak. Now let  $x_{gaze}$  represent the gaze angle. According to the experiments, changing the gaze angle only affects the amplitude of the response  $r$ . This may be described through a product,

$$r = f(x_{target} - a) g(x_{gaze}), \quad (1)$$

where the function  $g(x_{gaze})$  is precisely the gain field of the neuron. For a population of gain-modulated neurons, all responses may be described by the above equation, except that different neurons will have peaks at different locations (that is, different  $a$  terms) and somewhat different  $g$  functions as well. Now suppose that a downstream neuron is driven, through synaptic connections, by the population of gain-modulated neurons. A key mathematical result is that, under fairly mild assumptions about the functions involved and the values of the synaptic weights, the downstream response  $R$  will have the following form:

$$R = F(c_1 x_{target} + c_2 x_{gaze}), \quad (2)$$

where  $c_1$  and  $c_2$  are constants that depend on the synaptic weights and  $F$  is a peaked function that represents the receptive field of the downstream neuron (Salinas and Abbott 1995, 1997b). This result means that a set of neurons downstream may explicitly represent the quantity  $x_{target} + x_{gaze}$ , for instance, while another set of downstream neurons could represent  $x_{target} - x_{gaze}$ , both sets being driven by the same population of gain-modulated neurons. For the coordinate transformation illustrated in Figure 2,  $(x_{target} + x_{gaze})$  corresponds to a neuron that responds to stimulus location in head-centered coordinates (see Fig. 3). Thus, gain modulation at one stage can easily be converted into receptive field selectivity at the next stage. The same result is obtained when  $x_{target}$  and  $x_{gaze}$  are substituted by any other quantities, so the mechanism is very general.

changes in the synchrony of its inputs (Salinas and Sejnowski 2000). For this to work as a basis for gain modulation, a set of modulatory neurons should have the capacity of changing the synchrony of a separate population of neurons. It is not clear yet how this may happen, but the possibility is interesting considering that attention, which may give rise to changes in gain (Connor and others 1996, 1997; McAdams and Maunsell 1999; Treue

and Martínez-Trujillo 1999), also produces changes in synchrony (Steinmetz and others 2000; Fries and others 2001). Finally, more recently Chance and Abbott (2000) proposed that cortical neurons may receive two kinds of inputs, a mainly excitatory drive that generates the selectivity, and balanced inputs, which combine strong, simultaneous excitation and inhibition and may produce an approximately multiplicative modulation of the responses to the driving inputs. This mechanism is also based on well-documented properties of cortical neurons.

The problem of identifying the biophysical basis for gain modulation is an active topic of research, but we will not consider it any further. The rest of the article will gravitate around the second major question concerning gain fields: Why would this kind of multiplicative, distributed representation be useful?

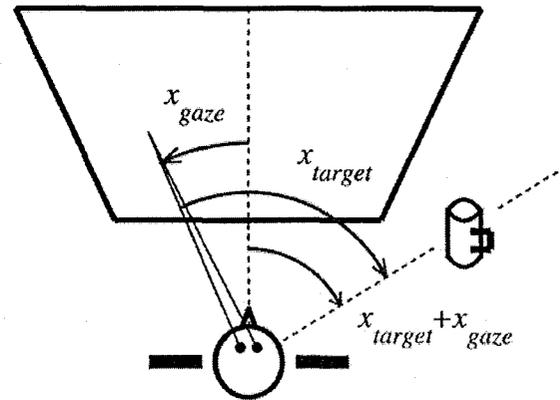
### Links between Gain Modulation and Coordinate Transformations

Figure 2 illustrates a simple coordinate transformation that the visual system solves routinely. To reach for an object, its location needs to be specified relative to the hand or to the body. However, this information must be extracted from the retinal image, which changes every time we move our eyes or our head. Therefore, information about eye and head position must be combined with visual information about the object to determine its location in a manner that is invariant with respect to eye or head position. The operation that so combines retinal and extra-retinal signals is a coordinate transformation. In the example shown in Figure 2, the computation required for this operation is the addition of two angles; how do neurons perform it?

Several lines of evidence indicate that gain-modulated neurons are ideally suited to add, subtract, and perform other computations that are fundamental for coordinate transformations.

### Results from Connectionist Networks

The computational power of gain modulation was first revealed by Zipser and Andersen (1988). They trained an artificial neural network to receive as inputs a gaze signal and the retinal location of a visual stimulus, and to generate an output equal to the stimulus location in body-centered coordinates. All input and output representations were modeled after experimentally measured neuronal responses, so the network essentially learned to perform the coordinate transformation illustrated in Figure 2. What the network adjusted during the training procedure (based on the back-propagation algorithm) were the properties of the "hidden units," which are a set of model neurons that stand between input and output stages. Once the network learned to perform the transformation accurately, the authors examined the responses of the hidden units and found that these had developed gain fields that depended on gaze, as in the parietal cortex. This suggested that gain modulation might be an efficient means to solve the coordinate transformation



**Fig. 2.** A coordinate transformation performed by the visual system. While reading a newspaper, you want to reach for the mug without shifting your gaze. The location of the mug relative to the body is given by the angle between the two dashed lines. For simplicity, assume that initially the hand is close to the body, at the origin of the coordinate system. The reaching movement should be generated in the direction of the mug regardless of where one is looking, that is, regardless of the gaze angle  $x_{gaze}$ . The location of the target in retinal coordinates (i.e., relative to the fixation point) is  $x_{target}$ , but this varies with gaze. However, the location relative to the body is given by  $x_{target} + x_{gaze}$ , which does not vary with gaze. Through this addition, a change from retinal, or eye-centered, to body-centered coordinates is performed.

problem given the neurophysiological properties of input and output representations.

