Neural models of stereoscopic vision

Randolph Blake and Hugh R. Wilson

Human stereopsis remains an enigma: how does the brain match features between the left and right eye images and compute disparity between these matched features? Developments in computational neuroscience and machine vision have led to several models of human stereopsis that provide insight into possible mechanisms underlying this phenomenon. These models, reviewed in this paper, adopt one of three general strategies. One class of models employs cooperative interactions, whereby a unique solution to the matching problem emerges from excitatory and inhibitory interactions among binocular neural elements. A second class of models implements matching and disparity computation serially over multiple spatial scales. A third class relies on local, non-interacting computations performed in parallel to overcome speed limitations inherent in the other models. Considered together, these theoretical developments offer fresh insights concerning the actual neural concomitants of binocular stereopsis.

Looking about the visual world, we see objects from two slightly different perspectives owing to the lateral separation of the eyes in the head. Evidently the brain is remarkably adept at combining these two monocular views, for binocular single vision and stereopsis occur effortlessly even when the two monocular views contain no discernible form information to guide the matching of the left and right eye images. Moreover, the brain is incredibly accurate in registering slight positional differences of image features from the left and right eyes, as evidenced by our keen sense of stereoscopic depth perception. How does the binocular visual system identify corresponding monocular features, measure positional disparities between these features and transform these measurements into a description of the three-dimensional layout of objects in the visual scene?

A first glimpse of the actual neural hardware underlying binocular vision was provided in the late 1960s when physiologists began recording from individual neurons in the visual cortex of mammals with frontally placed eyes. Shortly after Hubel and Wiesel described the existence of binocularly innervated cortical cells, several laboratories reported that many of these binocular cells in fact possessed receptive fields on non-corresponding areas of the two eyes. In other words, the optimum stimulus for activating these neurons would be an edge or surface marking located somewhere in depth other than the plane of fixation. By virtue of their disparity selectivity, these binocular neurons provided a plausible neural basis for stereopsis. This idea received further support from behavioral experiments demonstrating deficits in stereopsis in animals that lacked the normal complement of binocular cortical neurons.

Following these initial exciting discoveries, other laboratories concentrated on measuring the degree of disparity selectivity among single binocular neurons, correlating those measures with cortical cell types and searching for disparity-selective cells in extrastriate areas. Much of that work is reviewed elsewhere. Despite this enduring interest in disparity processing by the brain, however, real progress in our understanding of the neural bases of stereopsis has come slowly, partly because of the absence of a theoretical context in which to frame physiological questions. Recent developments in computational neuroscience and machine vision have provided new theoretical insights into the nature of stereoscopic vision, in turn leading to important psychophysical discoveries concerning human stereopsis. This article reviews those theoretical and empirical developments, in the hope of stimulating a refined assault on the actual neural concomitants of stereopsis.


Randolph Blake is at the Dept of Psychology, Vanderbilt University, Nashville, TN 37240, USA, and Hugh R. Wilson is at the Dept of Ophthalmology and Visual Science, University of Chicago, 939 E. 57th Street, Chicago, IL 60637, USA.
Neural models of stereopsis can be grouped conveniently into three major categories: cooperative models, models incorporating interactions across multiple spatial scales ('coarse-to-fine constraints' models), and feedforward models. Since many of these models have taken as their chief challenge a solution to the so-called 'false-target problem', we begin by defining that problem.

The false-target problem

As a first step in stereoscopic vision, the brain must determine which features contained in the left eye's image match up with features in the right eye's image. This determination poses a serious challenge when a complex scene is viewed binocularly. Consider, for instance, the formidable task posed by a random-dot stereogram (RDS) composed entirely of randomly placed black and white dots (Fig. 1): for any given horizontal line of the RDS, an individual black dot in the view of one eye could in principle be matched to any black dot in the view of the other eye. With 50 black dots per line, for example, the number of unique binocular pairings is 50! (i.e., $50 \times 49 \times 48 \ldots$), and all but 50 of this immense number are spurious! False targets arise from the inherent geometric ambiguity in binocular vision: a pair of two-dimensional images of a three-dimensional world contains insufficient information for the unique reconstruction of that world. How, then, can the false-target problem be surmounted? To eliminate false targets and accurately compute depth, additional constraints must be embodied within the neural substrate performing that computation. As summarized in the following sections, three general strategies have been proposed.

Cooperative models

Cooperative models solve the false-target problem and extend the disparity range processed using variants of two complementary themes: facilitation in the two-dimensional (i.e., fronto-parallel) plane and inhibition in the three-dimensional (i.e., depth) plane. Various cooperative models implement these interactions differently, and Fig. 2 summarizes schematically those various implementations. In all models, projections from the two retinas innervate an array of binocular cortical neurons sensitive to a range of crossed and uncrossed disparities. Without other connections, these binocular neurons would respond to correct binocular matches as well as to false targets. How do the models veto incorrect matches? According to Dev\textsuperscript{11} and to Nelson\textsuperscript{12}, each binocular cell inhibits all others tuned to different disparities at the same position in space (blue vertical line in left-hand portion of Fig. 2A), and excites binocular cells tuned to the same disparity at neighboring spatial locations (black horizontal line). Inhibition across disparities prevents the perception of transparent depth surfaces, while facilitation across space minimizes the disparity variation from point to point. Consequently, these two models will extract solid surfaces varying smoothly in depth. Although the Nelson model is purely qualitative, simulations show that the Dev model is capable of solving one-dimensional analogs of RDSs.

In a variant on this cooperative theme, Marr and Poggio\textsuperscript{13} proposed that the inhibition across disparities should run diagonally (blue diagonal lines in middle portion of Fig. 2A). This arrangement implements the constraint that a single monocular input can only generate one binocular response. In their model, incidentally, binocular units behave as logical AND gates, responding only to simultaneous stimulation of both eyes. This property further simplifies the false-target problem, although it renders the units in the model non-functional when one eye is closed. Simulations\textsuperscript{14} show that this model can solve RDSs. Sperling\textsuperscript{15} produced a qualitative neural model that differs from those described above in several interesting respects. While assuming the existence of inhibitory connections across disparities at each point in space, his model includes no direct excitatory connections among binocular units (which, incidentally, are AND gates). Instead, 'implicit positive feedback' occurs via disinhibition. For disinhibition to promote the spread of facilitation across space, however, the disparity-specific inhibition must spread spatially, as indicated by the blue regions in the right-hand portion of Fig. 2A. With this inhibitory spread, a binocular unit at one location will disinhibit not only itself (via inhibition of its potential inhibitors) but also its spatial neighbors tuned to the same disparity. Sperling's model also postulates two differently sized groups of monocular receptive fields, with binocular processing occurring in parallel in both groups of receptive fields, large and small. Rather than interacting, however, these two merely pool their depth estimates in an unspecified manner.

Mayhew and Frisby\textsuperscript{16} have developed a model that uses two different forms of cooperativity. An initial filtering stage extracts the positions and local orientations of edges in the monocular images. This filtering occurs in parallel on three spatial scales associated with three different ranges of spatial frequency tuning. Next, disparities are calculated between all horizontally disparate edge segments having similar orientations and matching polarities (light--dark or dark--light) in each monocular view. False targets are eliminated from this disparity field by application of two cooperative rules. One is a feature-specific continuity rule: cooperative interactions occur between oriented binocular edges that are roughly aligned along the direction of their orientation and at
the same disparity. This orientation-specific binocular facilitation is unique to this particular model. The other rule specifies cooperativity across spatial scales: binocular matches occurring at approximately the same orientation and position across all three scales are favored. This facilitation across scales works because an object contour will usually generate neural activity at similar orientations on all spatial scales. (This interaction across spatial scales is not unidirectional, in contrast to the coarse-to-fine strategies discussed below.) A refined version of this model employs a cooperative rule whereby each unit is facilitated by the sum of all neighboring units falling within a restricted cone-shaped volume of disparities across space (gray regions in Fig. 2B). Inhibitory connections spread within complementary cones in the disparity domain (blue regions in Fig. 2B). This pattern of connections thus extends cooperativity to surfaces curved smoothly in depth.