Xing and Andersen recently extended this approach to other kinds of transformations, such as those involved in localizing auditory stimuli (2000b), and to the temporal domain (2000a). For the latter, they trained similar but somewhat more complex artificial neural networks to represent two consecutive saccades separated by a delay interval, thus forcing the network not only to perform coordinate transformations but also to acquire certain temporal dynamics. After training, the hidden units developed gain fields and also showed sensory and memory responses whose timecourses matched those of recorded LIP neurons. Interestingly, the connectivity patterns observed between these units were consistent with the model by Salinas and Abbott (1996) in which gain modulation is the result of nonlinear interactions between recurrently connected neurons. It is also worth noting that such model networks with feedback connections also naturally give rise to a form of short-term memory (Salinas and Abbott 1996), which is what the network studied by Xing and Andersen (2000a) had to develop to maintain the responses throughout a delay period. These studies suggest that gain modulation and working memory may be related through common biophysical mechanisms at the level of cortical microcircuits. Such persistent activity is consistent with the idea that parietal areas also serve to maintain internal representations of the locations of various body parts relative to the world and to each other (Wolpert and others 1998).

## Theoretical Results

The relation between gain fields and coordinate transformations was also investigated by Salinas and Abbott (1995), who studied the problem illustrated in Figure 2 using a different approach. They asked the following question. Consider a population of parietal neurons that respond to a visual stimulus at retinal location  $x_{\text{target}}$  and are gain-modulated by gaze angle  $x_{\text{gaze}}$ . Now suppose that the gain-modulated neurons drive, through synaptic connections, another population of downstream neurons involved in generating an arm movement toward the target. To actually reach the target with any possible combination of target location and gaze angle, these neurons must encode target location in body-centered coordinates, that is, they must respond as functions of  $x_{\text{target}} + x_{\text{gaze}}$ . Under what conditions will this happen? Salinas and Abbott (1995) developed a purely mathematical description of the problem (see Box 1), found those conditions, and confirmed the results through additional simulations. First, they found a mathematical constraint on the synaptic connections; satisfying this constraint guarantees that the downstream neurons will explicitly encode the sum  $x_{\text{target}} + x_{\text{gaze}}$ . This is illustrated in Figure 3. Here a model network of four gain-modulated neurons drives a downstream unit. The downstream response as a function of  $x_{\text{target}}$  shifts when  $x_{\text{gaze}}$  changes, exactly as expected from a neuron whose response is actually a function of  $x_{\text{target}} + x_{\text{gaze}}$ . This example is extremely simplified, but it clearly shows how transformed responses may arise from gain-modulated ones. Similar shifts, associated with a variety of coordinate transformations, have been documented in several cortical areas (Jay and Sparks 1984; Graziano and others 1994; Stricanne and others 1996). One of the best demonstrations comes from the work of Graziano and others (1997) in the premotor cortex, where some cells have visual receptive fields anchored to the head. Their responses are invariant to eye position and depend on the location of a stimulus relative to the head (see also Duhamel and others 1997).

Salinas and Abbott (1995) also found that the connections satisfying the critical constraint may be established by Hebbian or correlation-based learning, which is about the simplest, biologically plausible mechanism for synaptic modification (Brown and others 1990). For Hebbian learning to work, a learning period is required during which correct movements toward the target are executed while synaptic modification takes place. Watching self-generated movements, like babies do (Van der Meer and others 1995), would suffice for this; the hand itself would act as the target, and input and output representations would be automatically aligned. Finally, a third result was that the downstream neurons can just as well encode  $x_{\text{target}} - x_{\text{gaze}}$  or any other linear combination of the two variables, as long as the synaptic weights satisfy the corresponding constraint. Therefore, the population of gain-modulated neurons may serve as a platform for multiple frames of reference.

Pouget and Sejnowski (1997b) extended these results by showing that gain-modulated neurons such as those in

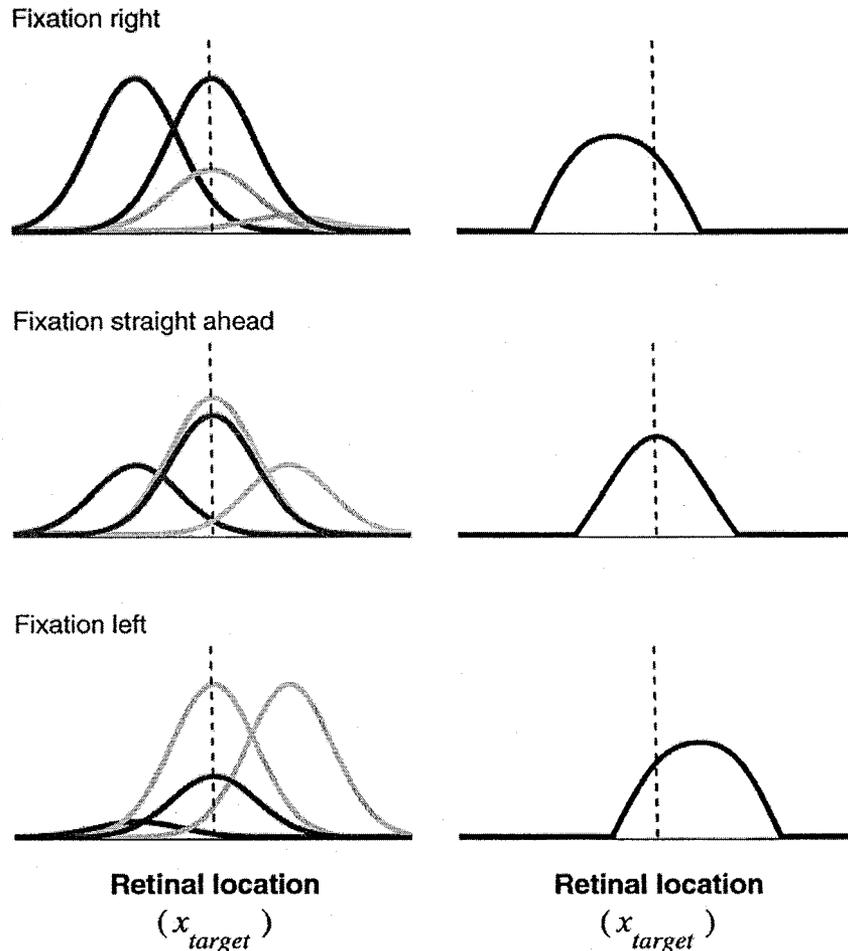
the parietal cortex are also ideally configured to generate nonlinear transformations. They showed that if there are enough different gain-modulated neurons—a “complete” set, in a mathematical sense—then any nonlinear function can be synthesized in a single set of synapses. That is, responses directly downstream from the gain-modulated neurons may be arbitrary functions of the encoded quantities. The mechanism by which the proper synaptic connections can be established in this general case is uncertain, but relatively simple rules may still be sufficient (Pouget and Sejnowski 1997b).