Grossberg and Marshall have also produced a cooperative model for what they call 'stereo boundary fusion'. Outputs from a monocular stage are summed at a binocular stage, within which recurrent excitation and inhibition sharpen the distribution of activity. These excitatory and inhibitory interactions are replicated independently over a range of size-tuned spatial scales. Although this model accounts for the transition from single vision to diplopia (i.e. double vision) with increasing disparity, it retains no information about disparity magnitude at a given spatial scale within the disparity range yielding single vision. Nor is it clear whether the binocular units in the model can discriminate crossed from uncrossed disparity (i.e. near from far depth) or whether they can segregate depth planes in a complex pattern (e.g. a RDS).

To summarize, cooperative stereo models succeed in solving the false-target problem through excitatory and inhibitory interactions. However, those same interactions preclude two readily perceived stereoscopic phenomena, i.e. depth transparency (multiple depth planes at a given visual direction) and depth averaging (forming a single depth plane from multiple proximate planes). These aspects of stereopsis are handled more naturally by models falling under the coarse-to-fine category.

Coarse-to-fine strategies

A key observation concerning the false-target problem, originally highlighted by Marr and Poggio, links disparity processing to receptive-field size (Fig. 3). Extensive physiological and psychophysical evidence implicates spatial filtering by cortical receptive fields that are responsive to a limited range of spatial frequencies, with the peak frequency (i.e. the value yielding maximum response) varying among cells. Now, suppose a pattern of spatial frequency wp is presented to the two eyes. Left and right eye images that differ sufficiently in pattern can be unambiguously matched so long as the disparity between the two images does not exceed 1/(2wp), i.e. 1/2 the period of the grating. However, cortical spatial filters respond to spatial frequencies up to about twice their peak spatial frequency wp. Therefore, to ensure that all matches are correct for any given cortical cell, the matching process must be restricted to disparities no greater than ±1/(4wp); i.e. ±1/2 the period of the cell's highest resolvable spatial frequency. (Because the

---

**Fig. 2. Illustration of excitatory and inhibitory connections postulated by different stereo models.** By convention, excitatory connections are black or gray and inhibitory connections are blue. The circles at the bottom of the drawings represent monocular neurons with horizontally displaced receptive fields. Circles within the ‘disparity’ matrix depict an array of binocular units with receptive fields centered at different horizontal locations, each tuned to a different crossed (near) or uncrossed (far) disparity. The thin dashed lines show the diagonal rows of binocular units (representing stimuli at different depths along the same sight line) innervated by an image falling two cells to the left of the fovea of the left eye (f_L) and two cells to the left of the fovea of the right eye (f_R). As each binocular unit in these models responds only when both monocular inputs are simultaneously active (AND units), only the cross-hatched unit in the matrix would respond to stimulation of these two monocular cells. (A) In the Dev and Nelson models (far left) each unit has reciprocal excitatory connections with its nearest spatial neighbors, which are tuned to the same disparity. Units at the same horizontal position tuned to different disparities are all mutually inhibitory, as illustrated by the vertical blue line. The Marr and Poggio model (middle of diagram) has excitatory connections identical to the Dev and Nelson models, but the inhibitory connections run diagonally through the network (blue lines). The Sperling model (far right) employs only inhibitory connections, but these must be assumed to have some spatial spread (blue regions) in order to permit a spatial spread of facilitation via disinhibition. (B) The model of Pollard et al. extends the region of facilitation to include a range of units falling within a disparity gradient limit of unity (dark gray regions). Inhibitory connections then extend into the cones of cells (blue) representing deep disparity gradients beyond unity.
range must encompass both crossed and uncrossed disparities, it is expressed as plus or minus the value within which unambiguous matches are possible. Restriction of disparity processing to this range will yield only correct matches: the false-target problem would be circumvented. We shall term this solution the ‘quarter-cycle limit’. The following paragraphs describe several ways this constraint could be implemented, starting with Marr and Poggio’s model.