In the example discussed above, neurons responded as functions of  $x_{\text{target}}$ , and were gain-modulated by  $x_{\text{gaze}}$ , but it is crucial to realize that the results are valid for any other pair of encoded variables. For example, Pouget and Sejnowski (1994) showed that egocentric distance may be extracted from the activity of visual neurons tuned to binocular disparity and gain-modulated by the distance of fixation. Another interesting generalization (Salinas and Abbott 1997a, 1997b), discussed below, shows that responses to highly structured images can also be shifted to other coordinate frames. The mechanism by which a set of gain-modulated neurons gives rise to transformed representations downstream is extremely general.

## Neglect: A Coordinate Frame Syndrome

Neglect is a disorder observed after damage to the parietal lobes (Rafal 1998; Pouget and Driver 2000b). Patients with unilateral neglect typically seem unaware of visual stimuli appearing on the side contralateral to the lesion, even though they can see and identify objects on that side when presented in isolation. For instance, a patient with damage to the right parietal cortex may fail to eat food on the left side of a plate or to shave the left side of his face. Neglect patients typically fail at the line cancellation task illustrated in Figure 4a. The problem may extend to auditory and somatic stimuli as well.

One of the puzzling aspects of neglect is the coordinate frame that it affects (Karnath and others 1993; Driver and others 1994; Rafal 1998; Behrmann and Tipper 1999; Driver and Pouget 2000): Do patients ignore objects to the left of the fixation point, to the left of their bodies, or to the left of some other reference object? Figures 4b and 4c show stimuli used to investigate this problem (see caption). Researchers have struggled with this question, often finding contradictory results. The answer seems to require a close examination of the neurophysiological properties of parietal neurons (Pouget and Driver 2000a). The theoretical studies (Salinas and Abbott 1995; Pouget and Sejnowski 1997b) had shown that multiple coordinate frames can be extracted from the same set of gain-modulated neurons. Pouget and Sejnowski (1997b) further observed that damage to the parietal representation could therefore alter downstream neurons that encode object location in a variety of coordinate frames. Another key idea is that the reference frame used to locate an object may not be unique, varying instead with ongoing behavior. For instance, one may be interested in the location of an



**Fig. 3.** Combining the activity of several gain-modulated neurons may give rise to responses that shift. The left column shows the responses of four idealized, gain-modulated parietal neurons as functions of stimulus location; the three rows correspond to three eye positions. The blue and orange curves correspond to gain fields that increase when gazing to the right and to the left, respectively. The right column shows the responses of a downstream neuron driven by the four gain-modulated cells. These response curves were computed as weighted sums of the four parietal response curves on the left minus a constant, where negative values were set to zero. The response of the downstream neuron clearly shifts as the gaze angle changes; this is because it is a function of  $x_{target} + x_{gaze}$ . Including more parietal cells eliminates distortions in the shape of the response curves and allows other downstream neurons with peaks at different locations to shift, encoding target location explicitly in body-centered coordinates (Salinas and Abbott 1995).

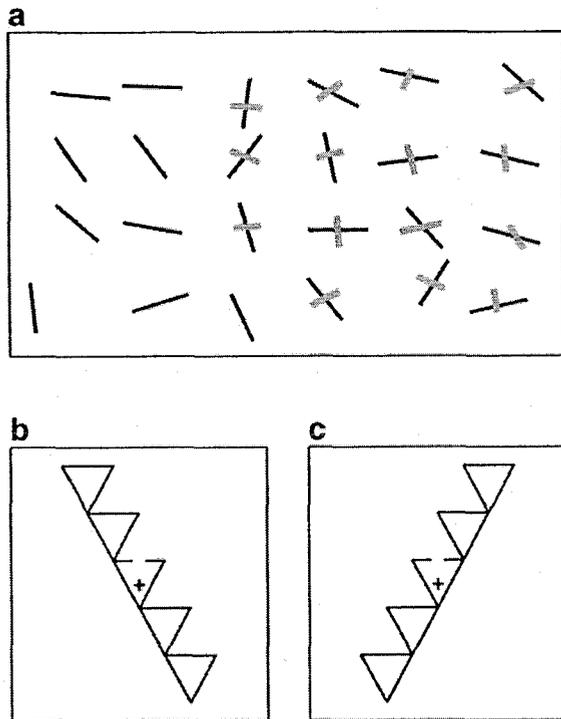
apple relative to the mouth, if one is about to eat it, or relative to the hand, if one wants to catch it, or relative to the ground, if one sees it falling from a tree. Pouget and Sejnowski (1997a, 2001) constructed a model of neglect based on a network where gain-modulated neurons give rise to representations of object location in a variety of coordinate frames. When some of the model neurons are eliminated in a way that mimics a typical parietal lesion, the model is able to reproduce a large number of effects found in neglect patients. In particular, it gives rise to deficits in multiple frames of reference. The model thus brings together neurophysiological, theoretical, and clinical findings into a unified picture.

One might still wonder, however, whether it is valid to extrapolate the findings from studies in monkeys to the parietal cortex of humans. But this seems to be appro-

priate: Studies using brain imaging techniques (Baker and others 1999; DeSouza and others 2000) have shown that parietal hand/arm movement regions in humans exhibit eye-position modulation that is consistent with its putative role in coordinate transformations and with the neurophysiological observations made in awake monkeys (Andersen and others 1985, 1990; Colby and Goldberg 1999).

#### *A Progression of Coordinate Transformations*

From the examples discussed above, one may draw the following picture. When an apple appears in the visual field, its location is initially specified relative to the fixation point, because the identity of the early visual cells that it activates corresponds to a specific position on the



**Fig. 4.** Neglect patients have difficulty locating objects in diverse frames of reference. *a*, In a line cancellation task, subjects with left neglect typically miss many of the targets on the left side of the display. *b* and *c*, Stimuli used by Driver and others (1994) to test for object-centered coordinates. The cross indicates the location where subjects had to fixate. The task consisted in detecting a gap like the one shown, except much smaller. In *b*, the triangles appear to point up and to the right, whereas in *c* they appear to point up and to the left. Left neglect patients performed significantly better in condition *c*, where the gap appears to the right of an axis parallel to the pointing direction, even though the central triangle containing the gap is identical in the two situations (Driver and others 1994). This indicates that the location of the gap is at least partly specified in object-centered coordinates.

retina. Suppose you want to bite the apple, for which you need to move your head. By combining the retinotopic responses with eye position information—through gain modulation (Andersen and others 1985, 1990)—neurons that respond to object location relative to the head (Graziano and others 1994, 1997; Graziano and Gross 1998) may be generated, and the correct direction of head movement can easily be read out from them.

Now suppose you want to reach for the apple. In this case, the early eye-centered visual information could be combined with information about the current location of the hand, also relative to the fixation point, to generate visual neurons that respond to object location relative to the hand (Graziano and others 1994, 1997; Graziano and Gross 1998; Buneo and others 2000). Again the key is that information about object and arm position should interact through gain modulation. Such gain-modulated responses are exactly what is found in the parietal reach region (Batista and others 1999; Buneo and others

2000), except that here the visual response to the object is only detected indirectly when there is an actual intention to move the arm toward it (Snyder and others 2000). But, remarkably, the coordinate transformation mechanism seems to work exactly as expected from the theory, with neurons in area 5, probably driven by the parietal reach region, responding as functions of the difference between hand and object locations (Buneo and others 2000). This difference is called motor error, because it directly encodes the direction in which the hand should move to in order to reach the apple.

Now take the example one step further. You see a tree with beautiful apples right next to a scarecrow; you go back to the car to fetch a bag, but now that you have moved, can you find the right tree in the middle of the orchard? Yes, because you know its position in the world, its position relative to other landmarks like the scarecrow. Some neurons in the hippocampus of monkeys are activated when gaze is directed toward a particular location in the world regardless of the position of the monkey (Rolls and O'Mara 1995), providing a neural correlate of this ability. According to the theory, these world-referenced responses could be synthesized from regular eye-centered visual responses that are gain-modulated by body position. Indeed, visual neurons with world-referenced gain fields are found in area 7a (Snyder and others 1998), which projects to the hippocampus.

A general principle seems to apply here. Information about object location originates from an eye-centered representation. At subsequent stages, new representations in other coordinate systems are created according to task requirements, and gain modulation always seems to mediate the generation of the new coordinate representations, whether they require simple or highly sophisticated transformations. The cascade of transformations may start already in V1, where gain-modulated cells have been found (Pouget and others 1993; Weyland and Malpeli 1993; Trotter and Celebrini 1999).

### Coordinate Transformations for Object Recognition

So far, we have discussed coordinate transformations that are useful for spatial perception and motor behavior, but representing information in different reference frames is also at the heart of other, apparently unrelated, cognitive functions such as object recognition, and in these cases gain modulation has also been implicated as the basis for key computational operations.

#### *Invariance to Object Location*

Object recognition is thought to depend to a significant degree on neurons in the IT (Goodale and Milner 1992; Gross 1992; Logothetis and others 1995). These neurons have two distinctive characteristics: First, they are strongly selective for highly complex images, such as faces (Desimone 1991; Gross 1992), and second, to a good extent their responses are insensitive to several image properties, in particular, to the exact location in

the visual field where images are presented (Desimone and others 1984; Tovee and others 1994). For instance, if a face subtending an 8-degree angle is shown, an IT cell may respond equally strongly to the face presented 15 degrees to the left or 15 degrees to the right of fixation, or at fixation, and may not respond at all to other kinds of images shown at these locations. This is known as translation invariance. These IT neurons provide a neural correlate of our capacity to recognize objects regardless of where they appear in the visual field, like when we are driving through a street and suddenly recognize, in the periphery, that a person is trying to cross. Translation-invariant responses must be the result of a tremendous amount of computation, because they have to be synthesized from the responses of early visual neurons that prefer simple illumination patterns and are not invariant at all; they have small, localized receptive fields (Shapley and Perry 1986).