Their model incorporated multiple spatial filters differing in size (or, in other words, peak spatial frequency). For a given sized spatial filter, disparity processing was restricted to a range defined by the quarter-cycle limit. Thus, small receptive fields (high spatial frequency) processed small disparities, while larger receptive fields (lower spatial frequency) processed progressively larger disparities. According to the model, low spatial frequencies in the image (registered by large receptive fields) controlled vergence eye movements, which, in turn, brought fine spatial detail (high spatial frequencies) within the quarter-cycle disparity limit of the smaller receptive fields. Disparities were then summed across spatial scales to produce a net disparity estimate. Computer simulation of this model accurately extracted depth from both RDSs and natural images, although some cooperative processing was found to be necessary in the implementation.

Despite its successful simulations, however, this model has three serious flaws. The first one is logical: any disparities greater than \( \pm 1/(4\omega_p) \) will be missed (where \( \omega_p \) refers to the peak frequency of the most coarse-scale cortical filters). Second, stereo targets composed entirely of high spatial-frequency information can trigger vergence eye movements, contrary to the model’s prediction. Third, stereo targets composed exclusively of high spatial frequencies can be combined binaurally at disparities well beyond the quarter-cycle limit (Fig. 4); this should be impossible according to the model. Hence, the Marr and Poggio assumption of a quarter-cycle limit is clearly violated at high spatial frequencies, and the assumption of vergence control by coarse spatial filters is also contradicted.

Other models involving coarse-to-fine strategies minimize the role of eye movements and instead implement spatial-scale interactions neurally. Models by Nishihara and Quam incorporate this interaction in the manner illustrated in Fig. 5. For any given region of the visual field, units with large receptive fields (coarse spatial scale) are selective for one of three values: crossed disparity for coarse-image (i.e., low spatial-frequency) features that are closer than the fixation plane, zero disparity for coarse features on or near the fixation plane, or uncrossed disparity for coarse features behind the fixation plane. By virtue of their large receptive fields, these coarse-scale units process large disparities, but only within the quarter-cycle limit, which avoids the false-target problem. For each given visual field location, the most active coarse unit among these three (crossed, zero or uncrossed) then constrains processing at the next
requires doubling their sampling density in space and coarse unit is most active, the 3x3 array of finer units on successively finer spatial scales, each having units disparity). This activated subset of finer scale units three-dimensional visual space as the most active smaller receptive fields analyse the same region of Simulations show that this class of coarse-to-fine and Halpern, D. L. J, which used D6 bars. The single finer-scale disparities to approximately the quarter-cycle limit for that spatial scale, again avoiding the false-negative columns in the visual cortex. Based on a signal-processing strategy termed 'cepstral filtering' used to detect radar echoes ('cepstral' being derived from 'spectral'), the algorithm implements an operation resembling autocorrelation. In the case of stereopsis, cepstral filtering is performed on the left and right eye neural images represented in a pair of neighboring ocular dominance columns. Any given pair of columns registers information over a small, local region of an image, and the filtering process is implemented on each pair with the entire ensemble processed in parallel. The filtering procedure itself involves three sequential steps: (1) computation of the power spectrum of the pair of matched images (a computation that could be approximated by spatial-frequency filtering); (2) a logarithmic transformation of this power spectrum (which could be accomplished with a compressive non-linearity); and

finest spatial scale to a 3x3 array of units whose smaller receptive fields analyse the same region of three-dimensional visual space as the most active coarse unit (shown in gray in Fig. 5). Thus, for instance, if the 'crossed' disparity (i.e., near depth) coarse unit is most active, the 3x3 array of finer units will be those preferring disparities clustered around this crossed disparity. These finer units have receptive fields of half the width of the coarse unit (which, to cover the same region of three-dimensional space, requires doubling their sampling density in space and disparity). This activated subset of finer scale units will now process disparities within the quarter-cycle limit for that spatial scale, again avoiding the false-target problem. This processing scheme is repeated on successively finer spatial scales, each having units packed twice as densely in both space and disparity. Simulations show that this class of coarse-to-fine constraint models can solve RDSs and natural images.