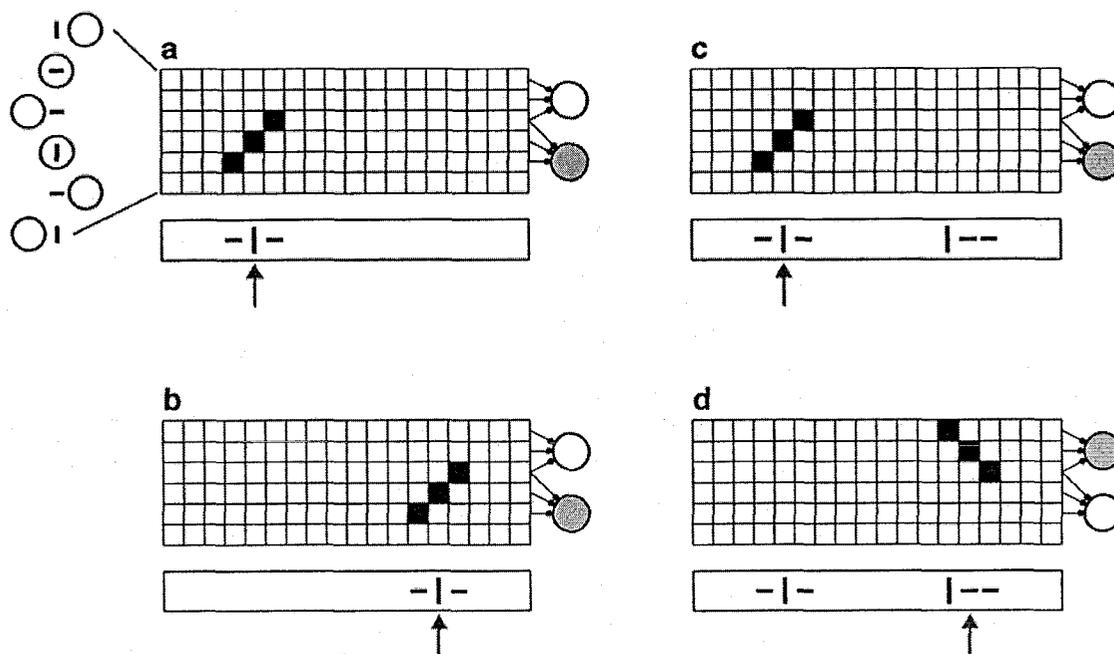
Translation-invariant responses may also be generated through gain modulation (Salinas and Abbott 1997a, 1997b), but in this case the modulatory quantity is attention. This is as follows. Attention is somewhat difficult to define precisely, but it essentially refers to our capacity to focus or concentrate on a specific part of the sensory world. For instance, stop for a second to think of how your shoe is applying pressure on your right toe, now. That must have shifted your attention to your right toe, decreasing the saliency of other sensations. Neuroscientists have developed ways to study visual attention (Parasuraman 1998). Two aspects of it are critical for this discussion: First, it can be directed to a particular location in a visual scene, and second, it can change independently of eye position. In fact, monkeys can be trained to direct their attention to specific locations while maintaining their gaze fixed. Using this kind of paradigm, Connor and others (1996, 1997) studied how attention affects neurons in area V4. These neurons have receptive fields in regular eye-centered coordinates and are reasonably well activated by oriented bars (Gallant and others 1996). In addition, Connor and collaborators found that V4 responses are gain-modulated by the location where attention is directed. Figure 5a shows some idealized examples based on the experimental results. A neuron in the top row, for instance, responds maximally when a vertical bar is shown inside its receptive field and attention is directed to the right; a neuron from the third row is maximally activated when a horizontal bar is presented in its receptive field and attention is focused to the left of it. In the experiments, neurons with many combinations of receptive field selectivities and preferred attentional locations were found. These interactions can be described through a multiplication, as in Equation (1), where the modulatory term now depends on attention. That is crucial because then it can be shown theoretically (Salinas and Abbott 1997a, 1997b) that downstream neurons driven by the gain-modulated ones can be selective for highly structured images and can respond in a coordinate frame centered on the location where attention is directed.

This mechanism is illustrated in Figure 5, which shows a highly simplified network that performs this transformation. Here the grid represents a set of visual neurons selective for horizontal or vertical bars that may appear at various horizontal positions. These neurons, like those in V4, are gain-modulated by the location where attention is focused, indicated by blue arrows. They drive two downstream neurons drawn as circles. Each downstream neuron is driven by three rows of gain-modulated units; the combination of rows determines the input pattern a downstream unit is most selective for. The downstream units mimic IT neurons, in the sense that they have two key properties: They are highly selective for specific patterns that are much larger than the receptive fields of upstream units, and they are insensitive to the absolute location where the patterns are presented. Instead, the responses illustrated depend on where a pattern appears relative to the point where attention is directed.

According to the full model (Salinas and Abbott 1997a, 1997b), translation-invariant responses in IT arise because they are driven by gain-modulated neurons in V4; note that IT is indeed the major target of V4. This explains translation-invariant object recognition as a change from eye-centered to attention-centered coordinates; it requires, however, that attention be focused on an object to be recognized (Olshausen and others 1993; Salinas and Abbott 2001). This might appear like a drawback, but it is actually consistent with psychophysical experiments (Mack and Rock 1998) showing that subjects have great difficulty identifying an image if attention is directed elsewhere.

### *Size Constancy*

The size of an image projected by an object on the retina depends not only on its real size but also on how far away the object is. Yet a person can compare a mouse lying on his or her hand with a cow standing on the other side of the road and tell which is larger, based only on retinal and eye position cues, even if the retinal image of the mouse is actually larger. This phenomenon is known as size constancy, and its neurophysiological basis has been investigated by Dobbins and others (1998). They recorded from neurons in V4 and tested their selectivity for object size by presenting visual stimuli at different distances, scaling them so that their retinal image size was constant. They found that V4 neurons do not prefer a specific object size—they do not exhibit size constancy—instead, they are selective for retinal images of given sizes and their responses are gain-modulated by viewing distance. It is not known whether neurons downstream from V4, in the IT, are selective for object size, but lesion experiments (Humphrey and Weiskrantz 1969; Ungerleider and others 1977) indicate that this area plays a key role in size discrimination. Therefore, IT neurons could become invariant to viewing distance by a process similar to those described before, namely, by pooling the responses of neurons with gain fields that



**Fig. 5.** A simple model showing how attention-centered coordinates can be useful for object recognition. For all panels, the bottom rectangle represents a one-dimensional visual field where patterns consisting of vertical and horizontal bars may appear. Squares in the above grid represent neurons that respond either to a horizontal or a vertical bar shown in their receptive field; these units are gain-modulated by attention. Blue arrows indicate the location where attention is directed. The circles on the right represent downstream neurons, each driven by three rows of gain-modulated units. Bars and circles to the left of *a* indicate, respectively, the selectivities and preferred attentional locations of the six rows of neurons. Thus, to obtain the maximal response from a neuron in the top row, a vertical bar must be shown in its receptive field and attention must be directed one receptive field to the right; to obtain the maximal response from a neuron in the second row, a horizontal bar must be shown in its receptive field and attention must be directed directly to it, and so on. Black squares show active neurons. *a* and *b*, The three-bar pattern shown activates the bottom downstream neuron (orange). The response is the same wherever the pattern is shown, as long as attention is directed to it. *c*, The bottom downstream neuron is still activated when a second pattern is presented simultaneously, if attention is focused on the preferred pattern. *d*, The bottom downstream neuron stops responding because attention is focused on the nonpreferred pattern, but another neuron selective for this pattern is now active. According to a more realistic version of this model (Salinas and Abbott 1997a, 1997b), neurons in V4 that are gain-modulated by attention (Connor and others 1996, 1997) may give rise to responses in the inferotemporal cortex that are highly selective and translation invariant (Desimone and others 1984; Tovee and others 1994).

are functions of viewing distance (Pouget and Sejnowski 1994).

### Gain Modulation and Motion Processing

When a person is strolling along the beach, his or her motion produces a pattern of optic flow that encodes the direction in which he or she is heading. However, turning the head and eyes to look at the approaching waves deforms the optic flow pattern. To maintain an accurate estimate of heading direction, the components of optic flow caused by eye and head movements need to be subtracted out. According to recent neurophysiological studies in awake monkeys, this is what neurons in area MSTd may be doing (Bradley and others 1996; Andersen and others 1997; Shenoy and others 1999). Some neurons in this area respond to optic flow stimuli and are gain-modulated by eye velocity, whereas other neurons encode heading direction in a manner that is independent of eye movements (Bradley and others 1996). This suggests that the responses of several gain-modulated neurons may be combined to produce

responses that are invariant to eye velocity, which is what the theoretical studies would suggest.

Shenoy and others (1999) recorded from monkeys trained to track a visual target in two conditions: either actively pursuing with their eyes or fixating on the target while their whole bodies were rotated, with the eyes and head fixed relative to the body. They showed that neurons in MSTd modify their tuning not only during pursuit, to compensate for the shift due to eye movements, but also during head rotations, which typically occur during gaze tracking. In the latter condition, it is information from the vestibular canals that gives rise to the compensation. Note, however, that not only vestibular information but also neck proprioception and efference copies of the motor commands influence the compensations observed behaviorally (Crowell and others 1998), so in fact several kinds of neural signals are integrated in MSTd. Thus, MSTd appears to be an area where eye-velocity and head-velocity signals originating from multiple sources are combined to eliminate self-motion and maintain a stable representation of external motion patterns (Andersen and others 1997). This may be impor-

tant for heading perception, pursuit or saccade commands, or even motion-based object recognition. Again there is strong evidence that gain modulation is the key mechanism at work.

### Final Remarks

Gain modulation is a widespread phenomenon in the cortex and in other subcortical structures (Van Opstal and others 1995; Stuphorn and others 2000) as well. According to a wide range of experimental results, gain-modulated responses are typically found upstream from areas that represent the same information in a different coordinate frame or in a way that is invariant, insensitive to some quantity. As a complement to this, theoretical and simulation work shows that transformed or invariant responses should be found downstream from neurons with gain fields, which provide a flexible and convenient basis for such transformations. This match between theory and experiment represents a great success for computational neuroscience: It shows how the measured neurophysiological properties of real neurons may underlie a certain class of nontrivial computations that are fundamental for a wide variety of behaviors. The neglect syndrome illustrates this most dramatically, spanning the whole range from single neuron neurophysiology to human behavior. Much work still lies ahead to refine our understanding of gain modulation: How does it arise biophysically? What are the actual learning rules involved in the transformations? What are its limitations? However, the way the functional significance of gain modulation has been established sets a standard for future investigations related to other aspects of brain function.