This coarse-to-fine constraint could be achieved in any of several ways. For instance, the most active coarse unit could facilitate activity in its associated array of fine-scale units. Alternatively, the most active coarse unit could exert spatially localized inhibition of all disparity-selective units tuned to the finer scale, except those centered on the region of three-dimensional space represented by that most active coarse unit (boxed subsets in Fig. 5). Less plausibly, monocular inputs to successively finer scales could literally be shifted, dependent on coarse-scale processing—a kind of neural 'shifter circuit'.

In fact, evidence exists for coarse-to-fine stereo constraints in human vision. As shown in Fig. 4, the fusion range for a band-limited bar asymptotes at about 15 min at spatial frequencies equal to or greater than two cycles per degree. These values were obtained for a bar presented against a uniform background. However, when a high spatial-frequency bar appears superimposed on a textured background, the spatial frequency of which is two octaves lower than that of the bar, the fusion limit for the bar drops to about 3.5 min. Moreover, this constricted disparity range for fusion remains centered on the disparity of the background texture: when the texture is imaged with crossed disparity, the 3.5 min disparity range of the bar is centered around that crossed disparity; but when the texture is imaged with uncrossed disparity, the bar's disparity range is shifted to that uncrossed value. The disparity of the coarse-scale image, in other words, constrains the range of disparities within which the fine-scale image can be fused binocularly. Now when the same bar is superimposed on a high spatial-frequency background texture, the bar's fusion limit is unaffected by the background—the fine-scale texture does not constrain binocular processing of the lower-scale image. Thus, coarse-scale disparities do reduce the fusion range for finer-scale disparities to approximately the quarter-cycle limit, consistent with the Nishihara and Quam models.

These models are not without their shortcomings, though. As currently formulated, the coarse-to-fine strategy depends on the presence of low-frequency information for processing of large disparities, yet we know that large disparities can be processed at the higher spatial scales alone (Fig. 4).

**Feedforward models**

The previous models all require either cooperative feedback or sequential disparity processing from coarse-to-fine spatial scale. Yeshurun and Schwartz have recently published a parallel, feedforward algorithm for disparity analysis, which has the advantage of speed. Their algorithm, while not touted as a complete model of human stereopsis, warrants mention for it was inspired by the existence of ocular dominance columns in the visual cortex. Based on a signal-processing strategy termed 'cepstral filtering' used to detect radar echoes ('cepstral' being derived from 'spectral'), the algorithm implements an operation resembling autocorrelation. In the case of stereopsis, cepstral filtering is performed on the left and right eye neural images represented in a pair of neighboring ocular dominance columns. Any given pair of columns registers information over a small, local region of an image, and the filtering process is implemented on each pair with the entire ensemble processed in parallel. The filtering procedure itself involves three sequential steps: (1) computation of the power spectrum of the pair of matched images (a computation that could be approximated by spatial-frequency filtering); (2) a logarithmic transformation of this power spectrum (which could be accomplished with a compressive non-linearity); and