### References

- Andersen RA. 1989. Visual and eye movement functions of posterior parietal cortex. *Annu Rev Neurosci* 12:377-403.
- Andersen RA, Bracewell RM, Barash S, Gnadt JW, Fogassi L. 1990. Eye position effects on visual, memory, and saccade-related activity in areas LIP and 7a of macaque. *J Neurosci* 10:1176-98.
- Andersen RA, Essick GK, Siegel RM. 1985. Encoding of spatial location by posterior parietal neurons. *Science* 230:450-8.
- Andersen RA, Mountcastle VB. 1983. The influence of the angle of gaze upon the excitability of light-sensitive neurons of the posterior parietal cortex. *J Neurosci* 3:532-48.
- Andersen RA, Snyder LH, Bradley DC, Xing J. 1997. Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annu Rev Neurosci* 20:303-30.
- Baker JT, Donoghue JP, Sanes JN. 1999. Gaze direction modulates finger movement activation patterns in human cerebral cortex. *J Neurosci* 19:10044-52.
- Batista AP, Buneo CA, Snyder LH, Andersen RA. 1999. Reach plans in eye-centered coordinates. *Science* 285:257-60.
- Behrmann M, Tipper SP. 1999. Attention accesses multiple reference frames: evidence from visual neglect. *J Exp Psychol Hum Percept Perform* 25:83-101.
- Bradley DC, Maxwell M, Andersen RA, Banks MS, Shenoy KV. 1996. Mechanisms in heading perception in primate visual cortex. *Science* 273:1544-7.
- Brotchie PR, Andersen RA, Snyder LH, Goodman SJ. 1995. Head position signals used by parietal neurons to encode locations of visual stimuli. *Nature* 375:232-5.
- Brown TH, Kairiss EW, Keenan CL. 1990. Hebbian synapses: biophysical mechanisms and algorithms. *Annu Rev Neurosci* 13:465-511.
- Buneo CA, Jarvis MR, Batista AP, Mitra PP, Andersen RA. 2000. Coding of reach variables in two parietal areas. *Soc Neurosci Abstr* 26:181.
- Chance FS, Abbott LF. 2000. Multiplicative gain modulation through balanced synaptic input. *Soc Neurosci Abstr* 26:1064.
- Colby CL, Goldberg ME. 1999. Space and attention in parietal cortex. *Annu Rev Neurosci* 22:319-49.
- Connor CE, Gallant JL, Preddie DC, Van Essen DC. 1996. Responses in area V4 depend on the spatial relationship between stimulus and attention. *J Neurophysiol* 75:1306-8.
- Connor CE, Preddie DC, Gallant JL, Van Essen DC. 1997. Spatial attention effects in macaque area V4. *J Neurosci* 17:3201-14.
- Crowell JA, Banks MS, Shenoy KV, Andersen RA. 1998. Visual self-motion perception during head turns. *Nat Neurosci* 1:732-7.
- Desimone R. 1991. Face-selective cells in the temporal cortex of monkeys. *J Cogn Neurosci* 3:1-8.
- Desimone R, Albright TD, Gross CG, Bruce C. 1984. Stimulus-selective properties of inferior temporal neurons in the macaque. *J Neurosci* 4:2051-62.
- DeSouza JFX, Dukelow SP, Gati JS, Menon RS, Andersen RA, Vilis T. 2000. Eye position signal modulates a human parietal pointing region during memory-guided movements. *J Neurosci* 20:5835-40.
- Dobbins AC, Jee RM, Fiser J, Allman JM. 1998. Distance modulation of neural activity in the visual cortex. *Science* 281:552-5.
- Driver J, Baylis GC, Goodrich SJ, Rafal RD. 1994. Axis-based neglect of visual shapes. *Neuropsychologia* 32:1353-65.
- Driver J, Pouget A. 2000. Object-centered visual neglect, or relative egocentric neglect? *J Cogn Neurosci* 12:542-5.
- Duhamel JR, Bremmer F, BenHamed S, Graf W. 1997. Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature* 389:845-8.
- Fries P, Reynolds JH, Rorie AE, Desimone R. 2001. Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291:1560-3.
- Gabbiani F, Krapp HG, Laurent G. 1999. Computation of object approach by a wide-field, motion-sensitive neuron. *J Neurosci* 19:1122-41.
- Gallant JL, Connor CE, Rakshit S, Lewis JW, Van Essen DC. 1996. Neural responses to polar, hyperbolic, and Cartesian gratings in area V4 of the macaque monkey. *J Neurophysiol* 76:2718-39.
- Goodale MA, Milner AD. 1992. Separate visual pathways for perception and action. *Trends Neurosci* 15:20-5.
- Graziano MSA, Gross CG. 1998. Visual responses with and without fixation: neurons in premotor cortex encode spatial locations independently of eye position. *Exp Brain Res* 118:373-80.
- Graziano MSA, Hu TX, Gross CG. 1997. Visuospatial properties of ventral premotor cortex. *J Neurophysiol* 77:2268-92.
- Graziano MSA, Yap GS, Gross CG. 1994. Coding of visual space by premotor neurons. *Science* 266:1054-7.
- Gross CG. 1992. Representation of visual stimuli in inferior temporal cortex. *Philos Trans R Soc Lond B* 335:3-10.
- Hubel DH. 1982. Exploration of the primary visual cortex, 1955-1978. *Nature* 299:515-24.
- Humphrey NK, Weiskrantz L. 1969. Size constancy in monkeys with inferotemporal lesions. *Q J Exp Psychol* 21:225-38.
- Jay MF, Sparks DL. 1984. Auditory receptive fields in primate superior colliculus shift with changes in eye position. *Nature* 309:345-7.
- Karnath HO, Christ K, Hartje W. 1993. Decrease of contralateral neglect by neck muscle vibration and spatial orientation of trunk midline. *Brain* 116:383-96.
- Logothetis NK, Pauls J, Poggio T. 1995. Shape representation in the inferior temporal cortex of monkeys. *Curr Biol* 5:552-63.
- Mack A, Rock I. 1998. *Inattention blindness*. Cambridge (MA): MIT Press.
- McAdams CJ, Maunsell JHR. 1999. Effects of attention on orientation tuning functions of single neurons in macaque cortical area V4. *J Neurosci* 19:431-41.