---

*[Fig. 4. The ordinate plots fusion limit (maximum horizontal disparity yielding binocular single vision) as a function of peak spatial frequency. Test stimuli were vertical bars whose luminance profile was defined either by a difference of two Gaussians of opposite signs (DOG) or by the sixth derivative of a Gaussian (D6). Both types of bars have limited spatial frequency bandwidth: the bandwidth at half amplitude is 1.7 octaves for a DOG and 1.0 octaves for a D6. Solid and open squares are data obtained with DOGs by Schor et al. (IW, I. Wood; CS, C. M. Schor), while circles are averaged data for the three subjects in the study by Wilson et al. (WB & H, Wilson, H. R., Blake, R. and Halpern, D. L.), which used D6 bars. The single diagonal line plots the quarter-cycle limit, i.e., the upper value of a range of disparities within which the probability of false matches is essentially zero. Measured fusion limits fall well above this limit at spatial frequencies greater than about 2.0 cycles per degree. In this high spatial-frequency range, the visual system can fuse disparities many times larger than the quarter-cycle limit; the false-target problem arises and must be solved by neural interactions.]*
The disparity between the monocular images will now be represented by the location of a peak in the amplitude spectrum from the third step (which could be accomplished by disparity-selective binocular neurons\(^3,4\)).

As a model of human stereopsis, there are two problems with this algorithm. First, the bandpass receptive fields in the visual cortex are too broadly tuned to provide a reasonable approximation to the power spectrum\(^23,24\); the use of broadband filters to approximate the power spectrum would blur the disparity signal to a degree incompatible with measured human stereocuity. Moreover, the cepstral algorithm can extract only one disparity value per ocular dominance pair, which themselves sublend roughly ten min of arc at the fovea. Because of this limitation, the cepstral algorithm cannot correctly analyse stereo targets depicting transparent depth planes\(^26,29\), as these targets would create multiple disparity signals within a given local pair of ocular dominance columns. In fairness, however, we should reiterate that Yeshurun and Schwartz offered this algorithm for solution of stereo problems in machine vision, not biological vision.

Finally, Ohzawa et al.\(^33\) have developed a feedforward model of disparity extraction that employs a hierarchical arrangement of neural elements. The actual disparity-detecting elements are complex cells, each of which receives input from four binocularly innervated simple cells, all with the same preferred contour orientation. For a given simple cell, the left and right eye receptive fields have matching spatial profiles, meaning the cell has the same preferred stimulus (e.g. a light bar) for left and right eyes. These four simple cells are grouped into two pairs, with members of a pair having receptive fields that differ in spatial phase by 90°; the two cells comprising a pair are multiply inhibitory. Thus, for instance, a simple cell maximally responsive to a light bar would inhibit, and be inhibited by, a simple cell responsive to a dark bar; the two cells comprising the other pair would respond to edges (one cell being responsive to a light-to-dark edge and the other to a dark-to-light edge). The outputs of these four simple cells are squared (to compensate for the absence of spontaneous activity in simple cells) and then summed by the complex cell. With this arrangement, the complex cell receiving inputs from these four simple cells responds maximally when a bar or an edge of the
same contrast to the two eyes forms an image at a
given disparity anywhere within the complex cell's
receptive field.

The model adequately predicts the measured dis-
parity selectivity of a sample of phase-selective
complex cells in cat visual cortex. The model predicts,
however, that these disparity-specific complex cells
should respond vigorously to opposite contrast bars
presented to the two eyes at a disparity different from
the preferred disparity measured with matched con-
trasts. Such behavior would imply that these complex
cells cannot uniquely signal a given retinal disparity.
Nor can these complex cells signal disparities beyond
the quarter-cycle limit of the input simple cells. In its
present form, the model makes no provision for
interactions across spatial scale.

Concluding remarks

The models outlined here focus primarily on two
problems of stereopsis - solution of the false-target
problem and disparity computation. Cooperativity and
course-to-fine processing, neither of which has been
studied neurophysiologically, are two computational
strategies that have achieved some success in solving
these problems. It should be possible, however, to
discover whether either strategy is biologically imple-
mented. For example, coarse-to-fine strategies pre-
dict that the size of the preferred disparity of a cell
should be directly related to the size of its receptive
field. Moreover, the response of such a disparity-
selective neuron should be altered by the presence of
spatial frequencies below its passband. Cooperative
interactions across space might be revealed by changes
in the response to a given disparity as a function of
the disparity of neighboring stimulus elements.