- Mel B. 1999. Why have dendrites? A computational perspective. In: Stuart G, Spruston N, Häusser M, editors. *Dendrites*. Oxford (UK): Oxford University Press. p 271–89.
- Mel BW. 1993. Synaptic integration in an excitable dendritic tree. *J Neurophysiol* 70:1086–101.
- Olshausen BA, Anderson CH, Van Essen DC. 1993. A neurobiological model of visual attention and invariant pattern recognition based on dynamical routing of information. *J Neurosci* 13:4700–19.
- Parasuraman R, editor. 1998. *The attentive brain*. Cambridge (MA): MIT Press.
- Pouget A, Dayan P, Zemel R. 2000. Information processing with population codes. *Nature Rev Neurosci* 2:125–32.
- Pouget A, Driver J. 2000a. Relating unilateral neglect to the neural coding of space. *Curr Opin Neurobiol* 10:242–9.
- Pouget A, Driver J. 2000b. Visual neglect. In: Wilson RA, Keil F, editors. *MIT encyclopedia of cognitive sciences*. Cambridge (MA): MIT Press. p 869–71.
- Pouget A, Fisher SA, Sejnowski TJ. 1993. Egocentric spatial representation in early vision. *J Cogn Neurosci* 5:151–61.
- Pouget A, Sejnowski TJ. 1994. A neural model of the cortical representation of egocentric distance. *Cereb Cortex* 4:314–29.
- Pouget A, Sejnowski TJ. 1997a. A new view of hemineglect based on the response properties of parietal neurones. *Philos Trans R Soc Lond B* 352:1449–59.
- Pouget A, Sejnowski TJ. 1997b. Spatial transformations in the parietal cortex using basis functions. *J Cogn Neurosci* 9:222–37.
- Pouget A, Sejnowski TJ. 2001. Simulating a lesion in a basis function model of spatial representations: comparison with hemineglect. *Psychol Rev*. Forthcoming.
- Rafal RD. 1998. Neglect. In: Parasuraman R, editor. *The attentive brain*. Cambridge (MA): MIT Press. p 489–525.
- Rolls ET, O'Mara SM. 1995. View-responsive neurons in the primate hippocampal complex. *Hippocampus* 5:409–24.
- Salinas E, Abbott LF. 1995. Transfer of coded information from sensory to motor networks. *J Neurosci* 15:6461–74.
- Salinas E, Abbott LF. 1996. A model of multiplicative neural responses in parietal cortex. *Proc Natl Acad Sci U S A* 93:11956–61.
- Salinas E, Abbott LF. 1997a. Attentional gain modulation as a basis for translation invariance. In: Bower J, editor. *Computational neuroscience: trends in research*. New York: Plenum. p 807–12.
- Salinas E, Abbott LF. 1997b. Invariant visual responses from attentional gain fields. *J Neurophysiol* 77:3267–72.
- Salinas E, Abbott LF. 2001. Coordinate transformations in the visual system: how to generate gain fields and what to compute with them. In: Nicolelis M, editor. *Principles of neural ensemble and distributed coding in the central nervous system. Progress in brain research*, vol. 130. Amsterdam: Elsevier. p 175–90.
- Salinas E, Sejnowski TJ. 2000. Impact of correlated synaptic input on output firing rate and variability in simple neuronal models. *J Neurosci* 20:6193–209.
- Salinas E, Thier P. 2000. Gain modulation: a major computational principle of the central nervous system. *Neuron* 27:15–21.
- Shapley R, Perry VH. 1986. Cat and monkey retinal ganglion cells and their visual functional roles. *Trends Neurosci* 9:229–35.
- Shenoy KV, Bradley DC, Andersen RA. 1999. Influence of gaze rotation on the visual response of primate MSTd neurons. *J Neurophysiol* 81:2764–86.
- Snyder LH, Batista AP, Andersen RA. 2000. Intention-related activity in the posterior parietal cortex: a review. *Vision Res* 40:1433–41.
- Snyder LH, Grieve KL, Brotchie P, Andersen RA. 1998. Separate body- and world-referenced representations of visual space in parietal cortex. *Nature* 394:887–91.
- Steinmetz PN, Roy A, Fitzgerald PJ, Hsiao SS, Johnson KO, Niebur E. 2000. Attention modulates synchronized neuronal firing in primate somatosensory cortex. *Nature* 404:187–90.
- Striccanne B, Andersen RA, Mazzoni P. 1996. Eye-centered, head-centered, and intermediate coding of remembered sound locations in area LIP. *J Neurophysiol* 76:2071–6.
- Stuphorn V, Bauswein E, Hoffmann KP. 2000. Neurons in the primate superior colliculus coding for arm movements in gaze-related coordinates. *J Neurophysiol* 83:1283–99.
- Tovee MJ, Rolls ET, Azzopardi P. 1994. Translation invariance in the responses to faces of single neurons in the temporal visual cortical areas of the alert macaque. *J Neurophysiol* 72:1049–60.
- Treue S, Martínez-Trujillo JC. 1999. Feature-based attention influences motion processing gain in macaque visual cortex. *Nature* 399:575–9.
- Trotter Y, Celebrini S. 1999. Gaze direction controls response gain in primary visual-cortex neurons. *Nature* 398:239–42.
- Ungerleider LG, Ganz L, Pribram KH. 1977. Effects of pulvinar, pre-striate and inferotemporal lesions. *Exp Brain Res* 27:251–69.
- Van der Meer ALH, Van der Weel FR, Lee DN. 1995. The functional significance of arm movements in neonates. *Science* 267:693–5.
- Van Opstal AJ, Hepp K, Suzuki Y, Henn V. 1995. Influence of eye position on activity in monkey superior colliculus. *J Neurophysiol* 74:1593–610.
- Weyland TG, Malpeli JG. 1993. Responses of neurons in primary visual cortex are modulated by eye position. *J Neurophysiol* 69:2258–60.
- Wolpert DM, Goodbody SJ, Husain M. 1998. Maintaining internal representations: the role of the human superior parietal lobe. *Nat Neurosci* 1:529–33.
- Xing J, Andersen RA. 2000a. Memory activity of LIP neurons for sequential eye movements simulated with neural networks. *J Neurophysiol* 84:651–65.
- Xing J, Andersen RA. 2000b. Models of the posterior parietal cortex which perform multimodal integration and represent space in several coordinate frames. *J Cogn Neurosci* 12:601–14.
- Zipser D, Andersen RA. 1988. A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature* 331:679–84.