There exist several problems as yet unsolved by
models of human stereopsis. These include the
derivation of accurate measures of three-dimensional
depth among objects in the visual scene from retinal
disparity information that is inherently ambiguous.
A given disparity may be associated with a range of
actual depth intervals between objects (see caption to
Fig. 1), and in order to scale disparity for perceived
depth, additional distance information is required.

How various sources of visual information are
combined to yield veridical depth perception remains
unsolved, although suggestions have been offered32,33.
In addition, most models of stereopsis fail
to account for stereopsis from disparities too large to
yield single vision, which is readily observed.

particularly, models in which binocular AND units register
disparity simply cannot handle stereopsis from diplo-
ic images. In fact, most binocular cortical neurons
respond to monocular stimulation in a fashion inco-
sistent with a logical AND operation2-4. Also largely
ignored by these models is binocular rivalry, the visual
alternation caused by dissimilar stimulation of corre-
ponding retinal areas. Because rivalry is a response of
the stereo system to aberrant stimulation, biologically
plausible models should generate rivalry under those
stimulus conditions36.

Finally, modellers and neurophysiologists have
worked under the assumption that stereoscopic com-
putations are based directly on positional disparities
among local stimulus features. However, it has been
claimed that pure temporal disparity information, in

the absence of positional disparity, creates a sensation
of stereopsis37. Moreover, some psychophysical
evidence38 suggests that disparity or depth curvature
(i.e. the second spatial derivative of disparity, in
contrast to the disparity gradient19) might provide the
basis for perception of three-dimensional surfaces.
Although there would certainly be no advantage for
visual processing to solve the false-target problem
first and then differentiate the result (retaining only
the second derivative), disparity curvature can be
directly computed by comparing contour curvatures in
the two monocular images. Disparity curvature and
temporal disparity processing certainly warrant further
exploration, both theoretically and physiologically.

So, computational models have not completely
solved the puzzle of human stereopsis. These models
have, however, sharpened our thinking about the
nature of the problem and about possible solutions.

Specifically, they have introduced the notions of
cooperativity, interactions across spatial scales and
the false-target problem. With these concepts,
neurophysiologists now have a framework to guide
their explorations of the hardware responsible for
stereopsis and binocular single vision.

Selected references

1 Julesz, B. (1971) Foundations of Cyclopean Perception,
University of Chicago Press
106–154
3 Barlow, H. B., Blakemore, C. and Pettigrew, J. D. (1967)
J. Physiol. 193, 327–342
4 Pettigrew, J. D., Nikara, T. and Bishop, P. O. (1968) Exp.
Brain Res. 6, 391–410
5 Blake, R. and Hirsch, H. V. B. (1975) Science 190,
1114–1116
6 Crawford, M. L. J., Smith, E. L., Harwerth, R. S. and
779–781
469–492
physiol. 49, 1148–1167
7, 379–412
223–239
16 Mayhew, J. E. W. and Frisby, J. P. (1980) Perception 9,
69–86
207, 187–217
Perception 14, 449–470
29–51
44, 421–432
204, 301–328
University Press
24 Wilson, H. R., Levi, D., Maffei, L., Vovamo, J. and Devalois,
Foundations (Spillman, L. and Werner, J. S., eds), pp.
231–272, Academic Press
Perception 10, 299–304
661–665

Acknowledgements
Supported by grants from NIH (EYO7760 and EY02158). This
review is dedicated to Befa Julesz, whose
development of the
random-dot stereogram
profoundly changed
the way visual
scientists think about
stereopsis. Andrew
Parker furnished
useful comments on
an earlier draft.
**Autoimmunity to glutamic acid decarboxylase (GAD) in Stiff-Man syndrome and insulin-dependent diabetes mellitus**

Michele Solimena and Pietro De Camilli

Stiff-Man syndrome (SMS) is a disorder of the CNS, characterized by rigidity of the body musculature, which has been hypothesized to result from an impairment of GABAergic neurotransmission. GABA is the main inhibitory neurotransmitter of the brain. It is also a putative signal molecule in the pancreas, where it is produced by β cells (insulin-secreting cells) - the autoimmune target in insulin-dependent diabetes mellitus (IDDM). Autoantibodies to the GABA-synthesizing enzyme glutamic acid decarboxylase (GAD) have been found in SMS and in IDDM. This review summarizes evidence suggesting that SMS may be an autoimmune disease and discusses the possible significance of the autoimmune response to GAD in SMS and IDDM.

Until recently, it was thought that anatomical barriers and functional properties normally protect neurons of the CNS from becoming the target of autoimmunity. Now, however, increasing evidence suggests that this is not the case. Paraneoplastic neurological disorders were the first neuronal diseases of the CNS for which an autoimmune origin was proposed. Evidence supporting an autoimmune pathogenesis has now been provided for a rare and severe neuronal CNS disease, Stiff-Man syndrome (SMS).

**Stiff-Man syndrome**

SMS was originally described by Moersch and Woltman in 1956, who reported 14 cases observed at the Mayo Clinic. Since then, more than a hundred cases of the disease have been reported in the literature. SMS, also recently referred to as Stiff-Person syndrome, is characterized by rigidity of skeletal muscles, primarily of the trunk and limbs, with superimposed painful spasm. It usually appears in adulthood and generally has a fluctuating, slowly progressive course. In a few cases, sudden death has been reported.

SMS resembles a chronic form of tetanus in many ways, although some differences, such as the absence of trismus in SMS, differentiate the two conditions. Symptoms result from the simultaneous activation of agonist and antagonist muscles. Several characteristics of the disease indicate its CNS origin.

Neurophysiological studies have demonstrated the presence of a continuous discharge of motor-unit potentials resembling a normal voluntary contraction of antagonist muscles. It was proposed that an imbalance between excitatory (catecholaminergic), and inhibitory (GABAergic) pathways controlling α-motoneuron activity causes the manifestations of the disease. High doses of various agonists of GABAergic neurotransmission, such as baclofen, sodium valproate and benzodiazepines, are generally effective in ameliorating rigidity by reducing α-motoneuron firing. These drugs, however, do not affect the course of the disease.

More recently, steroid treatment has been shown to be effective in some cases. The few autopsies carried out so far have failed to demonstrate pathological lesions of specific areas of the CNS. However, evidence for an inflammatory process in the spinal cord and brainstem has been reported in some cases (Ref. 6 and Refs listed there).

**The autoimmune hypothesis of SMS**

An autoimmune pathogenesis of SMS was suggested by the observation of sporadic cases in which this was associated with autoimmune diseases including IDDM. SMS is due to an autoimmune destruction of insulin-secreting β cells of pancreatic islets. The study of one SMS patient and subsequently of an additional 32 patients (who fitted the criteria for the diagnosis of the disease established by Gordon et al.) addressed directly the possibility that CNS autoimmunity was involved in SMS. In many of these patients, levels of IgG antibodies in the cerebrospinal fluid (CSF) were elevated or had an oligoclonal pattern or both, suggesting that IgG antibodies were produced locally within the blood–brain barrier. Furthermore, in 60% of these patients, autoantibodies directed against GABAergic neurons, and primarily their nerve terminals, were identified by an immunocytochemical assay in the serum and the available CSFs (Fig. 1). By western blotting and immunoprecipitation, the dominant autoantigen recognized by these autoantibodies in brain tissue was found to be the GABA synthesizing enzyme, glutamic acid decarboxylase (GAD) (Figs 2, 3). Sera and CSFs found to be negative by immunocytochemistry were also found to be negative for GAD antibodies by western blotting and immunoprecipitation. A few sera were found to be positive by immunocytochemistry and immunoprecipitation but negative by western blotting. This can probably be explained by the variable specificity of